Algebra & Number Theory

Volume 12

2018

No. 7

msp

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR

Bjorn Poonen
Massachusetts Institute of Technology
Cambridge, USA

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Richard E. Borcherds	University of California, Berkeley, USA
Antoine Chambert-Loir	Université Paris-Diderot, France
J-L. Colliot-Thélène	CNRS, Université Paris-Sud, France
Brian D. Conrad	Stanford University, USA
Samit Dasgupta	University of California, Santa Cruz, USA
Hélène Esnault	Freie Universität Berlin, Germany
Gavril Farkas	Humboldt Universität zu Berlin, Germany
Hubert Flenner	Ruhr-Universität, Germany
Sergey Fomin	University of Michigan, USA
Edward Frenkel	University of California, Berkeley, USA
Andrew Granville	Université de Montréal, Canada
Joseph Gubeladze	San Francisco State University, USA
Roger Heath-Brown	Oxford University, UK
Craig Huneke	University of Virginia, USA
Kiran S. Kedlaya	Univ. of California, San Diego, USA
János Kollár	Princeton University, USA
Philippe Michel	École Polytechnique Fédérale de Lausanne
Susan Montgomery	University of Southern California, USA
Shigefumi Mori	RIMS, Kyoto University, Japan
Martin Olsson	University of California, Berkeley, USA
Raman Parimala	Emory University, USA
Jonathan Pila	University of Oxford, UK
Anand Pillay	University of Notre Dame, USA
Michael Rapoport	Universität Bonn, Germany
Victor Reiner	University of Minnesota, USA
Peter Sarnak	Princeton University, USA
Joseph H. Silverman	Brown University, USA
Michael Singer	North Carolina State University, USA
Christopher Skinner	Princeton University, USA
Vasudevan Srinivas	Tata Inst. of Fund. Research, India
J. Toby Stafford	University of Michigan, USA
Pham Huu Tiep	University of Arizona, USA
Ravi Vakil	Stanford University, USA
Michel van den Bergh	Hasselt University, Belgium
Marie-France Vignéras	Université Paris VII, France
Kei-Ichi Watanabe	Nihon University, Japan
Shou-Wu Zhang	Princeton University, USA

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

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

The subscription price for 2018 is US $340/year for the electronic version, and $535/year (+$55, 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/

© 2018 Mathematical Sciences Publishers

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1559

Difference modules and difference cohomology

Marcin Chałupnik and Piotr Kowalski

We give some basics about homological algebra of difference representations. We consider both the difference discrete and the difference rational case. We define the corresponding cohomology theories and show the existence of spectral sequences relating these cohomology theories with the standard ones.

1. Introduction

In this article, we initiate a systematic study of module categories in the context of difference algebra. Our set-up is as follows. We call an object, such as a ring, a group or an affine group scheme, *difference* when it is additionally equipped with an endomorphism. Hence a difference ring is just a ring with the additional structure of a ring endomorphism. *Difference algebra* (that is, the theory of difference rings) was initiated by Ritt and developed further by Cohn [1965]. This general theory was motivated by the theory of *difference equations* (they may be considered as a discrete version of differential equations).

We introduce and investigate a suitable category of representations of difference (algebraic) groups which takes into account the extra difference structure. As far as we know, this quite natural field of research was explored only in [Kamensky 2013; Wibmer 2014]. We discuss the relation between their approach and ours in Section 5A.

We start by discussing the most general case of the category of difference modules over a difference ring in some detail (see Section 2). However, in the further part of the paper we concentrate on the theory of difference representations of a difference group and the parallel (yet more complicated) theory of difference representations of difference affine group schemes. The emphasis is put on developing the rudiments of homological algebra in these contexts, since our main motivation for studying difference representations is our idea of using difference language for comparing cohomology of affine group schemes and discrete groups. Let us now outline our program (further details can be found in Section 5B).

The basic idea is quite general. The Frobenius morphism extends to a self-transformation of the identity functor on the category of schemes over $\mathbb{F}_p$. Thus schemes over $\mathbb{F}_p$ can be naturally regarded as difference objects. We shall apply this approach to the classical problem of comparing rational and discrete cohomology of affine group schemes defined over $\mathbb{F}_p$. The main result in this area [Cline et al. 1977] establishes for a reductive algebraic group $\boldsymbol{G}$ defined over $\mathbb{F}_p$ an isomorphism between a certain

Chałupnik was supported by the Narodowe Centrum Nauki grant no. 2015/19/B/ST1/01150.
Kowalski was supported by the Narodowe Centrum Nauki grants no. 2015/19/B/ST1/01150 and 2015/19/B/ST1/01151.
MSC2010: primary 12H10; secondary 14L15, 20G05.
Keywords: rational cohomology, difference algebraic group, difference cohomology.

limit of its rational cohomology groups (called the *stable rational cohomology* of $\boldsymbol{G}$) and the discrete cohomology of the group of its $\overline{\mathbb{F}}_p$-rational points (for details, see Section 5B). The main results of our paper (Theorems 3.8 and 4.12) provide an interpretation of stable cohomology in terms of difference cohomology. Thus, the stable cohomology which was defined ad hoc as a limit is interpreted here as a genuine right derived functor in the difference framework. We hope to use this interpretation in a future work which aims to generalize the main theorem of [Cline et al. 1977] to the case of nonreductive group schemes. We also hope that this point of view together with Hrushovski's theory of generic Frobenius [2012] may lead to an independent and more conceptual proof of the main theorem of [Cline et al. 1977]. We provide more details of our program in Section 5B.

To summarize, the aim of our article is twofold. Firstly, we develop some basics of module theory and homological algebra in the difference setting. We believe that some interesting phenomena already can be observed at this stage. For example, in Remark 3.9 we point out a striking asymmetry between left and right difference modules, and in Section 5B we discuss the role of the process of inverting endomorphism. Thus we hope that our work will encourage further research in this subject. Secondly, we provide a formal framework for applying difference algebra to homological problems in algebraic geometry in the case of positive characteristic. We hope to use the tools we have worked out in the present paper in our future work exploring the relation between homological invariants of schematic and discrete objects.

The paper is organized as follows. In Section 2, we collect necessary facts about (noncommutative) difference rings. In Section 3, we deal with the difference discrete cohomology and in Section 4, we consider the difference rational cohomology. In Section 5, we compare our theory with the existing ones and with the theory of spectra from [Chałupnik 2015], and we also briefly describe another version of the notion of a difference rational representation (see Definition 5.1).

We would like to thank the referee for a careful reading of our paper and many useful suggestions.

2. Difference rings and modules

In this section, we introduce a suitable module category for difference rings. The theory of difference modules over commutative difference rings has been already considered (see, e.g., [Levin 2008, Chapter 3]), however our approach is different than the one from [Levin 2008] (we summarize the differences in Remark 2.2). We recall that a *difference ring* is a pair (R, σ), where R is a ring with a unit (not necessarily commutative), and $\sigma : R \to R$ is a ring homomorphism preserving the unit. A *homomorphism of difference rings* is a ring homomorphism commuting with the distinguished endomorphisms.

Let (R, σ) be a difference ring. We call a pair (M, σ_M) a *left difference* (R, σ)*-module* if it consists of a left R-module M with an additive map $\sigma_M : M \to M$ satisfying the condition

$$\sigma_M(\sigma(r) \cdot m) = r \cdot \sigma_M(m), \tag{†}$$

for any $r \in R$ and $m \in M$ (we explain why we choose such a condition in Remark 3.9). The condition (†) can be concisely rephrased as saying that the map

$$\sigma_M : M^{(1)} \to M$$

is a homomorphism of R-modules, where $M^{(1)}$ stands for M with the R-module structure twisted by σ, i.e., $r \cdot m := \sigma(r) \cdot m$, where $r \in R$ and $m \in M$. The left difference (R, σ)-modules form a category with the morphisms being the R-homomorphisms commuting with the fixed additive endomorphisms satisfying (†).

We have a parallel notion of a *right difference* (R, σ)*-module*. This time it is a right R-module M with an additive map $\sigma_M : M \to M$ satisfying the condition

$$\sigma_M(m \cdot r) = \sigma_M(m) \cdot \sigma(r), \tag{†′}$$

which, in terms of the induced R-modules, means that the map

$$\sigma_M : M \to M^{(1)}$$

is R-linear.

These categories can be interpreted as genuine module categories, which we explain below. We define the ring of twisted polynomial $R[\sigma]$ as follows. The underlying Abelian group is the same as in the usual polynomial ring $R[t]$. However, the multiplication is given by the formula

$$\Bigl(\sum t^i r_i\Bigr) \cdot \Bigl(\sum t^j r'_j\Bigr) := \sum_n t^n \Bigl(\sum_{i+j=n} \sigma^j(r_i) r'_j\Bigr).$$

Then we have the following.

Proposition 2.1. *The category of left and right difference* (R, σ)*-modules are equivalent* (*even isomorphic*) *to the category of left and right* $R[\sigma]$*-modules*, *respectively.*

Proof. Let M be a left difference R-module. Then we equip M with a structure of a left $R[\sigma]$-module by putting

$$\Bigl(\sum t^i r_i\Bigr) \cdot m := \sum \sigma_M^i(r_i \cdot m).$$

The condition (†) ensures that the commutativity relation in $R[\sigma]$ is respected. Conversely, for a left $R[\sigma]$-module N, we define σ_N by the formula

$$\sigma_N(n) := t \cdot n.$$

Then $\sigma_N : N \to N$ is clearly additive and satisfies (†). The proof for the right modules is similar. □

Remark 2.2. We summarize here how our definition of a difference module differs from the one in [Levin 2008].

(1) Our base ring of twisted polynomials (defined above) corresponds to the *opposite ring* to the ring of difference operators $\mathcal{D}$ considered in [Chapter 3.1]. Hence the left difference modules considered in [loc. cit.] correspond to our right difference modules.

(2) A possible notion of a *right* difference modules (which would correspond to our left difference modules, the choice on which we focus in this paper) is not considered in [loc. cit.].

We should warn the reader that the categories of left and right difference modules behave quite differently. For example, since $\sigma : R \to R^{(1)}$ may be thought of as a map of R-modules, R with $\sigma_R := \sigma$ is a right difference (R, σ)-module. If σ is an automorphism, then obviously R with $\sigma_R := \sigma^{-1}$ is a left difference (R, σ)-module. However, in the general case we do not have any natural structure of a left difference (R, σ)-module on R. Since in this paper we are mainly interested in left difference (R, σ)-modules (a technical explanation is provided in Remark 3.9), we would like to construct a left difference (R, σ)-module possibly closest to R. We achieve this goal by formally inverting the action of σ on R.

Definition 2.3. Let

$$\mathrm{R}_{1-t} : R[\sigma] \to R[\sigma]$$

be the right multiplication by $(1 - t)$. This is clearly a map of left $R[\sigma]$-modules and we define the following left $R[\sigma]$-module:

$$\tilde{R} := \mathrm{coker}(\mathrm{R}_{1-t}).$$

Our construction has the following properties.

Proposition 2.4. *Let $\sigma_{\tilde{R}}$ be the map provided by Proposition 2.1. Then we have the following*:

(1) *The map $\sigma_{\tilde{R}}$ is invertible.*

(2) *If σ is an automorphism, then*

$$(\tilde{R}, \sigma_{\tilde{R}}) \simeq (R, \sigma^{-1}).$$

Proof. Since we have the following relation in $\tilde{R}$:

$$\sum_{i=0}^{n} t^i r_i = \sum_{i=0}^{n} t^{i+1} \sigma(r_i),$$

we see that the map $\sum t^i r_i \mapsto \sum t^i \sigma(r_i)$ is the inverse of $\sigma_{\tilde{R}}$.

For the second part, we observe first that the map

$$\alpha : (R, \sigma^{-1}) \to \tilde{R},$$

given by the formula $\alpha(r) := r$, is a homomorphism of left $R[\sigma]$-modules, since the relation $\sigma^{-1}(r) = tr$ holds in $\tilde{R}$. Also, the map

$$\beta : \tilde{R} \to (R, \sigma^{-1})$$

given by

$$\beta\Big(\sum t^i r_i\Big) := \sum \sigma^{-i}(r_i)$$

is a homomorphism of left $R[\sigma]$-modules. We see now that α and β are mutually inverse. $\square$

From now on, we focus exclusively on left (difference) modules, hence we denote by Mod_R^σ the category of left difference (R,σ)-modules (or the equivalent category of left $R[\sigma]$-modules). Also, if it causes no confusion we will not refer to endomorphisms in our notation, i.e., we will usually say "M is a left difference R-module" (or even "M is a difference R-module") instead of saying "(M,σ_M) is a left difference (R,σ)-module".

We finish this section with an elementary homological computation, which explains (roughly speaking) the effect of adding a difference structure on homology. We will make this point more precise in the next section.

For a difference R-module M, let M^{σ_M} and M_{σ_M} stand for the Abelian groups of invariants and coinvariants of the action of σ_M, respectively. Explicitly, we have

$$M^{\sigma_M} = \{m \in M \mid \sigma_M(m) = m\} \quad \text{and} \quad M_{\sigma_M} = M/\langle \sigma_M(m) - m \mid m \in M\rangle.$$

Then we have the following.

Proposition 2.5. *For a difference R-module M, we have*

$$\begin{aligned}\mathrm{Hom}_{\mathrm{Mod}_R^\sigma}(\tilde{R}, M) &= M^{\sigma_M},\\ \mathrm{Ext}^1_{\mathrm{Mod}_R^\sigma}(\tilde{R}, M) &= M_{\sigma_M},\\ \mathrm{Ext}^{>1}_{\mathrm{Mod}_R^\sigma}(\tilde{R}, M) &= 0.\end{aligned}$$

Proof. Since the map R_{1-t} is injective, the complex

$$0 \to R[\sigma] \xrightarrow{\mathrm{R}_{1-t}} R[\sigma] \to 0$$

is a free resolution of $\tilde{R}$. Then the complex of Abelian groups

$$0 \to \mathrm{Hom}_{\mathrm{Mod}_R^\sigma}(R[\sigma], M) \xrightarrow{(\mathrm{R}_{1-t})^*} \mathrm{Hom}_{\mathrm{Mod}_R^\sigma}(R[\sigma], M) \to 0,$$

which computes our Ext-groups, may be identified with the complex

$$0 \to M \xrightarrow{\mathrm{L}_{1-t}} M \to 0$$

where L_{1-t} stands for the left multiplication by the element $(1-t)$. Thus, the proposition follows. □

3. Difference representations and cohomology

Let (A,σ_A) be a difference commutative ring and G be a group with an endomorphism σ_G. In this section, we apply the results of Section 2 to the ring $R := A[G]$, the group ring of G with coefficients in A. The ring R with the map

$$\sigma\left(\sum a_i g_i\right) := \sum \sigma_A(a_i)\sigma_G(g_i)$$

is clearly a difference ring. We will often say "difference representation of G (over A)" for "difference $A[G]$-module". We observe now that the augmentation map $\epsilon : A[G] \to A$ is a homomorphism of difference rings (by this we mean a ring homomorphism commuting with σ and σ_A). Hence, we can endow the left difference A-module $\tilde{A}$ (see Definition 2.3) with the "trivial" structure of a left difference $A[G]$-module, i.e., we put

$$\left(\sum a_i g_i\right) \cdot a := \sum a_i \cdot a.$$

Remark 3.1. We would like to warn the reader that in contrast to the classical representation theory, difference representations (M, σ_M) correspond to homomorphisms into the group $\mathrm{GL}_A(M)$ only if σ_M is an automorphism. More precisely, if (M, σ_M) is a difference A-module and σ_M is an automorphism, then we have the automorphism $\widetilde{\sigma_M}$ on $\mathrm{GL}_A(M)$ given by the conjugation:

$$\widetilde{\sigma_M}(\alpha) := \sigma_M^{-1} \circ \alpha \circ \sigma_M.$$

It is easy to see then that endowing (M, σ_M) with the structure of a difference $A[G]$-module is the same as constructing a homomorphism of difference groups

$$\Phi : (G, \sigma_G) \to (\mathrm{GL}_A(M), \widetilde{\sigma_M}).$$

We are ready now to define the notion of a difference group cohomology.

Definition 3.2. Let M be a difference $A[G]$-module. We define:

$$H^j_\sigma(G, M) := \mathrm{Ext}^j_{\mathrm{Mod}^\sigma_R}(\tilde{A}, M).$$

We show below that the zeroth difference cohomology can be described in terms of invariants.

Proposition 3.3. *For any difference $A[G]$-module M, we have*

$$H^0_\sigma(G, M) = M^G \cap M^{\sigma_M}.$$

Proof. We observe first that by the (†)-condition from Section 2, the A-module M^G is preserved by σ_M. Indeed, for any $m \in M^G$ we have:

$$g \cdot (\sigma_M(m)) = \sigma_M(\sigma_G(g) \cdot m) = \sigma_M(m).$$

Thus M^G is a difference A-module and, since G acts on $\tilde{A}$ trivially, we have

$$\mathrm{Hom}_{\mathrm{Mod}^\sigma_{A[G]}}(\tilde{A}, M) = \mathrm{Hom}_{\mathrm{Mod}^\sigma_A}(\tilde{A}, M^G).$$

By Proposition 2.5, we obtain

$$\mathrm{Hom}_{\mathrm{Mod}^\sigma_A}(\tilde{A}, M^G) = (M^G)^{\sigma_M} = M^G \cap M^{\sigma_M},$$

which completes the proof. □

This description shows possibility of factoring the difference cohomology functor as the composite of two left exact functors. To make this precise, let us consider the chain of left exact functors

$$\mathrm{Mod}^\sigma_{A[G]} \xrightarrow{K} \mathrm{Mod}^\sigma_A \xrightarrow{L} \mathrm{Mod}_A,$$

where

$$K(M) := \mathrm{Hom}_{\mathrm{Mod}_{A[G]}}(A, M) = M^G \quad \text{and} \quad L(N) := \mathrm{Hom}_{\mathrm{Mod}^{\sigma}_{A}}(\tilde{A}, N) = N^{\sigma_N}.$$

We recall here the fact observed in the proof of Proposition 3.3 that the target category of K is indeed the category $\mathrm{Mod}^{\sigma}_{A}$. Now, Proposition 3.3 can be understood as the following factorization

$$H^0_{\sigma}(G, -) = L \circ K.$$

We would like now to associate the Grothendieck spectral sequence to the above factorization. To achieve this, we need the following fact.

Lemma 3.4. *The functor* $\epsilon^* : \mathrm{Mod}^{\sigma}_{A} \to \mathrm{Mod}^{\sigma}_{A[G]}$ *is left adjoint to* K. *Consequently*, *the functor* K *preserves injectives.*

Proof. The desired adjunction is a natural isomorphism

$$\mathrm{Hom}_{\mathrm{Mod}^{\sigma}_{A[G]}}(\epsilon^*(N), M) \simeq \mathrm{Hom}_{\mathrm{Mod}^{\sigma}_{A}}(N, M^G),$$

which immediately follows from the fact that G acts trivially on $\epsilon^*(N)$. Thus K has an exact left adjoint functor, hence it preserves injectives. □

The description of the functor K above also shows that for any difference $A[G]$-module M, each $H^j(G, M)$ has a natural structure of a difference A-module. The endomorphism of $H^j(G, M)$ can be explicitly described as the composite of the following two arrows:

$$H^j(G, M) \xrightarrow{\sigma_G^*} H^j(G, M^{(1)}) \xrightarrow{(\sigma_M)_*} H^j(G, M),$$

where the first one is the restriction map along σ_G [Weibel 1994, Chapter 6.8], and the second one is the map induced by the G-invariant map $\sigma_M : M^{(1)} \to M$.

Then we have the following result, where the invariants and the coinvariants are taken with respect to the difference structure which was just described.

Theorem 3.5. *For any difference* $A[G]$*-module* M *and* $j \geqslant 0$, *there is a short exact sequence* (*setting* $H^{-1}(G, M) := 0$)

$$0 \to H^{j-1}(G, M)_{\sigma} \to H^j_{\sigma}(G, M) \to H^j(G, M)^{\sigma} \to 0.$$

Proof. Since L, K are left exact functors and K takes injective objects to L-acyclic ones by Lemma 3.4, we can construct the Grothendieck spectral sequence (see e.g., [Weibel 1994, Chapter 5.8]) associated to the composite functor $L \circ K$. This spectral sequence converges to $H^{p+q}_{\sigma}(G, M)$, and its second page has the following form:

$$E_2^{pq} = \mathrm{Ext}^p_{\mathrm{Mod}^{\sigma}_{A}}(\tilde{A}, H^q(G, M)).$$

By Proposition 2.5, there are only two nontrivial columns in this page where we have

$$E_2^{0j} = H^j(G, M)^{\sigma} \quad \text{and} \quad E_2^{1j} = H^j(G, M)_{\sigma}.$$

Thus all the differentials vanish and we get the result. □

The above theorem is an efficient tool for computations of difference cohomology groups. Let us look at some simple examples.

Example 3.6. Let $G = \mathbb{Z}/p$ be the cyclic group of prime order $p > 2$ with an automorphism σ_G given by the formula $\sigma_G(a) := ta$ for some integer t such that $0 < t < p$. Let r be the order of t in the multiplicative group of the field $\mathbb{F}_p$ and let further $A = \boldsymbol{k}$ be a field of characteristic p.

(1) Let us take $\sigma_A = \mathrm{id}$. We would like to compute

$$H^*_\sigma(\mathbb{Z}/p, \boldsymbol{k}) := \bigoplus_{n=0}^{\infty} H^n_\sigma(\mathbb{Z}/p, \boldsymbol{k})$$

for $(\boldsymbol{k}, \mathrm{id})$ regarded as the trivial difference $\boldsymbol{k}[G]$-module. In order to apply Theorem 3.5, we need to explicitly describe the endomorphism of $H^*(\mathbb{Z}/p, \boldsymbol{k})$, let us call it σ_{H^*}, which comes from the difference structure. When M is a trivial G-module, we have $H^1(G, M) = \mathrm{Hom}_{\mathrm{Ab}}(G, M)$ and we obtain

$$\sigma_{H^1}(\phi) = \sigma_M \circ \phi \circ \sigma_G.$$

Coming back to our example, let us fix a nonzero $y \in H^1(\mathbb{Z}/p, \boldsymbol{k})$ and let $x \in H^2(\mathbb{Z}/p, \boldsymbol{k})$ be the image of y by the Bockstein homomorphism. It is well known (see e.g., [Weibel 1994, Exercise 6.7.5]) that we have a ring isomorphism

$$H^*(\mathbb{Z}/p, \boldsymbol{k}) = S(\boldsymbol{k}x) \otimes \Lambda(\boldsymbol{k}y),$$

where $S(M)$ is the symmetric power and $\Lambda(M)$ is the exterior power of a $\boldsymbol{k}$-module M. Thus we see that $\sigma_{H^1}(y) = ty$ and, by the naturality of the Bockstein homomorphism, also $\sigma_{H^2}(x) = tx$. Therefore, by the naturality of the multiplicative structure on group cohomology, for all $j > 0$ we obtain

$$\sigma_{H^{2j}}(x^j) = t^j x^j, \sigma_{H^{2j-1}}(x^{j-1} \otimes y) = t^j (x^{j-1} \otimes y). \tag{$\star$}$$

Hence we see that $H^{2j}(\mathbb{Z}/p, \boldsymbol{k})^\sigma = \boldsymbol{k}x^j$ if and only if $r \mid j$ (recall that r is the multiplicative order of t), and $H^{2j}(\mathbb{Z}/p, \boldsymbol{k})^\sigma = 0$ otherwise. A similar conclusion holds for $H^{2j-1}(\mathbb{Z}/p, \boldsymbol{k})^\sigma$, $H^{2j}(\mathbb{Z}/p, \boldsymbol{k})_\sigma$ and $H^{2j-1}(\mathbb{Z}/p, \boldsymbol{k})_\sigma$. Applying Theorem 3.5, we get that $H^0_\sigma(\mathbb{Z}/p, \boldsymbol{k}) = \boldsymbol{k}$ and, for $n > 0$, we obtain the following

$$H^n_\sigma(\mathbb{Z}/p, \boldsymbol{k}) = \begin{cases} \boldsymbol{k} \oplus \boldsymbol{k} & \text{for } 2r \mid n, \\ \boldsymbol{k} & \text{for } 2r \mid n-1, \\ \boldsymbol{k} & \text{for } 2r \mid n+1, \\ 0 & \text{otherwise.} \end{cases}$$

(2) Let us now elaborate on the above example by adding an automorphism of scalars to the picture. Hence, let F be an automorphism of $\boldsymbol{k}$. Then $(\boldsymbol{k}, F^{-1})$ is a difference $(\boldsymbol{k}, F)[G]$-module and we are interested in its difference cohomology. We recall that $H^1(\mathbb{Z}/p, \boldsymbol{k}) = \mathrm{Hom}_{\mathrm{Ab}}(\mathbb{Z}/p, \boldsymbol{k})$, which is naturally identified with $\boldsymbol{k}$. After choosing $y \in \mathbb{F}_p$, we get the same formulas as in ($\star$) from the item (1) above.

Since each $H^n(\mathbb{Z}/p, \boldsymbol{k})$ is a difference $(\boldsymbol{k}, F)$-module, for $c \in \boldsymbol{k}$ we obtain the following

$$\sigma_{H^{2j}}(cx^j) = F^{-1}(c)t^j x^j,$$
$$\sigma_{H^{2j-1}}(cx^{j-1} \otimes y) = F^{-1}(c)t^j (x^{j-1} \otimes y).$$

For $a \in \mathbb{F}_p \setminus \{0\}$, let $\boldsymbol{k}^a$ stand for the eigenspace of F regarded as an $\mathbb{F}_p$-linear automorphism of $\boldsymbol{k}$ for the eigenvalue a. Dually, let $\boldsymbol{k}_a$ be the corresponding "coeigenspace", i.e., the quotient $\mathbb{F}_p$-linear space

$$\boldsymbol{k}_a = \boldsymbol{k}/\langle F(c) - ca \mid c \in \boldsymbol{k}\rangle.$$

Therefore, for any nonnegative integer j, we get by Theorem 3.5

$$H^{2j}_\sigma(\mathbb{Z}/p, \boldsymbol{k}) = \boldsymbol{k}^{t^j} \oplus \boldsymbol{k}_{t^j}, \quad H^{2j+1}_\sigma(\mathbb{Z}/p, \boldsymbol{k}) = \boldsymbol{k}^{t^{j+1}} \oplus \boldsymbol{k}_{t^j}.$$

(3) If we consider a special case of the situation considered in the item (2) above, where $A = \boldsymbol{k} = \mathbb{F}_p^{\mathrm{alg}}$ and $\sigma_A = \mathrm{Fr}_{\boldsymbol{k}}$ is the Frobenius map, then by the results of [Kowalski and Pillay 2007, §3], the difference module $H^*(\mathbb{Z}/p, \boldsymbol{k})$ is *σ-isotrivial*, i.e., we have the following isomorphism of difference modules

$$H^*(\mathbb{Z}/p, \boldsymbol{k}) \simeq (\boldsymbol{k}, \mathrm{Fr}_{\boldsymbol{k}}^{-1}) \otimes_{(\mathbb{F}_p, \mathrm{id})} (H^*(\mathbb{Z}/p, \boldsymbol{k})^\sigma, \mathrm{id}).$$

(To apply [Kowalski and Pillay 2007, Fact 3.4(ii)], we need to know that σ_{H^*} is a bijection, but it is the case since both σ_G and F are automorphisms.) Since $\boldsymbol{k}^{\mathrm{Fr}} = \mathbb{F}_p$, $\boldsymbol{k}_{\mathrm{Fr}} = 0$ and each $H^n(\mathbb{Z}/p, \boldsymbol{k})^\sigma$ is a 1-dimensional vector space over $\mathbb{F}_p$, we immediately (i.e., using neither the item (1) nor the item (2) above) get (by Theorem 3.5) the following isomorphism of $\mathbb{F}_p$-linear spaces:

$$H^*_\sigma(\mathbb{Z}/p, \boldsymbol{k}) \simeq S(\mathbb{F}_p x) \otimes \Lambda(\mathbb{F}_p y) = H^*(\mathbb{Z}/p, \mathbb{F}_p),$$

which coincides with the computations made in the item (2).

For a left $A[G]$-module M, let us denote by M^∞ the induced difference $A[G]$-module, i.e.,

$$M^\infty := A[G][\sigma] \otimes_{A[G]} M.$$

In order to describe M^∞ more explicitly, we slightly extend the notation introduced in Section 2, by setting $M^{(i)}$ to be the $A[G]$-module M with the structure twisted by σ^i. Then, we have an isomorphism of $A[G]$-modules

$$M^\infty \simeq \bigoplus_{i \geqslant 0} M^{(i)}.$$

Under this identification, the difference structure on M^∞ is given by the following shift:

$$\sigma_{M^\infty}(m_0, \dots, m_i, 0, \dots) = (0, m_0, \dots, m_i, 0, \dots).$$

Let us now investigate the exact sequence from Theorem 3.5 for the difference module M^∞. For this, we introduce the "stable cohomology groups" as

$$H^j_{\mathrm{st}}(G, M) := \operatorname*{colim}_i H^j(G, M^{(i)}),$$

where the maps in the direct system are the restriction maps along σ_G.

Remark 3.7. We give an interpretation of the stable cohomology in small dimensions.

(1) The zeroth stable cohomology group

$$H^0_{\mathrm{st}}(G, N) = \bigcup_{n=1}^{\infty} N^{\mathrm{Im}(\sigma_G^n)}$$

may be thought of as the group of "weak invariants" of the action of G on N.

(2) Suppose that N is a trivial G-module. Then we have

$$H^1_{\mathrm{st}}(G, N) := \mathrm{colim}(\mathrm{Hom}(G, N) \to \mathrm{Hom}(G, N) \to \cdots),$$

where the map producing the direct system is induced by σ_G. Hence $H^1_{\mathrm{st}}(G, N)$ can be considered as the effect of inverting formally the above endomorphism on $\mathrm{Hom}(G, N)$.

These stable cohomology groups play an important role in the comparison between rational and discrete cohomology in [Cline et al. 1977]. The fact that, as we will see in a moment, they appear as difference cohomology groups is one of the main motivations for the present work. Namely, when we explicitly describe the action of σ on

$$H^*(G, M^\infty) \simeq H^*\Big(G, \bigoplus_{i\geqslant 0} M^{(i)}\Big) \simeq \bigoplus_{i\geqslant 0} H^*(G, M^{(i)}),$$

we obtain that (after restricting to the summand $H^*(G, M^{(i)})$) this action is given by the map

$$\sigma_* : H^*(G, M^{(i)}) \to H^*(G, M^{(i+1)})$$

induced by σ on the cohomology. Thus we see that $H^*(G, M^\infty)^\sigma = 0$, and using Theorem 3.5 we get the following.

Theorem 3.8. *For any $A[G]$-module M and $j > 0$, there is an isomorphism*

$$H^j_\sigma(G, M^\infty) \simeq H^{j-1}_{\mathrm{st}}(G, M).$$

Remark 3.9. Apparently, there is no similar description of the stable cohomology in terms of cohomology of right difference modules. The technical obstacle for this is the fact that for a right difference $A[G]$-module M, the module of invariants M^G are not preserved by σ_M. Therefore, there is no Grothendieck spectral sequence analogous to the one which we used in the proof of Theorem 3.5. This is the main reason we have chosen to work with left difference modules in this paper, despite the fact that condition (†′) looks more natural than condition (†) (both of which can be found before Proposition 2.1).

4. Difference rational representations and cohomology

In this section, we introduce difference rational modules and difference rational cohomology. As rational representations and rational cohomology concern representations of algebraic groups, we will consider here representations of *difference algebraic groups*, so we recall this notion first. Let $\boldsymbol{k}$ be our ground field.

4A. ***Difference algebraic groups.*** We take the categorical definition of a difference algebraic group appearing in [Wibmer 2014]. When we say "algebraic group", we mean "affine group scheme". We do not care here about the finite-generation (or finite type) issues: neither in the schematic nor in the difference-schematic meaning. We comment about other possible approaches in Section 5C.

Let $\sigma : \boldsymbol{k} \to \boldsymbol{k}$ be a field homomorphism. The category of difference $(\boldsymbol{k}, \sigma)$-algebras (denoted here by $\mathrm{Alg}_{(\boldsymbol{k},\sigma)}$) consists of commutative $\boldsymbol{k}$-algebras A equipped with ring endomorphisms σ_A such that $(\sigma_A)|_{\boldsymbol{k}} = \sigma$. A morphism between two $(\boldsymbol{k}, \sigma)$-algebras (A_1, σ_1) and (A_2, σ_2) is a $\boldsymbol{k}$-algebra morphism $f : A_1 \to A_2$ such that

$$\sigma_2 \circ f = f \circ \sigma_1.$$

An *affine difference algebraic group* is defined as a representable functor from the category $\mathrm{Alg}_{(\boldsymbol{k},\sigma)}$ to the category of groups. Note that it is in an exact analogy with the pure algebraic case. Such a functor is represented by a *difference Hopf algebra* which may be defined as (H, σ_H), where H is a Hopf algebra over $\boldsymbol{k}$, $\sigma^*(H)$ is obtained from H using the base extension $\sigma : \boldsymbol{k} \to \boldsymbol{k}$ (i.e., $\sigma^*(H) = H \otimes_{\boldsymbol{k}} (\boldsymbol{k}, \sigma)$) and $\sigma_H : \sigma^*(H) \to H$ is a Hopf algebra morphism [Wibmer 2014, Definition 2.2]. Dualizing, we see that a difference algebraic group $\mathcal{G}$ is the same as a pair $(\boldsymbol{G}, \sigma_{\boldsymbol{G}})$ where $\boldsymbol{G}$ is an affine group scheme over $\boldsymbol{k}$ and $\sigma_{\boldsymbol{G}} : \boldsymbol{G} \to \sigma^*(\boldsymbol{G})$ is a group scheme morphism, where $\sigma^*(\boldsymbol{G})$ is again obtained from $\boldsymbol{G}$ using the base extension $\sigma : \boldsymbol{k} \to \boldsymbol{k}$.

Difference algebraic groups appeared first in the context of model theory (of difference fields) and yielded important applications to number theory (related to the Manin–Mumford conjecture) and algebraic dynamics, see e.g., [Chatzidakis and Hrushovski 2008a; 2008b; Hrushovski 2001; Medvedev and Scanlon 2014; Kowalski and Pillay 2007]. Difference algebraic groups also appear as the Galois groups of certain linear differential equations [Di Vizio et al. 2014] and linear difference equations [Ovchinnikov and Wibmer 2015].

We are mostly interested in the case when $\boldsymbol{G}$ is defined over the field of constants of σ (see Section 5B). In such a case, one can replace the difference field $(\boldsymbol{k}, \sigma)$ with the difference field $(\mathrm{Fix}(\sigma), \mathrm{id})$. Therefore, in the rest of Section 4, we assume that $\sigma = \mathrm{id}_{\boldsymbol{k}}$. In Section 5C, we discuss our attempts to define a more general notion of a difference rational representation, which covers the case of an arbitrary base difference field $(\boldsymbol{k}, \sigma)$ (see also Remark 4.4).

4B. ***Difference rational representations.*** Let $\boldsymbol{G}$ be a $\boldsymbol{k}$-affine group scheme with an endomorphism $\sigma_{\boldsymbol{G}}$. Its representing ring $\boldsymbol{k}[\boldsymbol{G}]$ is a Hopf algebra over $\boldsymbol{k}$ with a $\boldsymbol{k}$-Hopf algebra endomorphism, denoted here by the same symbol $\sigma_{\boldsymbol{G}}$. We would like to introduce the notion of a difference rational $\boldsymbol{G}$-module. We recall from classical algebraic geometry [Jantzen 2003] that for a $\boldsymbol{k}$-affine group scheme $\boldsymbol{G}$, a left rational $\boldsymbol{G}$-module (or a rational representation of $\boldsymbol{G}$) is a functor

$$M : \mathrm{Alg}_{\boldsymbol{k}} \to \mathrm{Mod}_{\boldsymbol{k}}$$

such that for any $\boldsymbol{k}$-algebra A, we have $M(A) = M(\boldsymbol{k}) \otimes A$, and each $M(A)$ is equipped with a natural (in $A \in \mathrm{Alg}_{\boldsymbol{k}}$) left action of the group $\boldsymbol{G}(A)$ through A-linear transformations. The left rational $\boldsymbol{G}$-modules with the morphisms being the natural transformations form the Abelian category $\mathrm{Mod}_{\boldsymbol{G}}$. Given $M \in \mathrm{Mod}_{\boldsymbol{G}}$, one can construct a natural structure of a *right* $\boldsymbol{k}[G]$-comodule on $M(\boldsymbol{k})$. The assignment $M \mapsto M(\boldsymbol{k})$ gives an equivalence between the category $\mathrm{Mod}_{\boldsymbol{G}}$ and the category of right $\boldsymbol{k}[\boldsymbol{G}]$-comodules [Jantzen 2003, §I.2.8]. The inverse is explicitly given by the following construction. An element

$$g \in \boldsymbol{G}(A) = \mathrm{Hom}_{\mathrm{Alg}_{\boldsymbol{k}}}(\boldsymbol{k}[\boldsymbol{G}], A)$$

acts on $M(A) = M(\boldsymbol{k}) \otimes A$ by the composite

$$(\mathrm{id} \otimes m) \circ (\mathrm{id} \otimes g \otimes \mathrm{id}) \circ (\Delta_M \otimes \mathrm{id}),$$

where

$$\Delta_M : M(\boldsymbol{k}) \to M(\boldsymbol{k}) \otimes \boldsymbol{k}[\boldsymbol{G}]$$

is the comodule map on $M(\boldsymbol{k})$, and m is the multiplication on A. From now on, if no confusion can arise, we will identify M with $M(\boldsymbol{k})$.

Let us come back to the situation when $\boldsymbol{G}$ is additionally equipped with an endomorphism $\sigma_{\boldsymbol{G}}$. A natural adaptation of the concept of a difference representation to the context of difference algebraic groups is the following.

Definition 4.1. A *difference rational representation* of a difference group $(\boldsymbol{G}, \sigma_{\boldsymbol{G}})$ is a pair (M, σ_M) consisting of a left rational $\boldsymbol{G}$-module M and a natural transformation $\sigma_M : M \to M$ such that for each $A \in \mathrm{Alg}_{\boldsymbol{k}}$, the A-module $M(A)$ becomes a left difference $A[\boldsymbol{G}(A)]$-module with $\sigma_{M(A)}$ being $\sigma_M(A)$, and $\sigma_{A[\boldsymbol{G}(A)]}$ is given by the following formula:

$$\sigma_{A[\boldsymbol{G}(A)]}\Big(\sum a_i g_i\Big) := \sum a_i \sigma_{\boldsymbol{G}(A)}(g_i).$$

Let (M, σ_M) and (N, σ_N) be rational difference $(\boldsymbol{G}, \sigma_{\boldsymbol{G}})$-modules. We call a transformation of functors $f : M \to N$ a *difference $\boldsymbol{G}$-homomorphism*, if for any $\boldsymbol{k}$-algebra A,

$$f(A) : M(A) \to N(A)$$

is a homomorphism of difference $A[\boldsymbol{G}(A)]$-modules.

Similarly as in Section 3, we will often skip the endomorphisms from the notation and simply say that M is a difference rational representation of $\boldsymbol{G}$. The difference rational representations of $\boldsymbol{G}$ with difference $\boldsymbol{G}$-homomorphisms obviously form a category, which we denote by $\mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$.

Remark 4.2. We can find a similar interpretation of our difference rational representations as the one in Remark 3.1. We consider $\mathrm{GL}(M)$ as a $\boldsymbol{k}$-group functor, see [Jantzen 2003, §I.2.2]. In the case when $\sigma_M : M \to M$ is a $\boldsymbol{k}$-linear automorphism, it induces the inner automorphism of this $\boldsymbol{k}$-group functor:

$$\sigma_{\mathrm{GL}(M)} : \mathrm{GL}(M) \to \mathrm{GL}(M).$$

Then enhancing (M, σ_M) with the structure of a $(\boldsymbol{G}, \sigma_{\boldsymbol{G}})$-module is the same as giving a morphism of difference $\boldsymbol{k}$-group functors as below:

$$(\boldsymbol{G}, \sigma_{\boldsymbol{G}}) \to (\mathrm{GL}(M), \sigma_{\mathrm{GL}(M)}).$$

Keeping in mind the results of Section 3 and the case of rational representations, we obtain two equivalent descriptions of the category $\mathrm{Mod}^{\sigma}_{\boldsymbol{G}}$. Analogously as in Section 2, for a rational $\boldsymbol{G}$-module M, we denote by $M^{(1)}$ the $\boldsymbol{G}$-module structure on M twisted by $\sigma_{\boldsymbol{G}}$. If we take the comodule point of view, then the comodule map on $M^{(1)}$ is given by the following composite:

$$(\mathrm{id} \otimes \sigma_{\boldsymbol{G}}) \circ \Delta_M : M^{(1)} \to M^{(1)} \otimes \boldsymbol{k}[\boldsymbol{G}].$$

Then we have the following.

Proposition 4.3. *Let $\boldsymbol{G}$ be an affine difference group scheme. Then the following categories are equivalent*:

(1) *The category* $\mathrm{Mod}^{\sigma}_{\boldsymbol{G}}$.

(2) *The category of pairs (M, σ_M), where M is a rational $\boldsymbol{G}$-module and $\sigma_M : M^{(1)} \to M$ is a $\boldsymbol{G}$-homomorphism.*

(3) *The category of pairs (M, σ_M), where M is a right $\boldsymbol{k}[\boldsymbol{G}]$-comodule and $\sigma_M : M \to M$ is a $\boldsymbol{k}$-linear map satisfying the following identity*:

$$\Delta_M \circ \sigma_M = (\sigma_M \otimes \sigma_{\boldsymbol{G}}) \circ \Delta_M. \tag{$*$}$$

Remark 4.4. A difference rational representation is a natural (in $A \in \mathrm{Alg}_{\boldsymbol{k}}$) collection of difference $A[\boldsymbol{G}(A)]$-modules. Hence we see that we work in a less general context than the one in Section 3, since we have no endomorphism on A and neither on $\boldsymbol{k}$. It would be tempting to introduce difference rational representations as functors on the category of difference algebras over $\boldsymbol{k}$ or even over a difference field $(\boldsymbol{k}, \sigma)$. The resulting category is much more complicated, e.g., we have not even succeeded yet in showing that it is Abelian. Since the simpler approach in this section is sufficient for homological applications we have in mind, we decided to stick to it in this paper. We discuss possible generalizations of difference representation theory and its relations with the other approaches in Section 5.

Example 4.5. We point out here three important examples of difference rational $\boldsymbol{G}$-modules:

(1) The *trivial difference $\boldsymbol{G}$-module*. Clearly, the $\boldsymbol{k}$-algebra unit map $\boldsymbol{k} \to \boldsymbol{k}[\boldsymbol{G}]$ endows $(\boldsymbol{k}, \mathrm{id})$ with the structure of a difference rational $\boldsymbol{G}$-module.

(2) The *regular difference $\boldsymbol{G}$-module* is defined as follows. We put

$$M := \boldsymbol{k}[\boldsymbol{G}], \quad \sigma_M := \sigma_{\boldsymbol{G}}.$$

Then the condition $(*)$ in Proposition 4.3(3) is satisfied, since $\sigma_{\boldsymbol{G}}$ is a homomorphism of coalgebras.

(3) The last example corresponds to the induced module $\boldsymbol{k}[G][\sigma] \otimes_{\boldsymbol{k}[G]} M$ from Section 3. It could be described in terms of cotensor product, but we prefer the following explicit description. For a rational $\boldsymbol{G}$-module M, we set

$$M^{\infty} := \bigoplus_{i=0}^{\infty} M^{(i)}$$

as a rational $\boldsymbol{G}$-module. Since $(M^{\infty})^{(1)} = \bigoplus_{i=1}^{\infty} M^{(i)}$, the inclusion map

$$\bigoplus_{i=1}^{\infty} M^{(i)} \subset \bigoplus_{i=0}^{\infty} M^{(i)}$$

defines the structure of a difference rational $\boldsymbol{G}$-module on M^{∞}. Note that this inclusion map is the same as the "right-shift" map appearing before Remark 3.7.

In certain simple cases, the category $\mathrm{Mod}^{\sigma}_{\boldsymbol{G}}$ can be fully described. The following example should be thought of as the first step towards understanding difference rational representations of reductive groups with the Frobenius endomorphism.

Let $\boldsymbol{k}$ be a field of positive characteristic p, $\mathbb{G}_{\mathrm{m}}$ be the multiplicative group over $\boldsymbol{k}$ and $\mathrm{Fr} : \mathbb{G}_{\mathrm{m}} \to \mathbb{G}_{\mathrm{m}}$ be the (relative) Frobenius morphism. Then the category $\mathrm{Mod}^{\sigma}_{\mathbb{G}_{\mathrm{m}}}$ can be explicitly described. Let $\mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x]}$ denote the category of $\mathbb{Z}$-graded $\boldsymbol{k}[x]$-modules satisfying the following condition (for each $j \in \mathbb{Z}$):

$$xM^{j} \subseteq M^{pj}.$$

We set $X := (\mathbb{Z} \setminus p\mathbb{Z}) \cup \{0\}$, and for $j \in X$, we define $\mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x],j}$ as the full subcategory of the category $\mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x]}$ consisting of modules concentrated in the degrees of the form $p^n j$ for $n \in \mathbb{N}$. Then we have the following:

Proposition 4.6. *The category* $\mathrm{Mod}^{\sigma}_{\mathbb{G}_{\mathrm{m}}}$ *admits the following description*:

(1) *There is an equivalence of categories*

$$\mathrm{Mod}^{\sigma}_{\mathbb{G}_{\mathrm{m}}} \simeq \mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x]}.$$

(2) *There is a decomposition into infinite product*

$$\mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x]} \simeq \prod_{j \in X} \mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x],j}.$$

(3) *The category* $\mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x],0}$ *is equivalent to the category of* $\boldsymbol{k}[x]$*-modules, while the category* $\mathrm{Mod}^{\mathbb{Z},p}_{\boldsymbol{k}[x],j}$ *for* $j \neq 0$ *is equivalent to the category of* $\mathbb{N}$*-graded modules over the graded* $\boldsymbol{k}$*-algebra* $\boldsymbol{k}[x]$, *where* $|x| = 1$.

Proof. Since $\mathbb{G}_{\mathrm{m}} = \mathrm{Diag}(\mathbb{Z})$, we can use the results from [Jantzen 2003, §I.2.11]. For $M \in \mathrm{Mod}^{\sigma}_{\mathbb{G}_{\mathrm{m}}}$, we take a decomposition of the rational module $M \simeq \bigoplus M_j$ into isotypical rational representations of $\mathbb{G}_{\mathrm{m}}$, i.e., each M_j is a direct sum of equivalent irreducible representations such that for each $A \in \mathrm{Alg}_{\boldsymbol{k}}$, $a \in \mathbb{G}_{\mathrm{m}}(A)$ and $m \in M_j(A)$, we have

$$a \cdot m = a^j m.$$

Then, since $(M_j)^{(1)} = (M^{(1)})_{pj}$, we have $\sigma_M(M_j) \subseteq M_{pj}$. This turns M into an object of the category $\mathrm{Mod}_{\boldsymbol{k}[x]}^{\mathbb{Z},p}$. The rest is straightforward. □

4C. *Difference rational cohomology.* We would like to develop now some homological algebra in the category $\mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$. Firstly, it is obvious that $\mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$ is an Abelian category with the kernels and cokernels inherited from the category $\mathrm{Mod}_{\boldsymbol{G}}$. However, the existence of enough injectives is not a priori obvious. We shall construct injective objects in the category $\mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$ by using a particular case of induction. Let (M, σ_M) be a $\boldsymbol{k}$-linear vector space with an endomorphism. Then, $M \otimes \boldsymbol{k}[\boldsymbol{G}]$ with the comodule map $\mathrm{id} \otimes \Delta_G$ and the endomorphism $\sigma_M \otimes \sigma_{\boldsymbol{G}}$ satisfies the condition $(*)$ from Proposition 4.3(3), hence this data defines a difference $\boldsymbol{G}$-module. This construction is clearly natural, hence it gives rise to a functor

$$\sigma\ \mathrm{ind}_1^{\boldsymbol{G}} : \mathrm{Mod}_{\boldsymbol{k}[x]} \to \mathrm{Mod}_{\boldsymbol{G}}^{\sigma}.$$

We will show (similarly to the classical context) that this difference induction functor is right adjoint to the forgetful functor

$$\sigma\ \mathrm{res}_1^{\boldsymbol{G}} : \mathrm{Mod}_{\boldsymbol{G}}^{\sigma} \to \mathrm{Mod}_{\boldsymbol{k}[x]}.$$

Proposition 4.7. *The functor $\sigma\ \mathrm{ind}_1^{\boldsymbol{G}}$ is right adjoint to the functor $\sigma\ \mathrm{res}_1^{\boldsymbol{G}}$. Consequently, the functor $\sigma\ \mathrm{ind}_1^{\boldsymbol{G}}$ preserves injective objects.*

Proof. We take $(N, \sigma_N) \in \mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$ and $(M, \sigma_M) \in \mathrm{Mod}_{\boldsymbol{k}[x]}$. After forgetting the endomorphisms σ_N and σ_M, we have (by the classical adjunction) a natural isomorphism

$$\mathrm{Hom}_{\mathrm{Mod}_{\boldsymbol{k}}}(N, M) \simeq \mathrm{Hom}_{\mathrm{Mod}_{\boldsymbol{G}}}(N, M \otimes \boldsymbol{k}[\boldsymbol{G}]).$$

This isomorphism can be explicitly described as taking a $\boldsymbol{k}$-linear map $f : N \to M$ to the composite $(f \otimes \mathrm{id}) \circ \Delta_N$. The inverse is given by postcomposing with the counit in $\boldsymbol{k}[\boldsymbol{G}]$. Then an explicit calculation shows that the both assignments preserve morphisms satisfying the condition $(*)$ from Proposition 4.3(3), which proves our adjunction. Preserving injectives is a formal consequence of having exact left adjoint. □

Now we construct injective objects in $\mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$ by a standard argument.

Corollary 4.8. *Any object M in the category $\mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$ embeds into an injective object.*

Proof. Let $\sigma\ \mathrm{res}_1^{\boldsymbol{G}}(M) \to I$ be an embedding in the category $\mathrm{Mod}_{\boldsymbol{k}[x]}$, where I is injective. Then we take the chain of embeddings

$$M \to \sigma\ \mathrm{ind}_1^{\boldsymbol{G}} \circ \sigma\ \mathrm{res}_1^{\boldsymbol{G}}(M) \to \sigma\ \mathrm{ind}_1^{\boldsymbol{G}}(I),$$

and observe that $\sigma\ \mathrm{ind}_1^{\boldsymbol{G}}(I)$ is injective by Proposition 4.7. □

Since we have enough injective objects, we can develop now homological algebra in the category $\mathrm{Mod}_{\boldsymbol{G}}^{\sigma}$.

Definition 4.9. For a difference rational $\boldsymbol{G}$-module M, we define the *difference rational cohomology groups* (see Example 4.5(1)) as follows:

$$H^n_\sigma(\boldsymbol{G}, M) := \mathrm{Ext}^n_{\mathrm{Mod}^\sigma_{\boldsymbol{G}}}(\boldsymbol{k}, M).$$

We would like to obtain a short exact sequence relating difference rational and rational cohomology groups. We proceed similarly as in Section 3. First, we recall that for a rational $\boldsymbol{G}$-module M, the $\boldsymbol{k}$-vector space $\mathrm{Hom}_{\mathrm{Mod}^\sigma_{\boldsymbol{G}}}(\boldsymbol{k}, M)$ can be identified with

$$M^{\boldsymbol{G}} := \{m \in M \mid \Delta_M(m) = m \otimes 1\}.$$

By the condition $(*)$ from Proposition 4.3(3), we immediately get that for a difference rational $\boldsymbol{G}$-module M, the $\boldsymbol{k}$-module of invariants $M^{\boldsymbol{G}}$ is preserved by σ_M. Therefore, the functor $(-)^{\boldsymbol{G}}$ can be thought of as a functor from $\mathrm{Mod}^\sigma_{\boldsymbol{G}}$ to $\mathrm{Mod}_{\boldsymbol{k}[x]}$. Since we can make the following identification:

$$\mathrm{Hom}_{\mathrm{Mod}^\sigma_{\boldsymbol{G}}}(\boldsymbol{k}, M) = M^{\boldsymbol{G}} \cap M^{\sigma_M},$$

we can factor the above Hom-functor through the category $\mathrm{Mod}_{\boldsymbol{k}[x]}$ as

$$\mathrm{Hom}_{\mathrm{Mod}^\sigma_{\boldsymbol{G}}}(\boldsymbol{k}, -) = (-)^{\sigma_M} \circ (-)^{\boldsymbol{G}}.$$

Now, we recall from the proof of Corollary 4.8 that for an injective cogenerator I of $\mathrm{Mod}_{\boldsymbol{k}[x]}$, $I \otimes \boldsymbol{k}[\boldsymbol{G}]$ is an injective cogenerator of $\mathrm{Mod}^\sigma_{\boldsymbol{G}}$. Then we see that

$$(I \otimes \boldsymbol{k}[\boldsymbol{G}])^{\boldsymbol{G}} = I,$$

hence the functor $(-)^{\boldsymbol{G}}$ preserves injectives. Therefore, we can apply the Grothendieck spectral sequence to our factorization of the functor $\mathrm{Hom}_{\mathrm{Mod}^\sigma_{\boldsymbol{G}}}(\boldsymbol{k}, -)$ and, similarly as in Theorem 3.5, we get the following.

Theorem 4.10. *Let M be a difference rational $\boldsymbol{G}$-module. Then for any $j \geqslant 0$, there is a short exact sequence* (*where $H^{-1}(\boldsymbol{G}, M) := 0$*)

$$0 \to H^{j-1}(\boldsymbol{G}, M)_\sigma \to H^j_\sigma(\boldsymbol{G}, M) \to H^j(\boldsymbol{G}, M)^\sigma \to 0.$$

Proof. The proof of Theorem 3.5 carries over to this situation replacing the ring $A[\sigma_A]$ with the ring $\boldsymbol{k}[x]$ and the discrete cohomology with the rational cohomology. □

Example 4.11. We compute rational difference cohomology in the following special case. As a difference rational group, we consider the additive group scheme $\mathbb{G}_{\mathrm{a}}$ over $\mathbb{F}_p$ ($p > 2$) with the Frobenius endomorphism Fr, and we take the trivial difference rational $(\mathbb{G}_{\mathrm{a}}, \mathrm{Fr})$-module $(\mathbb{F}_p, \mathrm{id})$.

The ring $H^*(\mathbb{G}_{\mathrm{a}}, \mathbb{F}_p)$ was computed in [Cline et al. 1977, Theorem 4.1] together with a description of the rational action of $\mathbb{G}_{\mathrm{m}}$. In particular, $H^1(\mathbb{G}_{\mathrm{a}}, \mathbb{F}_p)$ is an infinite dimensional vector space over $\mathbb{F}_p$ with a basis $\{a_i\}_{i \geqslant 0}$, which can be chosen in such a way that in the action of $\mathbb{F}_p[\sigma]$ ($\simeq \mathbb{F}_p[x]$) on

$$H^1(\mathbb{G}_{\mathrm{a}}, \mathbb{F}_p) = \mathrm{Hom}(\mathbb{G}_{\mathrm{a}}, \mathbb{G}_{\mathrm{a}}),$$

we have $\sigma(a_i) = a_{i+1}$.

Thus we see that $H^1(\mathbb{G}_a, \mathbb{F}_p)^\sigma = 0$ and $\dim(H^1(\mathbb{G}_a, \mathbb{F}_p)_\sigma) = 1$. Since σ acts trivially on $H^0(\mathbb{G}_a, \mathbb{F}_p)$, we get $\dim(H^0(\mathbb{G}_a, \mathbb{F}_p)_\sigma) = 1$, and we obtain by Theorem 4.10 that

$$\dim(H^1_\sigma(\mathbb{G}_a, \mathbb{F}_p)) = 1.$$

In order to extend our computation, we will use the following description of the graded ring $H^*(\mathbb{G}_a, \mathbb{F}_p)$ from [Cline et al. 1977, Theorem 4.1]:

$$H^*(\mathbb{G}_a, \mathbb{F}_p) \simeq \Lambda(H^1(\mathbb{G}_a, \mathbb{F}_p)) \otimes S(\tilde{H}^1(\mathbb{G}_a, \mathbb{F}_p)),$$

where Λ and S stand respectively for the exterior and symmetric algebra over $\mathbb{F}_p$, $\tilde{H}^1(\mathbb{G}_a, \mathbb{F}_p)$ is a space with a basis $\{a_i\}_{i\geqslant 1}$ and its nonzero elements have degree 2. Since Fr commutes with algebraic group homomorphisms, the action of σ on $H^*(\mathbb{G}_a, \mathbb{F}_p)$ is multiplicative. Hence σ acts on decomposable elements of $H^*(\mathbb{G}_a, \mathbb{F}_p)$ diagonally. Therefore, we have that $H^j(\mathbb{G}_a, \mathbb{F}_p)^\sigma = 0$ for all $j > 0$, and we obtain by Theorem 4.10 that

$$H^j_\sigma(\mathbb{G}_a, \mathbb{F}_p) \simeq H^{j-1}(\mathbb{G}_a, \mathbb{F}_p)_\sigma$$

for all $j > 0$. Taking these facts into account, we can summarize our computations as follows:

$$\dim(H^j_\sigma(\mathbb{G}_a, \mathbb{F}_p)) = \begin{cases} 1 & \text{for } j = 0, 1, 2, \\ \infty & \text{for } j > 2. \end{cases}$$

This final outcome may look a bit bizarre, but it coincides with the general philosophy that "invariants reduce the infinite part of the difference dimension by 1" (this can be made precise using the notion of an SU-rank, see [Chatzidakis and Hrushovski 1999, §2.2]).

Continuing the analogy with the discrete situation, we can apply Theorem 4.10 to the induced difference rational module M^∞ (see Example 4.5(3)). We define, analogously to the discrete case, the "stable rational cohomology groups" as

$$H^j_{\mathrm{st}}(\boldsymbol{G}, M) := \operatorname*{colim}_i H^j(\boldsymbol{G}, M^{(i)}).$$

Similarly as in Theorem 3.8, we obtain the following.

Theorem 4.12. *For any rational $\boldsymbol{G}$-module M and $j > 0$, there is an isomorphism*

$$H^j_\sigma(\boldsymbol{G}, M^\infty) \simeq H^{j-1}_{\mathrm{st}}(\boldsymbol{G}, M).$$

5. Applications, alternative approaches and possible generalizations

In this section, we discuss applications of our results to the problem of comparing rational and discrete group cohomology. We also compare our approach with the theories of difference representations in [Kamensky 2013; Wibmer 2014], and sketch another (in a way more ambitious) approach to difference representations.

5A. ***Comparison with earlier approach to difference representations.*** Let us compare our construction of difference representations with the existing theories of representations of difference groups in [Wibmer 2014; Kamensky 2013]. One sees that Lemma 5.2 in [Wibmer 2014] amounts to saying that the category of difference rational representations of $(\boldsymbol{G}, \sigma_{\boldsymbol{G}})$ considered in [Wibmer 2014] is equivalent to the category of rational representations of $\boldsymbol{G}$. In fact, in the approach in [Wibmer 2014; Kamensky 2013], the difference structure on $\boldsymbol{G}$ is not encoded in a single representation but rather in some extra structure on the whole category of representations, namely in the functor $M \mapsto M^{(1)}$ which twists the $\boldsymbol{G}$-action by $\sigma_{\boldsymbol{G}}$. For example, when the difference group is reconstructed from its representation category through the Tannakian formalism [Kamensky 2013], this extra structure is used in an essential way. Hence our approach is, in a sense, more direct. In particular, it allows us to introduce the difference group cohomology which differs from the cohomology of the underlying algebraic group. Actually, both of the approaches build on the same structure. Abstractly speaking, we have a category $\mathscr{C}$ with endofunctor F. Then one can consider just the category $\mathscr{C}$ and investigate the effect of the action of F on it; this is, essentially, the approach initiated in [Wibmer 2014; Kamensky 2013]. On the other hand, one can introduce, like in our approach, the category $\mathscr{C}^F$, whose objects are the arrows

$$\sigma_M : F(M) \to M$$

for $M \in \mathscr{C}$. This approach generalizes the first one, since the construction M^∞ (which can be performed in any category with countable coproducts) produces a faithful functor

$$\mathscr{C} \to \mathscr{C}^F.$$

On the other hand, our functor $\sigma \operatorname{ind}_{\boldsymbol{G}}$ produces important objects like injective cogenerators which do not come from $\mathscr{C}$, hence this approach is potentially more flexible and rich.

5B. ***Comparing cohomology, inverting Frobenius and spectra.*** As we mentioned in Section 1, the main motivation for the present work was its possible application to the problem of comparing rational and discrete cohomology. More specifically, let $\boldsymbol{G}$ be an affine group scheme defined over $\mathbb{F}_p$ and let M be a rational $\boldsymbol{G}$-module. Then, it is natural to compare the rational cohomology groups $H^j(\boldsymbol{G}, M)$ and the discrete cohomology groups $H^j(\boldsymbol{G}(\mathbb{F}_{p^n}), M)$. For $\boldsymbol{G}$ reductive and split over $\mathbb{F}_p$, the comparison is given by the celebrated Cline–van der Kallen–Parshall–Scott theorem [Cline et al. 1977] saying that

$$H^j_{\mathrm{st}}(\boldsymbol{G}, M) := \operatorname*{colim}_i H^j(\boldsymbol{G}, M^{(i)}) \simeq \lim_n H^j(\boldsymbol{G}(\mathbb{F}_{p^n}), M),$$

and that the both limits stabilize for any fixed $j \geqslant 0$. Then it was observed [Parshall 1987, Theorem 4(d)] that the right-hand side above (called sometimes *generic cohomology*) coincides with the discrete group cohomology $H^j(\boldsymbol{G}(\overline{\mathbb{F}}_p), M)$. Our work allows one to interpret the left-hand side as a right derived functor as well (see Theorem 4.12). We hope to use this description in a future work aiming to generalize the main theorem from [Cline et al. 1977] to nonreductive algebraic groups. We expect a theorem on difference cohomology expressing generic cohomology as a sort of completion of rational cohomology. We hope that the comparison on difference level should be easier because the limit with respect to the twists is built into the difference theory. Then, one could obtain the theorem on algebraic groups by taking the M^{∞}-construction (we recall that there is no need for taking "stable discrete cohomology" because the Frobenius morphism on a perfect field is an automorphism). This is a subject of our future work.

We would like to point out certain unexpected similarities between Hrushovski's work [2012] and the homological results from [Cline et al. 1977]. In both cases, the situation somehow "smooths out" after taking higher and higher powers of Frobenius. It is visible in the twisted Lang–Weil estimates from [Hrushovski 2012, Theorem 1.1] and in the main theorem of [Cline et al. 1977] above.

At the time being, we can offer another heuristic reasoning supporting our belief that the difference formalism is an adequate tool for the problem of comparing rational and discrete cohomology. Namely, the principal reason why one should not hope for the existence of an isomorphism between rational and generic group cohomology in general is the fact that the Frobenius morphism becomes an automorphism after restricting to the group of rational points over a perfect field. Hence we have

$$H^*(\boldsymbol{G}(\mathbb{F}_q), M) \simeq H^*(\boldsymbol{G}(\mathbb{F}_q), M^{(1)}),$$

while, in general, there is no reason for the map

$$\sigma_* : H^*(\boldsymbol{G}, M) \to H^*(\boldsymbol{G}, M^{(1)})$$

to be an isomorphism. However, the colimit defining $H^*_{\mathrm{st}}(\boldsymbol{G}, M)$ can be thought of as the result of making the map σ_* invertible (see an example of this phenomenon in Remark 3.7(2)). On the other hand, the process of inverting the endomorphism σ is built into the homological algebra of left difference modules through the construction of the module $\tilde{R}$ defined in Section 2. This supports our belief that the category of left difference modules is a relevant tool in this context.

Actually, the first author succeeded in making the connection between the stable cohomology and the process of inverting Frobenius morphism more precise in an important special case [Chałupnik 2015]. To explain this idea better, let us come back for a moment to a general categorical context of Section 5A. We assume that we have a category $\mathscr{C}$ with an endofunctor F and a family $\{\mathscr{C}_j\}_{j\in\mathbb{Z}}$ of full orthogonal subcategories such that any object in $\mathscr{C}$ is a direct sum of objects from $\{\mathscr{C}_j\}_{j\in\mathbb{Z}}$. Thus we have an equivalence of categories

$$\mathscr{C} \simeq \prod_{j\in\mathbb{Z}} \mathscr{C}_j.$$

Moreover, we assume that F takes $\mathscr{C}_j$ into $\mathscr{C}_{pj}$. This situation is quite common in representation theory over $\mathbb{F}_p$. For example, any central element of infinite order in $\boldsymbol{G}$ produces such a decomposition of the category of rational representations of $\boldsymbol{G}$ with F being the functor of twisting by the Frobenius morphism (see e.g., Proposition 4.6). Then we can grade the category

$$\mathscr{C}^* := \prod_{j\neq 0} \mathscr{C}_j$$

by positive integers, putting

$$\mathscr{C}_i^* := \prod_{d \in Y} \mathscr{C}_{p^i d}$$

for $i \geqslant 0$, where $Y := \mathbb{Z} \setminus p\mathbb{Z}$. Let us take now $M = \bigoplus_{i \geqslant 0} M_i$, where $M_i \in \mathscr{C}_i^*$. Then we see that an object in $(\mathscr{C}^*)^F$ is just a sequence of maps

$$F(M_i) \to M_{i+1},$$

hence it produces a "spectrum of objects of $\mathscr{C}^*$" [Hovey 2001]. The formalism of spectra is a classical tool which is used to formally invert an endofunctor, hence it fits well into our context. In [Chałupnik 2015], the first author considered $\mathscr{C}$ as the category $\hat{\mathscr{P}}$ of "completed" strict polynomial functors in the sense of [Friedlander and Suslin 1997], which is closely related to the category of representations of GL_n. The category $\hat{\mathscr{P}}$ has an orthogonal decomposition

$$\hat{\mathscr{P}} \simeq \prod_{j \geqslant 0} \mathscr{P}_j$$

into the subcategories of strict polynomial functors homogeneous of degree j, and F is the "precomposition with the Frobenius twist functor".

The first author managed to find [Chałupnik 2015, Corollary 4.7] an interpretation of "stable Ext-groups" in $\mathscr{P}$ in terms of Ext-groups in the corresponding category of spectra. He also obtained a version of the main theorem of [Cline et al. 1977] in $\mathscr{P}$ as an analogue of the Freudenthal theorem [Chałupnik 2015, Theorem 5.3(3)].

Let us now try to compare spectra and difference modules in general. Although the starting categories are very close, one introduces homological structures in each case in a different way. Namely, in the case of the category of spectra, the formalism of Quillen model categories is used, while in the case of the category of difference modules, we just use its obvious structure of an Abelian category. The important point here is that the resulting Ext-groups are not the same, since in the interpretation of stable cohomology in terms of difference cohomology there is a shift of degree (see Theorem 4.12). Hence, the relation between these two constructions remains quite mysterious.

5C. ***Functors on the category of difference algebras.*** We finish our paper with discussing another version of the notion of a difference rational representation. In fact, there is a certain ambiguity at the very core of difference algebraic geometry. Namely, there are two natural choices for the kind of functors which could be considered as difference schemes:

(1) Functors from the category of rings to the category of difference sets.

(2) Functors from the category of difference rings to the category of sets.

In the case of representable functors (i.e., affine difference schemes) both of the choices above are equivalent by the Yoneda lemma. Thanks to this, a difference group scheme can be unambiguously defined as (the dual of) a difference Hopf algebra. Unfortunately, this "several choices" problem reappears when one tries to introduce the appropriate notion of a difference representation. In fact, we made in Section 4 the "first choice" which is simpler and sufficient for the main objectives of our article. The drawback of this approach is that the difference structure on the module $M(A)$ from Section 4B does not depend on a possible difference structure on A. In other words: there is no natural way of turning the functor M into a functor on the category of difference $\boldsymbol{k}$-algebras. For this reason, the framework of Section 4 is less general than the one in Section 3. Thus, it would be tempting to introduce the notion of a difference rational representation corresponding to the "second choice" above.

We will outline now an alternative approach, which is potentially richer but is also much more involved technically. We fix a difference field $(\boldsymbol{k}, \sigma)$ and consider the category $\mathrm{Alg}_{(\boldsymbol{k},\sigma)}$ of difference commutative algebras over $\boldsymbol{k}$ as in Section 4A. Then, undoubtedly, we want our difference representation to be some sort of a functor

$$M : \mathrm{Alg}_{(\boldsymbol{k},\sigma)} \to \mathrm{Mod}_{\boldsymbol{k}[\sigma]},$$

such that $M(A)$ is naturally a difference (A, σ_A)-module. Now we need an analogue of the fact that an ordinary rational representation sends a $\boldsymbol{k}$-algebra A to $A \otimes M(\boldsymbol{k})$. A reasonable choice here seems to be the following:

$$M(A) = A[\sigma] \otimes_{\boldsymbol{k}[\sigma]} M(\boldsymbol{k}),$$

since in that case the structure of an $A[\sigma]$-module on $M(A)$ depends both on (A, σ_A) and on (M, σ_M).

When we add to this framework a group action, we obtain the following definition.

Definition 5.1. Let $(\boldsymbol{G}, \sigma_{\boldsymbol{G}})$ be a difference algebraic group. We call a functor

$$M : \mathrm{Alg}_{(\boldsymbol{k},\sigma)} \to \mathrm{Mod}_{\boldsymbol{k}[\sigma]},$$

such that

$$M(A) = A[\sigma] \otimes_{\boldsymbol{k}[\sigma]} M(\boldsymbol{k})$$

a $\boldsymbol{G}$*-difference representation* (or a $\boldsymbol{G}$*-difference module*), if there is a natural (in $A \in \mathrm{Alg}_{(\boldsymbol{k},\sigma)}$) structure of a difference $A[\boldsymbol{G}(A)]$-module on $M(A)$.

With the above definition, we achieve the level of generality we had in the discrete case of Section 3. However, in order to make the category of such difference representations usable, one would like to obtain its algebraic description in terms of comodules over coalgebras etc. Unfortunately, the formulae we have obtained so far are quite complicated and do not fit easily into known patterns. For example, it is not clear how to use them even to show that the category under consideration has enough injective objects. For this reason, in this paper, we decided to adopt the approach corresponding to the "first choice".

References

[Chałupnik 2015] M. Chałupnik, "On spectra and affine strict polynomial functors", 2015. arXiv

[Chatzidakis and Hrushovski 1999] Z. Chatzidakis and E. Hrushovski, "Model theory of difference fields", *Trans. Amer. Math. Soc.* **351**:8 (1999), 2997–3071. MR Zbl

[Chatzidakis and Hrushovski 2008a] Z. Chatzidakis and E. Hrushovski, "Difference fields and descent in algebraic dynamics, I", *J. Inst. Math. Jussieu* **7**:4 (2008), 653–686. MR Zbl

[Chatzidakis and Hrushovski 2008b] Z. Chatzidakis and E. Hrushovski, "Difference fields and descent in algebraic dynamics, II", *J. Inst. Math. Jussieu* **7**:4 (2008), 687–704. MR Zbl

[Cline et al. 1977] E. Cline, B. Parshall, L. Scott, and W. van der Kallen, "Rational and generic cohomology", *Invent. Math.* **39**:2 (1977), 143–163. MR Zbl

[Cohn 1965] R. M. Cohn, *Difference algebra*, Interscience Publishers, New York, 1965. MR Zbl

[Di Vizio et al. 2014] L. Di Vizio, C. Hardouin, and M. Wibmer, "Difference Galois theory of linear differential equations", *Adv. Math.* **260** (2014), 1–58. MR Zbl

[Friedlander and Suslin 1997] E. M. Friedlander and A. Suslin, "Cohomology of finite group schemes over a field", *Invent. Math.* **127**:2 (1997), 209–270. MR Zbl

[Hovey 2001] M. Hovey, "Spectra and symmetric spectra in general model categories", *J. Pure Appl. Algebra* **165**:1 (2001), 63–127. MR Zbl

[Hrushovski 2001] E. Hrushovski, "The Manin–Mumford conjecture and the model theory of difference fields", *Ann. Pure Appl. Logic* **112**:1 (2001), 43–115. MR Zbl

[Hrushovski 2012] E. Hrushovski, "The Elementary Theory of the Frobenius Automorphisms", preprint, 2012, Available at http://www.ma.huji.ac.il/∼ehud/FROB.pdf.

[Jantzen 2003] J. C. Jantzen, *Representations of algebraic groups*, 2nd ed., Mathematical Surveys and Monographs **107**, American Mathematical Society, Providence, RI, 2003. MR Zbl

[Kamensky 2013] M. Kamensky, "Tannakian formalism over fields with operators", *Int. Math. Res. Not.* **2013**:24 (2013), 5571–5622. MR Zbl

[Kowalski and Pillay 2007] P. Kowalski and A. Pillay, "On algebraic σ-groups", *Trans. Amer. Math. Soc.* **359**:3 (2007), 1325–1337. MR Zbl

[Levin 2008] A. Levin, *Difference algebra*, Algebra and Applications **8**, Springer, 2008. MR Zbl

[Medvedev and Scanlon 2014] A. Medvedev and T. Scanlon, "Invariant varieties for polynomial dynamical systems", *Ann. of Math.* (2) **179**:1 (2014), 81–177. MR Zbl

[Ovchinnikov and Wibmer 2015] A. Ovchinnikov and M. Wibmer, "σ-Galois theory of linear difference equations", *Int. Math. Res. Not.* **2015**:12 (2015), 3962–4018. MR Zbl

[Parshall 1987] B. J. Parshall, "Cohomology of algebraic groups", pp. 233–248 in *The Arcata Conference on Representations of Finite Groups* (Arcata, California, 1986), edited by P. Fong, Proc. Sympos. Pure Math. **47**, Amer. Math. Soc., Providence, RI, 1987. MR Zbl

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Mathematics **38**, Cambridge University Press, 1994. MR Zbl

[Wibmer 2014] M. Wibmer, "Affine difference algebraic groups", preprint, 2014. arXiv

Communicated by Anand Pillay
Received 2017-01-03 Revised 2018-02-03 Accepted 2018-06-27

mchal@mimuw.edu.pl *Instytut Matematyki, Uniwersytet Warszawski, Warszawa, Poland*

pkowa@math.uni.wroc.pl *Instytut Matematyczny, Uniwersytet Wrocławski, Wrocław, Poland*

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1581

Density theorems for exceptional eigenvalues for congruence subgroups

Peter Humphries

Using the Kuznetsov formula, we prove several density theorems for exceptional Hecke and Laplacian eigenvalues of Maaß cusp forms of weight 0 or 1 for the congruence subgroups $\Gamma_0(q)$, $\Gamma_1(q)$, and $\Gamma(q)$. These improve and extend upon results of Sarnak and Huxley, who prove similar but slightly weaker results via the Selberg trace formula.

1. Introduction

Let $\kappa \in \{0, 1\}$, let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$, and let χ be a congruence character of Γ satisfying $\chi(-I) = (-1)^\kappa$ should $-I$ be a member of Γ. Denote by $\mathcal{A}_\kappa(\Gamma, \chi)$ the space spanned by Maaß cusp forms of weight κ, level Γ, and nebentypus χ, namely the L^2-closure of the space of smooth functions $f : \mathbb{H} \to \mathbb{C}$ satisfying

- $f(\gamma z) = \chi(\gamma) j_\gamma(z)^\kappa f(z)$ for all $\gamma \in \Gamma$ and $z \in \mathbb{H}$, where for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$,
$$j_\gamma(z) := \frac{cz+d}{|cz+d|},$$
- f is an eigenfunction of the weight κ Laplacian
$$\Delta_\kappa := -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + i\kappa y \frac{\partial}{\partial x},$$
- f is of moderate growth, and
- the constant term is zero in the Fourier expansion of f at every cusp $\mathfrak{a}$ of $\Gamma\backslash\mathbb{H}$ that is singular with respect to χ.

We may choose a basis $\mathcal{B}_\kappa(\Gamma, \chi)$ of the complex vector space $\mathcal{A}_\kappa(\Gamma, \chi)$ consisting of Hecke eigenforms. For $f \in \mathcal{B}_\kappa(\Gamma, \chi)$, we let $\lambda_f = \frac{1}{4} + t_f^2$ denote the eigenvalue of the weight κ Laplacian, where either $t_f \in [0, \infty)$ or $it_f \in \left(0, \frac{1}{2}\right)$. Similarly, we let $\lambda_f(p)$ denote the eigenvalue of the Hecke operator T_p at a prime p, so that $|\lambda_f(p)| < p^{\frac{1}{2}} + p^{-\frac{1}{2}}$. The generalised Ramanujan conjecture states that t_f is real and that $|\lambda_f(p)| \le 2$ for every prime p. Exceptions to this conjecture are called exceptional eigenvalues. It is known that exceptional Laplacian eigenvalues cannot occur if $\kappa = 1$, while for $\kappa = 0$ there are no exceptional Laplacian eigenvalues for Maaß cusp forms of squarefree conductor less than 857 [Booker

MSC2010: primary 11F72; secondary 11F30.
Keywords: Selberg eigenvalue conjecture, Ramanujan conjecture.

and Strömbergsson 2007, Theorem 1]. The best current bounds towards the generalised Ramanujan conjecture are due to Kim and Sarnak [2003]; they show that

$$\lambda_f \geq \tfrac{1}{4} - \left(\tfrac{7}{64}\right)^2, \quad |\lambda_f(p)| \leq p^{\frac{7}{64}} + p^{-\frac{7}{64}}.$$

Results. In this paper, we use the Kuznetsov formula to prove density results for exceptional eigenvalues for the congruence subgroups

$$\Gamma_0(q) := \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : c \equiv 0 \pmod{q}\right\},$$
$$\Gamma_1(q) := \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : a, d \equiv 1 \pmod{q},\ c \equiv 0 \pmod{q}\right\},$$
$$\Gamma(q) := \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : a, d \equiv 1 \pmod{q},\ b, c \equiv 0 \pmod{q}\right\},$$

with χ equal to the trivial character for the latter two congruence subgroups. Recall that

$$\mathrm{vol}(\Gamma\backslash\mathbb{H}) = \frac{\pi}{3}[\mathrm{SL}_2(\mathbb{Z}) : \Gamma] = \begin{cases} \frac{\pi}{3} q \prod_{p \mid q}\left(1 + \frac{1}{p}\right) & \text{if } \Gamma = \Gamma_0(q), \\ \frac{\pi}{3} q^2 \prod_{p \mid q}\left(1 - \frac{1}{p^2}\right) & \text{if } \Gamma = \Gamma_1(q), \\ \frac{\pi}{3} q^3 \prod_{p \mid q}\left(1 - \frac{1}{p^2}\right) & \text{if } \Gamma = \Gamma(q). \end{cases}$$

When χ is the trivial character, we write $\mathcal{B}_\kappa(\Gamma)$ in place of $\mathcal{B}_\kappa(\Gamma, \chi)$, while when $\Gamma = \Gamma_0(q)$, we write this as $\mathcal{B}_\kappa(q, \chi)$. Given positive integers q and q_χ with $q_\chi \mid q$, we factorise $q = \prod_{p^\alpha \| q} p^\alpha$ and $q_\chi = \prod_{p^\gamma \| q_\chi} p^\gamma$, and define

$$\dot{Q} = \dot{Q}(q, q_\chi) = \prod_{\substack{p^\alpha \| q \\ p^\gamma \| q_\chi}} \dot{Q}(p^\alpha, p^\gamma), \quad \ddot{Q} = \ddot{Q}(q, q_\chi) = \prod_{\substack{p^\alpha \| q \\ p^\gamma \| q_\chi}} \ddot{Q}(p^\alpha, p^\gamma),$$

with

$$\dot{Q}(p^\alpha, p^\gamma) := \begin{cases} p^{\lfloor (3\alpha+1)/4 \rfloor - \alpha/2} & \text{if } p \text{ is odd and } \alpha = \gamma \geq 3, \\ 2^{\lfloor (3\alpha+1)/4 \rfloor - \alpha/2} & \text{if } p = 2 \text{ and } \gamma + 1 \geq \alpha \geq 3, \\ 1 & \text{otherwise,} \end{cases}$$

$$\ddot{Q}(p^\alpha, p^\gamma) := \begin{cases} p & \text{if } p \text{ is odd and } \alpha = \gamma \geq 3, \\ 4 & \text{if } p = 2 \text{ and } \alpha = \gamma \geq 3, \\ 2 & \text{if } p = 2 \text{ and } \alpha = \gamma + 1 \geq 3, \\ 1 & \text{otherwise.} \end{cases}$$

Theorem 1.1. *For any fixed finite collection of primes $\mathcal{P}$ not dividing q, any $\alpha_p \in (2, p^{\frac{1}{2}} + p^{-\frac{1}{2}})$ and $0 \leq \mu_p \leq 1$ for all $p \in \mathcal{P}$ with $\sum_{p \in \mathcal{P}} \mu_p = 1$, we have that*

$$\begin{aligned} &\#\left\{f \in \mathcal{B}_\kappa(\Gamma_1(q)) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\right\} \\ &\qquad \ll_\varepsilon \mathrm{vol}(\Gamma_1(q)\backslash\mathbb{H})^{1 - 3\sum_{p \in \mathcal{P}} \mu_p (\log \alpha_p/2)/\log p + \varepsilon} (T^2)^{1 - 4\sum_{p \in \mathcal{P}} \mu_p (\log \alpha_p/2)/\log p + \varepsilon}, \end{aligned} \tag{1.2}$$

$$\#\big\{f \in \mathcal{B}_\kappa(\Gamma(q)) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \textit{ for all } p \in \mathcal{P}\big\} \\ \ll_\varepsilon \operatorname{vol}(\Gamma(q)\backslash\mathbb{H})^{1-\frac{8}{3}\sum_{p\in\mathcal{P}} \mu_p (\log\alpha_p/2)/\log p+\varepsilon} (T^2)^{1-4\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p+\varepsilon}, \quad (1.3)$$

$$\#\big\{f \in \mathcal{B}_\kappa(q, \chi) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \textit{ for all } p \in \mathcal{P}\big\} \\ \ll_\varepsilon \operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})^{1-4\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p+\varepsilon} (T^2)^{1-4\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p+\varepsilon} \\ \times \min\big\{\dot{Q}^{4\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p},\ \ddot{Q}^{1-4\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p}\big\}. \quad (1.4)$$

Theorem 1.1 should be compared to the Weyl law, which states that

$$\#\big\{f \in \mathcal{B}_\kappa(\Gamma, \chi) : t_f \in [0, T]\big\} \sim \frac{\operatorname{vol}(\Gamma\backslash\mathbb{H})}{4\pi} T^2.$$

For $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, so that χ is the trivial character, and $\mathcal{P}$ consisting of a single prime p, Theorem 1.1 is a result of Blomer, Buttcane, and Raulf [Blomer et al. 2014, Proposition 1], improving on a slightly weaker result of Sarnak [1987, Theorem 1.1], who uses the Selberg trace formula in place of the Kuznetsov formula and obtains instead (see [Blomer et al. 2014, Footnote 1])

$$\#\big\{f \in \mathcal{B}_0(\mathrm{SL}_2(\mathbb{Z})) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha\big\} \ll (T^2)^{1-2(\log\alpha/2)/\log p}.$$

Theorem 1.5. *For any fixed finite (possibly empty) collection of primes $\mathcal{P}$ not dividing q, any $\alpha_0 \in \big(0, \frac{1}{2}\big)$, $\alpha_p \in (2, p^{\frac{1}{2}} + p^{-\frac{1}{2}})$, and $0 \leq \mu_0, \mu_p \leq 1$ for all $p \in \mathcal{P}$ with $\mu_0 + \sum_{p\in\mathcal{P}} \mu_p = 1$, we have that*

$$\#\big\{f \in \mathcal{B}_0(\Gamma_1(q)) : it_f \in \big(\alpha_0, \tfrac{1}{2}\big),\ |\lambda_f(p)| \geq \alpha_p \textit{ for all } p \in \mathcal{P}\big\} \\ \ll_\varepsilon \operatorname{vol}(\Gamma_1(q)\backslash\mathbb{H})^{1-3\left(\mu_0\alpha_0+\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p\right)+\varepsilon}, \quad (1.6)$$

$$\#\big\{f \in \mathcal{B}_0(\Gamma(q)) : it_f \in \big(\alpha_0, \tfrac{1}{2}\big),\ |\lambda_f(p)| \geq \alpha_p \textit{ for all } p \in \mathcal{P}\big\} \\ \ll_\varepsilon \operatorname{vol}(\Gamma(q)\backslash\mathbb{H})^{1-\frac{8}{3}\left(\mu_0\alpha_0+\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p\right)+\varepsilon}. \quad (1.7)$$

$$\#\big\{f \in \mathcal{B}_0(q, \chi) : it_f \in \big(\alpha_0, \tfrac{1}{2}\big),\ |\lambda_f(p)| \geq \alpha_p \textit{ for all } p \in \mathcal{P}\big\} \\ \ll_\varepsilon \operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})^{1-4\left(\mu_0\alpha_0+\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p\right)+\varepsilon} \\ \times \min\big\{\dot{Q}^{4\left(\mu_0\alpha_0+\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p\right)},\ \ddot{Q}^{1-4\left(\mu_0\alpha_0+\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p\right)}\big\}. \quad (1.8)$$

When $\mathcal{P}$ is empty and χ is the trivial congruence character, Theorem 1.5 improves upon a result of Huxley [1986], who uses the Selberg trace formula in place of the Kuznetsov formula and obtains instead this result with the exponent 2 for each of the three congruence subgroups instead of 3, $\frac{8}{3}$, and 4 respectively. When $\mathcal{P}$ is empty and χ is the trivial congruence character, (1.8) is a result of Iwaniec [2002, Theorem 11.7] ; see also [Iwaniec and Kowalski 2004, (16.61)].

Since

$$\left\lfloor \frac{3\alpha+1}{4} \right\rfloor - \frac{\alpha}{2} \leq \frac{3\alpha}{10},$$

so that $\dot{Q} \ll \operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})^{\frac{3}{10}}$, the right-hand side of (1.4) is bounded by

$$\operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})^{1-\frac{14}{5}\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p+\varepsilon} (T^2)^{1-4\sum_{p\in\mathcal{P}} \mu_p(\log\alpha_p/2)/\log p+\varepsilon},$$

while the right-hand side of (1.8) is bounded by

$$\operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})^{1-\frac{14}{5}(\mu_0\alpha_0+\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p)+\varepsilon}.$$

On the other hand, taking $\mathcal{P}$ to consist of a single prime in (1.4) recovers the Selberg bound $\lambda_f(p) \ll_\varepsilon p^{\frac{1}{4}+\varepsilon}$ for an individual element $f \in \mathcal{B}_\kappa(q,\chi)$ by taking T sufficiently large, while taking $\mathcal{P}$ to be empty in (1.8) recovers the Selberg bound $\lambda_f \geq \frac{3}{16}$ by embedding f in $\mathcal{B}_\kappa(qQ,\chi)$ and taking Q sufficiently large.

Finally, we also prove the following improvements of Theorems 1.1 and 1.5 for $\Gamma_1(q)$ with q squarefree via a twisting argument.

Theorem 1.9. *When q is squarefree*, (1.2) *and* (1.6) *hold with the exponent* 3 *replaced by* 4.

Idea of Proof. By Rankin's trick (which is to say Chebyshev's inequality), it suffices to find bounds for

$$\sum_{\substack{f\in\mathcal{B}_\kappa(\Gamma,\chi)\\ t_f\in[0,T]}}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}, \qquad \sum_{\substack{f\in\mathcal{B}_0(\Gamma,\chi)\\ it_f\in(0,\frac{1}{2})}} X^{2it_f}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}$$

for nonnegative integers ℓ_p and a positive real number $X \geq 1$ to be chosen. To bound these quantities, we begin with the Kuznetsov formula for $\mathcal{B}_\kappa(q,\chi)$; we then use the Atkin–Lehner decomposition to turn this into a Kuznetsov formula for $\mathcal{B}_\kappa(\Gamma,\chi)$. We take a test function in the Kuznetsov formula that localises the spectral sum to cusp forms with $t_f \in [0,T]$ in the case of Theorem 1.1 and to cusp forms with $it_f \in \left(0,\frac{1}{2}\right)$ in the case of Theorem 1.5. We use the Hecke relations to introduce powers of the Hecke eigenvalues into the Kuznetsov formula. By positivity, we discard the contribution of the continuous spectrum, and we are left with bounding the right-hand side of the Kuznetsov formula.

The chief novelty of the proof is the bounds for sums of Kloosterman sums in the Kuznetsov formula for each congruence subgroup. As well as the usual Weil bound, we use character orthogonality for $\Gamma_1(q)$ and $\Gamma(q)$, at which point we only use the trivial bound for the resulting sum of Kloosterman sums. For $\Gamma_0(q)$ and χ the principal character, we may also use the Weil bound, but for χ nonprincipal, additional difficulties arise in bounding the Kloosterman sum, with the bound possibly depending on the conductor of χ; it is for this reason that the bounds (1.4) and (1.8) involve $\dot{Q}$, for $\dot{Q}$ arises when only weaker bounds than the Weil bound are possible for the Kloosterman sums involved.

We also highlight the key trick to proving Theorem 1.9, namely that the Laplacian eigenvalue and absolute value of a Hecke eigenvalue of a Maaß form remain unchanged under twisting by a Dirichlet character. Twisting may alter the level of a Maaß form, yet Theorem 1.9 involves a favourable situation in which the resulting family of twisted Maaß forms are sufficiently well-behaved that we are able to improve the exponent in the density theorem.

It is worth mentioning that the results in this paper ought to generalise naturally to cusp forms on GL_2 over arbitrary number fields F. Bruggeman and Miatello [2009] prove a form of the Kuznetsov formula for GL_2 over a totally real field and use this to prove weighted Weyl law for cusp forms. Similarly, Maga [2013] proves a semiadèlic version of the Kuznetsov formula for GL_2 over an arbitrary number field. In the former case, this formula is valid for congruence subgroups of the form $\Gamma_0(\mathfrak{q})$ for a nonzero integral ideal $\mathfrak{q}$ of the ring of integers $\mathcal{O}_F$ of F and arbitrary congruence characters χ modulo $\mathfrak{q}$, while the latter only treats the case of trivial congruence character but should easily be able to be generalised to arbitrary congruence character; this is precisely what is required for density theorems for the congruence subgroups $\Gamma_0(\mathfrak{q})$, $\Gamma_1(\mathfrak{q})$, and $\Gamma(\mathfrak{q})$.

2. The Kuznetsov formula

The background on automorphic forms and notation in this section largely follows [Duke et al. 2002]; see [Duke et al. 2002, Section 4] for more details. Let $\kappa \in \{0, 1\}$, and let χ be a primitive Dirichlet character modulo q_χ, where q_χ divides q, satisfying $\chi(-1) = (-1)^\kappa$; this defines a congruence character of $\Gamma_0(q)$ via $\chi(\gamma) := \chi(d)$ for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(q)$. We denote by $L^2(\Gamma_0(q)\backslash\mathbb{H}, \kappa, \chi)$ the L^2-completion of the space of all smooth functions $f : \mathbb{H} \to \mathbb{C}$ that are of moderate growth and satisfy $f(\gamma z) = \chi(\gamma) j_\gamma(z)^\kappa f(z)$. This space has the spectral decomposition

$$L^2(\Gamma_0(q)\backslash\mathbb{H}, \kappa, \chi) = \mathcal{A}_\kappa(q, \chi) \oplus \mathcal{E}_\kappa(q, \chi)$$

with respect to the weight κ Laplacian, where $\mathcal{A}_\kappa(q, \chi) := \mathcal{A}_\kappa(\Gamma_0(q), \chi)$ is the space spanned by Maaß cusp forms of weight κ, level q, and nebentypus χ, and $\mathcal{E}_\kappa(q, \chi)$ is the space spanned by incomplete Eisenstein series parametrised by the cusps $\mathfrak{a}$ of $\Gamma_0(q)\backslash\mathbb{H}$ that are singular with respect to χ.

We denote by $\mathcal{B}_\kappa(q, \chi)$ an orthonormal basis of Maaß cusp forms $f \in \mathcal{A}_\kappa(q, \chi)$ normalised to have L^2-norm 1:

$$\langle f, f\rangle_q := \int_{\Gamma_0(q)\backslash\mathbb{H}} |f(z)|^2 \, d\mu(z) = 1,$$

where $d\mu(z) = dx\,dy/y^2$ is the $\mathrm{SL}_2(\mathbb{R})$-invariant measure on $\mathbb{H}$. Later we will use the Atkin–Lehner decomposition of $\mathcal{A}_\kappa(q, \chi)$ in order to specify that $\mathcal{B}_\kappa(q, \chi)$ can be chosen to consist of linear combinations of Hecke eigenforms. The Fourier expansion of $f \in \mathcal{B}_\kappa(q, \chi)$ is

$$f(z) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \rho_f(n) W_{\mathrm{sgn}(n)\kappa/2, it_f}(4\pi|n|y) e(nx),$$

where $W_{\alpha,\beta}$ is the Whittaker function and

$$\rho_f(n) W_{\mathrm{sgn}(n)\kappa/2, it_f}(4\pi|n|y) = \int_0^1 f(z) e(-nx)\, dx.$$

For a singular cusp $\mathfrak{a}$, we define the Eisenstein series

$$E_{\mathfrak{a}}(z, s, \chi) := \sum_{\gamma \in \Gamma_{\mathfrak{a}}\backslash\Gamma_0(q)} \overline{\chi}(\gamma) j_{\sigma_{\mathfrak{a}}^{-1}\gamma}(z)^{-\kappa} \Im(\sigma_{\mathfrak{a}}^{-1}\gamma z)^s,$$

which is absolutely convergent for $\Re(s) > 1$ and extends meromorphically to $\mathbb{C}$, with the Fourier expansion

$$\delta_{\mathfrak{a},\infty} y^{\frac{1}{2}+it} + \varphi_{\mathfrak{a},\infty}\left(\tfrac{1}{2}+it, \chi\right) y^{\frac{1}{2}-it} + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \rho_{\mathfrak{a}}(n, t, \chi) W_{\mathrm{sgn}(n)\kappa/2, it}(4\pi|n|y) e(nx)$$

for $s = \frac{1}{2} + it$ with $t \in \mathbb{R} \setminus \{0\}$, where

$$\delta_{\mathfrak{a},\infty} y^{\frac{1}{2}+it} + \varphi_{\mathfrak{a},\infty}\big(\tfrac{1}{2}+it, \chi\big) y^{\frac{1}{2}-it} := \int_0^1 E_{\mathfrak{a}}\big(z, \tfrac{1}{2}+it, \chi\big)\, dx,$$

$$\rho_{\mathfrak{a}}(n, t, \chi) W_{\operatorname{sgn}(n)\kappa/2, it}(4\pi |n| y) := \int_0^1 E_{\mathfrak{a}}\big(z, \tfrac{1}{2}+it, \chi\big) e(-nx)\, dx.$$

The subspace $\mathcal{E}_\kappa(q, \chi)$ consists of functions $g \in L^2(\Gamma_0(q)\backslash\mathbb{H}, \kappa, \chi)$ that are orthogonal to every Maaß cusp form $f \in \mathcal{A}_\kappa(q, \chi)$; it is the L^2-closure of the space spanned by incomplete Eisenstein series, which are functions of the form

$$E_{\mathfrak{a}}(z, \psi, \chi) := \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} E_{\mathfrak{a}}(z, s, \chi) \widehat{\psi}(s)\, ds \tag{2.1}$$

for some singular cusp $\mathfrak{a}$ and some smooth function of compact support $\psi : \mathbb{R}^+ \to \mathbb{C}$, where $\sigma > 1$ and

$$\widehat{\psi}(s) := \int_0^\infty \psi(x) x^{-s} \,\frac{dx}{x}.$$

Theorem 2.2 [Duke et al. 2002, Proposition 5.2]. *For $m, n \geq 1$ and $r \in \mathbb{R}$,*

$$\sum_{f \in \mathcal{B}_\kappa(q,\chi)} \frac{4\pi \sqrt{mn}\, \overline{\rho_f}(m) \rho_f(n)}{\cosh \pi(r - t_f) \cosh \pi(r + t_f)} + \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} \frac{\sqrt{mn}\, \overline{\rho_{\mathfrak{a}}}(m, t, \chi) \rho_{\mathfrak{a}}(n, t, \chi)}{\cosh \pi(r - t) \cosh \pi(r + t)}\, dt$$
$$= \frac{|\Gamma(1 - \kappa/2 - ir)|^2}{\pi^2} \Bigg(\delta_{m,n} + \sum_{\substack{c=1 \\ c \equiv 0 \ (\mathrm{mod}\ q)}}^{\infty} \frac{S_\chi(m, n; c)}{c} I_\kappa\Big(\frac{4\pi\sqrt{mn}}{c}, r\Big)\Bigg),$$

where

$$S_\chi(m, n; c) := \sum_{d \in (\mathbb{Z}/c\mathbb{Z})^\times} \chi(d) e\Big(\frac{md + n\bar{d}}{c}\Big),$$

$$I_\kappa(t, r) := -2t \int_{-i}^{i} (-i\zeta)^{\kappa-1} K_{2ir}(\zeta t)\, d\zeta,$$

with the latter integral being over the semicircle $|z| = 1$, $\Re(z) > 0$.

By the reflection formula for the gamma function, we have that for $r \in \mathbb{R}$,

$$\Big|\Gamma\Big(1 - \frac{\kappa}{2} - ir\Big)\Big|^2 = \begin{cases} \pi r/\sinh \pi r & \text{if } \kappa = 0, \\ \pi/\cosh \pi r & \text{if } \kappa = 1. \end{cases}$$

Given a sufficiently well-behaved function h, we may multiply both sides of the pre-Kuznetsov formula for $\kappa = 0$ by

$$\frac{1}{2}\Big(h\Big(r + \frac{i}{2}\Big) + h\Big(r - \frac{i}{2}\Big)\Big) \cosh \pi r$$

and then integrate both sides from $-\infty$ to ∞ with respect to r. This yields the following Kuznetsov formula (see [Blomer et al. 2007, Section 2.1.4; Iwaniec and Kowalski 2004, Theorem 16.3; Knightly and Li 2013, Equation (7.32)]):

Theorem 2.3. *Let $\delta > 0$, and let h be a function that is even, holomorphic in the horizontal strip $|\Im(t)| \leq \frac{1}{2} + \delta$, and satisfies $h(t) \ll (|t|+1)^{-2-\delta}$. Then*

$$\sum_{f \in \mathcal{B}_0(q,\chi)} 4\pi \sqrt{mn}\,\overline{\rho_f}(m)\rho_f(n) \frac{h(t_f)}{\cosh \pi t_f} + \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} \sqrt{mn}\,\overline{\rho_{\mathfrak{a}}}(m,t,\chi)\rho_{\mathfrak{a}}(n,t,\chi) \frac{h(t)}{\cosh \pi t}\,dt$$
$$= \delta_{mn} g_0 + \sum_{\substack{c=1 \\ c \equiv 0 \pmod q}}^{\infty} \frac{S_\chi(m,n;c)}{c} g_0\Big(\frac{4\pi\sqrt{mn}}{c}\Big),$$

where

$$g_0 := \frac{1}{\pi} \int_{-\infty}^{\infty} r h(r) \tanh \pi r\,dr, \qquad g_0(x) := 2i \int_{-\infty}^{\infty} J_{2ir}(x) \frac{r h(r)}{\cosh \pi r}\,dr.$$

The left-hand side of the Kuznetsov formula is called the spectral side; the first term is the contribution from the discrete spectrum, while the second term is the contribution from the continuous spectrum. The right-hand side of the Kuznetsov formula is called the geometric side; the first term is the delta term and the second term is the Kloosterman term.

3. Decomposition of spaces of modular forms

Eisenstein series and Hecke operators. The space $\mathcal{E}_\kappa(q,\chi)$ is spanned by incomplete Eisenstein series of the form (2.1), which are obtained by integrating test functions against Eisenstein series indexed by singular cusps $\mathfrak{a}$; in this sense, the Eisenstein series $E_{\mathfrak{a}}(z,s,\chi)$ are a spanning set for $\mathcal{E}_\kappa(q,\chi)$. We may instead choose a different spanning set of Eisenstein series for $\mathcal{E}_\kappa(q,\chi)$; in place of the set of Eisenstein series $E_{\mathfrak{a}}(z,s,\chi)$ with $\mathfrak{a}$ a singular cusp, we may instead choose a spanning set of Eisenstein series of the form $E(z,s,f)$ with Fourier expansion

$$c_{1,f}(t) y^{\frac{1}{2}+it} + c_{2,f}(t) y^{\frac{1}{2}-it} + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \rho_f(n,t,\chi) W_{\operatorname{sgn}(n)\kappa/2,it}(4\pi|n|y) e(nx)$$

for $s = \frac{1}{2} + it$ with $t \in \mathbb{R} \setminus \{0\}$, where $\mathcal{B}(\chi_1,\chi_2) \ni f$ with $\chi_1\chi_2 = \chi$ is some finite set depending on χ_1, χ_2 corresponding to an orthonormal basis in the space of the induced representation constructed out of the pair (χ_1,χ_2); see [Blomer et al. 2007, Section 2.1.1] or [Knightly and Li 2013, Chapter 5]. For our purposes, we need not be more specific about $\mathcal{B}(\chi_1,\chi_2)$, other than noting that for each $f \in \mathcal{B}(\chi_1,\chi_2)$, the Eisenstein series $E\big(z, \frac{1}{2}+it, f\big)$ is an eigenfunction of the Hecke operators T_n for $(n,q) = 1$ with Hecke eigenvalues

$$\lambda_f(n,t) = \sum_{ab=n} \chi_1(a) a^{it} \chi_2(b) b^{-it},$$

where for $g : \mathbb{H} \to \mathbb{C}$ a periodic function of period one,

$$(T_n g)(z) := \frac{1}{\sqrt{n}} \sum_{ad=n} \chi(a) \sum_{b \pmod d} g\Big(\frac{az+b}{d}\Big).$$

So for $f \in \mathcal{B}(\chi_1, \chi_2)$,

$$\lambda_f(m,t)\lambda_f(n,t) = \sum_{d\mid(m,n)} \chi(d)\lambda_f\Big(\frac{mn}{d^2}, t\Big), \tag{3.1}$$

$$\overline{\lambda_f}(n,t) = \overline{\chi}(n)\lambda_f(n,t), \tag{3.2}$$

$$\rho_f(1,t)\lambda_f(n) = \sqrt{n}\rho_f(n,t), \tag{3.3}$$

whenever $m, n \geq 1$ with $(mn, q) = 1$ and $s = \frac{1}{2} + it$.

Lemma 3.4 (cf. [Conrey et al. 1997, Lemma 3; Hughes and Miller 2007, Lemma 2.8; Petrow and Young 2018, Section 6]). *For any prime $p \nmid q$ and positive integer ℓ, we have that*

$$|\lambda_f(p,t)|^{2\ell} = \sum_{j=0}^{\ell} \alpha_{2j,2\ell}\overline{\chi}(p)^j \lambda_f(p^{2j}, t) \tag{3.5}$$

for any $f \in \mathcal{B}(\chi_1, \chi_2)$ and $s = \frac{1}{2} + it$, where

$$\alpha_{2j,2\ell} = \frac{2j+1}{\ell+j+1}\binom{2\ell}{\ell+j} = \begin{cases} \binom{2\ell}{\ell-j} - \binom{2\ell}{\ell-j-1} & \text{if } 0 \leq j \leq \ell-1, \\ 1 & \text{if } j = \ell, \end{cases} \tag{3.6}$$

so that each $\alpha_{2j,2\ell}$ is positive and satisfies

$$\sum_{j=0}^{\ell} \alpha_{2j,2\ell} = \binom{2\ell}{\ell} \leq 2^{2\ell}. \tag{3.7}$$

Proof. That (3.7) follows from (3.6) is clear. For (3.5), we have that

$$\overline{\chi}(p)^{j/2}\lambda_f(p^j, t) = U_j\bigg(\frac{\overline{\chi}(p)^{\frac{1}{2}}\lambda_f(p,t)}{2}\bigg),$$

where U_j is the j-th Chebyshev polynomial of the second kind, because U_j satisfies $U_0(x/2) = 1$, $U_1(x/2) = x$, and the recurrence relation

$$U_{j+1}\Big(\frac{x}{2}\Big) = xU_j\Big(\frac{x}{2}\Big) - U_{j-1}\Big(\frac{x}{2}\Big)$$

for all $j \geq 1$, and $\overline{\chi}(p)^{j/2}\lambda_f(p^j, t)$ satisfies the same recurrence relation from (3.1). Since

$$\frac{2}{\pi}\int_{-1}^{1} U_j(x)U_k(x)\sqrt{1-x^2}\,dx = \delta_{j,k},$$

we have that

$$x^{2\ell} = \sum_{j=0}^{2\ell} \alpha_{j,2\ell} U_j\Big(\frac{x}{2}\Big),$$

where

$$\alpha_{j,2\ell} = \frac{2^{2\ell+1}}{\pi}\int_{-1}^{1} x^{2\ell}U_j(x)\sqrt{1-x^2}\,dx.$$

This vanishes if j is odd as $U_j(-x) = (-1)^j U_j(x)$, while for j even we have the identity (3.6) from [Gradshteyn and Ryzhik 2007, 7.311.2]. Combined with (3.2), this proves (3.5). $\square$

Atkin–Lehner decomposition for $\Gamma_0(q)$. Similarly, we may choose a basis of $\mathcal{A}_\kappa(q,\chi)$ consisting of linear combinations of Hecke eigenforms. Let $\mathcal{B}_\kappa^*(q,\chi)$ denote the set of newforms of weight κ, level q, and nebentypus χ, and let $\mathcal{A}_\kappa^*(q,\chi)$ denote the subspace of $\mathcal{A}_\kappa(q,\chi)$ spanned by such newforms. Recall that a newform $f \in \mathcal{B}_\kappa^*(q,\chi)$ is an eigenfunction of the weight κ Laplacian Δ_κ with eigenvalue $\frac{1}{4}+t_f^2$ and of every Hecke operator T_n, $n \geq 1$, with eigenvalue $\lambda_f(n)$, as well as the operator $Q_{\frac{1}{2}+it_f,\kappa}$ as defined in [Duke et al. 2002, Section 4], with eigenvalue $\epsilon_f \in \{-1,1\}$; we say that f is even if $\epsilon_f = 1$ and f is odd if $\epsilon_f = -1$. In particular,

$$\lambda_f(m)\lambda_f(n) = \sum_{\substack{d\mid(m,n)\\(d,q)=1}} \chi(d)\lambda_f\Big(\frac{mn}{d^2}\Big), \tag{3.8}$$

$$\rho_f(1)\lambda_f(n) = \sqrt{n}\rho_f(n) \tag{3.9}$$

whenever $m, n \geq 1$, and

$$\overline{\lambda_f}(n) = \overline{\chi}(n)\lambda_f(n) \tag{3.10}$$

for $n \geq 1$ with $(n,q) = 1$. Using (3.8) and (3.10), we have the following:

Lemma 3.11. *For any prime $p\nmid q$ and positive integer ℓ, we have that*

$$|\lambda_f(p)|^{2\ell} = \sum_{j=0}^{\ell} \alpha_{2j,2\ell}\overline{\chi}(p)^j\lambda_f(p^{2j}) \tag{3.12}$$

for any $f \in \mathcal{B}_\kappa^*(q,\chi)$, where once again $\alpha_{2j,2\ell}$ is given by (3.6).

The Atkin–Lehner decomposition states that

$$\mathcal{A}_\kappa(q,\chi) = \bigoplus_{\substack{q_1q_2=q\\ q_1\equiv 0 \ (\mathrm{mod}\ q_\chi)}} \ \bigoplus_{f\in\mathcal{B}_\kappa^*(q_1,\chi)} \ \bigoplus_{d\mid q_2} \mathbb{C}\cdot\iota_{d,q_1,q}f,$$

where $\iota_{d,q_1,q} : \mathcal{A}_\kappa(q_1,\chi) \to \mathcal{A}_\kappa(q,\chi)$ is the map $\iota_{d,q_1,q}f(z) = f(dz)$. The map $\iota_{d,q_1,q}$ commutes with the weight k Laplacian Δ_κ and the Hecke operators T_n whenever $n \geq 1$ and $(n,q) = 1$. It follows that if $g = \iota_{d,q_1,q}f$ for some $f \in \mathcal{B}_\kappa^*(q_1,\chi)$, then $t_g = t_f$ and $\lambda_g(n) = \lambda_f(n)$ whenever $n \geq 1$ and $(n,q) = 1$. Note, however, that $\rho_g(1) = 0$ unless $d = 1$, in which case $\rho_g(1) = \rho_f(1)$.

Unfortunately, the inner Atkin–Lehner decomposition

$$\bigoplus_{d\mid q_2} \mathbb{C}\cdot\iota_{d,q_1,q}f$$

is not an orthogonal decomposition. Nonetheless, one may make use of this decomposition in determining an orthonormal basis of $\mathcal{A}_\kappa(q,\chi)$. For squarefree q and principal nebentypus, this is a result of Iwaniec, Luo, and Sarnak [Iwaniec et al. 2000, Lemma 2.4], while Blomer and Milićević [2015, Lemma 9] have generalised this to nonsquarefree q. Here we generalise this further to nonprincipal nebentypus; this has also independently been derived by Schulze-Pillot and Yenirce [2018] via a different method.

Lemma 3.13 (cf. [Iwaniec et al. 2000, Lemma 2.4; Blomer and Milićević 2015, Lemma 9]). *Suppose that χ has conductor $q_\chi \mid q$, and suppose that $q_1 q_2 = q$ with $q_1 \equiv 0 \pmod{q_\chi}$. For $f \in \mathcal{B}_\kappa^*(q_1, \chi)$ and $\ell_1, \ell_2 \mid q_2$, we have that*

$$\frac{\langle \iota_{\ell_1,q_1,q} f,\ \iota_{\ell_2,q_1,q} f\rangle_q}{\langle \iota_{1,q_1,q} f,\ \iota_{1,q_1,q} f\rangle_q} = A_f\left(\frac{\ell_2}{(\ell_1,\ell_2)}\right)\overline{A_f}\left(\frac{\ell_1}{(\ell_1,\ell_2)}\right),$$

where $A_f(n)$ is the multiplicative function defined on prime powers by

$$A_f(p^t) = \begin{cases} \dfrac{\lambda_f(p)}{\sqrt{p}(1+\chi_{0(q_1)}(p)p^{-1})} & \text{if } t = 1, \\ \dfrac{\lambda_f(p^t)-\chi_{(q_1)}(p)\lambda_f(p^{t-2})p^{-1}}{p^{t/2}(1+\chi_{0(q_1)}(p)p^{-1})} & \text{if } t \geq 2, \end{cases}$$

where $\chi_{0(q_1)}$ denotes the principal character modulo q_1 and $\chi_{(q_1)} := \chi\,\chi_{0(q_1)}$ denotes the Dirichlet character modulo q_1 induced from χ.

Proof. For $\Re(s) > 1$, consider the integral

$$F(s) := \int_{\Gamma_0(q)\backslash\mathbb{H}} f(\ell_1 z)\bar{f}(\ell_2 z)E(z,s)\,d\mu(z), \quad \text{where } E(z,s) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(q)} \Im(\gamma z)^s.$$

Unfolding the integral and using Parseval's identity,

$$F(s) = \int_0^\infty y^{s-1} \sum_{\substack{n_1=-\infty \\ n_1\neq 0}}^{\infty} \sum_{\substack{n_2=-\infty \\ n_2\neq 0 \\ \ell_1 n_1 = \ell_2 n_2}}^{\infty} \rho_f(n_1)\overline{\rho_f}(n_2) W_{\mathrm{sgn}(n_1)\kappa/2,it_f}(4\pi\ell_1|n_1|y)^2 \,\frac{dy}{y}.$$

From (3.9) and the fact from [Duke et al. 2002, Equation (4.70)] that

$$\rho_f(-n) = \epsilon_f \frac{\Gamma((1+\kappa)/2+it_f)}{\Gamma((1-\kappa)/2+it_f)}\rho_f(n)$$

for $n \geq 1$, where $\epsilon_f \in \{-1, 1\}$, we find that

$$F(s) = \frac{|\rho_f(1)|^2}{(4\pi[\ell_1,\ell_2])^{s-1}\sqrt{\ell'\ell''}} \sum_{n=1}^{\infty} \frac{\lambda_f(\ell'' n)\overline{\lambda_f}(\ell' n)}{n^s}$$
$$\times \int_0^\infty y^{s-1}\left(W_{\kappa/2,it_f}(y)^2 + \left|\frac{\Gamma((1+\kappa)/2+it_f)}{\Gamma((1-\kappa)/2+it_f)}\right|^2 W_{-\kappa/2,it_f}(y)^2\right)\frac{dy}{y},$$

where we have written $n_1 = \ell'' n,\ n_2 = \ell' n$, with $\ell' = \ell_1/(\ell_1,\ell_2)$ and $\ell'' = \ell_2/(\ell_1,\ell_2)$.

Next, by the multiplicativity of the Hecke eigenvalues of f together with the fact that $(\ell', \ell'') = 1$, the sum over n is equal to

$$\sum_{\substack{n=1\\(n,\ell'\ell'')=1}}^{\infty} \frac{|\lambda_f(n)|^2}{n^s} \prod_{p^t \| \ell''} \sum_{r=0}^{\infty} \frac{\lambda_f(p^{r+t})\overline{\lambda_f}(p^r)}{p^{rs}} \prod_{p^t \| \ell'} \sum_{r=0}^{\infty} \frac{\lambda_f(p^r)\overline{\lambda_f}(p^{r+t})}{p^{rs}}.$$

From (3.8) and (3.10), we find that

$$\sum_{r=0}^{\infty} \frac{\lambda_f(p^{r+t})\overline{\lambda_f}(p^r)}{p^{rs}} = B_f(p^t; s) \sum_{r=0}^{\infty} \frac{|\lambda_f(p^r)|^2}{p^{rs}},$$
$$\sum_{r=0}^{\infty} \frac{\lambda_f(p^r)\overline{\lambda_f}(p^{r+t})}{p^{rs}} = \overline{B_f}(p^t; \bar{s}) \sum_{r=0}^{\infty} \frac{|\lambda_f(p^r)|^2}{p^{rs}},$$

where $B_f(n; s)$ is defined to be the multiplicative function

$$B_f(p^t; s) = \begin{cases} \dfrac{\lambda_f(p)}{1 + \chi_{0(q_1)}(p)p^{-s}} & \text{if } t = 1, \\ \dfrac{\lambda_f(p^t) - \chi_{(q_1)}(p)\lambda_f(p^{t-2})p^{-s}}{1 + \chi_{0(q_1)}(p)p^{-s}} & \text{if } t \geq 2, \end{cases}$$

so that $A_f(n) = n^{-\frac{1}{2}} B_f(n; 1)$. We surmise that $F(s)$ is equal to

$$\frac{|\rho_f(1)|^2}{(4\pi[\ell_1, \ell_2])^{s-1}\sqrt{\ell'\ell''}} B_f(\ell''; s)\overline{B_f}(\ell'; \bar{s}) \sum_{n=1}^{\infty} \frac{|\lambda_f(n)|^2}{n^s}$$
$$\times \int_0^{\infty} y^{s-1} \left(W_{\kappa/2, it_f}(y)^2 + \left| \frac{\Gamma((1+\kappa)/2 + it_f)}{\Gamma((1-\kappa)/2 + it_f)} \right|^2 W_{-\kappa/2, it_f}(y)^2 \right) \frac{dy}{y}. \quad (3.14)$$

The result follows by taking the residue at $s = 1$, noting that $E(z, s)$ has residue equal to $1/\operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})$ at $s = 1$ independently of $z \in \Gamma_0(q)\backslash\mathbb{H}$, and comparing to the case $\ell_1 = \ell_2 = 1$. □

Lemma 3.15 (cf. [Blomer and Milićević 2015, Lemma 9]). *An orthonormal basis of $\mathcal{A}_\kappa(q, \chi)$ is given by*

$$\mathcal{B}_\kappa(q, \chi) = \bigsqcup_{\substack{q_1 q_2 = q\\ q_1 \equiv 0 \pmod{q_\chi}}} \bigsqcup_{f \in \mathcal{B}_\kappa^*(q_1, \chi)} \bigsqcup_{d \mid q_2} \left\{ f_d = \sum_{\ell \mid d} \xi_f(\ell, d) \iota_{\ell, q_1, q} f \right\}, \quad (3.16)$$

where each $f \in \mathcal{B}_\kappa^*(q_1, \chi)$ is normalised such that $\langle \iota_{1,q_1,q} f, \iota_{1,q_1,q} f\rangle_q = 1$, and the function $\xi_f(\ell, d)$ is jointly multiplicative.

$$\text{For } 0 \le r \le t, \quad \xi_f(p^r, p^t) = \begin{cases} 1 & \text{if } r = t = 0, \\ -\dfrac{\overline{A_f}(p)}{\sqrt{1-|A_f(p)|^2}} & \text{if } r = 0 \text{ and } t = 1, \\ \dfrac{1}{\sqrt{1-|A_f(p)|^2}} & \text{if } r = t = 1, \\ \dfrac{\overline{\chi}_{(q_1)}(p)}{p}\dfrac{1}{\sqrt{(1-\chi_{0(q_1)}(p)p^{-2})(1-|A_f(p)|^2)}} & \text{if } r = t-2 \text{ and } t \ge 2, \\ -\dfrac{\overline{\lambda_f}(p)}{\sqrt{p}}\dfrac{1}{\sqrt{(1-\chi_{0(q_1)}(p)p^{-2})(1-|A_f(p)|^2)}} & \text{if } r = t-1 \text{ and } t \ge 2, \\ \dfrac{1}{\sqrt{(1-\chi_{0(q_1)}(p)p^{-2})(1-|A_f(p)|^2)}} & \text{if } r = t \text{ and } t \ge 2, \\ 0 & \text{if } 0 \le r \le t-3 \text{ and } t \ge 3. \end{cases}$$

The key point is that the coefficients $\xi_f(\ell, d)$ are chosen such that the ratio of inner products

$$\delta_f(d_1, d_2) := \frac{\langle f_{d_1}, f_{d_2}\rangle_q}{\langle \iota_{1,q_1,q} f, \iota_{1,q_1,q} f\rangle_q} = \sum_{\ell_1 \mid d_1}\sum_{\ell_2 \mid d_2} \xi_f(\ell_1, d_1)\overline{\xi_f}(\ell_2, d_2)\frac{\langle \iota_{\ell_1,q_1,q} f, \iota_{\ell_2,q_1,q} f\rangle_q}{\langle \iota_{1,q_1,q} f, \iota_{1,q_1,q} f\rangle_q}$$

is equal to 1 if $d_1 = d_2$ and 0 otherwise.

Proof. The proof follows the same lines as [Blomer and Milićević 2015, Proof of Lemma 9]; we omit the details. □

Explicit Kuznetsov formula. We may use the explicit basis (3.16) together with (3.10) and (3.9) to rewrite the discrete part of the Kuznetsov formula, noting that for $f \in \mathcal{B}_\kappa^*(q_1, \chi)$, $d \mid q_2$, and $n \ge 1$ coprime to q,

$$\rho_{f_d}(n) = \xi_f(1, d)\rho_f(1)\frac{\lambda_f(n)}{\sqrt{n}}.$$

Similarly, the continuous part can be rewritten in terms of the Eisenstein spanning set $\mathcal{B}(\chi_1, \chi_2)$ with $\chi_1\chi_2 = \chi$ together with (3.2) and (3.3). This yields the following explicit versions of the pre-Kuznetsov and Kuznetsov formulæ.

Proposition 3.17. *When $m, n \ge 1$ with $(mn, q) = 1$, the pre-Kuznetsov formula has the form*

$$\begin{aligned} &\sum_{\substack{q_1q_2=q \\ q_1 \equiv 0 \pmod{q_\chi}}} \sum_{f \in \mathcal{B}_\kappa^*(q_1,\chi)} 4\pi\xi_f|\rho_f(1)|^2 \frac{\overline{\chi}(m)\lambda_f(m)\lambda_f(n)}{\cosh\pi(r-t_f)\cosh\pi(r+t_f)} \\ &\quad + \sum_{\substack{\chi_1,\chi_2 \pmod{q} \\ \chi_1\chi_2=\chi}} \sum_{f \in \mathcal{B}(\chi_1,\chi_2)} \int_{-\infty}^{\infty} |\rho_f(1,t)|^2 \frac{\overline{\chi}(m)\lambda_f(m,t)\lambda_f(n,t)}{\cosh\pi(r-t)\cosh\pi(r+t)}\,dt \\ &\qquad = \frac{|\Gamma(1-\kappa/2-ir)|^2}{\pi^2}\Biggl(\delta_{mn} + \sum_{\substack{c=1 \\ c\equiv 0 \pmod{q}}}^{\infty} \frac{S_\chi(m,n;c)}{c} I_\kappa\biggl(\frac{4\pi\sqrt{mn}}{c}, r\biggr)\Biggr) \end{aligned} \tag{3.18}$$

for $\kappa \in \{0, 1\}$, *where we define*

$$\xi_f := \sum_{d \mid q_2} |\xi_f(1, d)|^2,$$

while the Kuznetsov formula for $\kappa = 0$ *has the form*

$$\begin{aligned}\sum_{\substack{q_1 q_2 = q \\ q_1 \equiv 0 \ (\mathrm{mod}\ q_\chi)}} \sum_{f \in \mathcal{B}_0^*(q_1, \chi)} \frac{4\pi \xi_f |\rho_f(1)|^2}{\cosh \pi t_f} \overline{\chi}(m) \lambda_f(m) \lambda_f(n) h(t_f) \\ + \sum_{\substack{\chi_1, \chi_2 \ (\mathrm{mod}\ q) \\ \chi_1 \chi_2 = \chi}} \sum_{f \in \mathcal{B}(\chi_1, \chi_2)} \int_{-\infty}^{\infty} \frac{|\rho_f(1, t)|^2}{\cosh \pi t} \overline{\chi}(m) \lambda_f(m, t) \lambda_f(n, t) h(t) \, dt \\ = \delta_{mn} g_0 + \sum_{\substack{c = 1 \\ c \equiv 0 \ (\mathrm{mod}\ q)}}^{\infty} \frac{S_\chi(m, n; c)}{c} g_0\left(\frac{4\pi \sqrt{mn}}{c}\right). \end{aligned} \tag{3.19}$$

In both formulæ, each $f \in \mathcal{B}_\kappa^*(q_1, \chi)$ *is normalised such that* $\langle \iota_{1, q_1, q} f, \iota_{1, q_1, q} f \rangle_q = 1$.

Atkin–Lehner decomposition for $\mathbf{\Gamma_1(q)}$. We recall the decomposition

$$\mathcal{A}_\kappa(\Gamma_1(q)) = \bigoplus_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi(-1) = (-1)^\kappa}} \mathcal{A}_\kappa(q, \chi),$$

which follows from the fact that $\Gamma_1(q)$ is a normal subgroup of $\Gamma_0(q)$ with quotient group isomorphic to $(\mathbb{Z}/q\mathbb{Z})^\times$, noting that $\mathcal{A}_\kappa(q, \chi) = \{0\}$ if $\chi(-1) \neq (-1)^\kappa$. From this, we obtain the natural basis of $\mathcal{A}_\kappa(\Gamma_1(q))$ given by

$$\mathcal{B}_\kappa(\Gamma_1(q)) = \bigsqcup_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi(-1) = (-1)^\kappa}} \bigsqcup_{\substack{q_1 q_2 = q \\ q_1 \equiv 0 \ (\mathrm{mod}\ q_\chi)}} \bigsqcup_{f \in \mathcal{B}_\kappa^*(q_1, \chi)} \bigsqcup_{d \mid q_2} \left\{ f_d = \sum_{\ell \mid d} \xi_f(\ell, d) \iota_{\ell, q_1, q} f \right\}. \tag{3.20}$$

This allows us to use the pre-Kuznetsov and Kuznetsov formulæ (3.18) and (3.19) for $\mathcal{B}_\kappa(\Gamma_1(q))$ and $\mathcal{B}_0(\Gamma_1(q))$, even though ostensibly these two formulæ are only set up for $\mathcal{B}_\kappa(q, \chi)$ and $\mathcal{B}_0(q, \chi)$.

Atkin–Lehner decomposition for $\mathbf{\Gamma(q)}$. A similar decomposition also holds for $\mathcal{A}_\kappa(\Gamma(q))$. In this case, the fact that

$$\begin{aligned}\Gamma_0(q^2) \cap \Gamma_1(q) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : a, d \equiv 1 \ (\mathrm{mod}\ q), \ c \equiv 0 \ (\mathrm{mod}\ q^2) \right\} \\ &= \begin{pmatrix} q^{-1} & 0 \\ 0 & 1 \end{pmatrix} \Gamma(q) \begin{pmatrix} q & 0 \\ 0 & 1 \end{pmatrix}\end{aligned}$$

implies that

$$\mathcal{A}_\kappa(\Gamma(q)) = \iota_{q^{-1}} \mathcal{A}_\kappa\big(\Gamma_0(q^2) \cap \Gamma_1(q)\big),$$

where $\iota_{q^{-1}} : \mathcal{A}_\kappa\big(\Gamma_0(q^2) \cap \Gamma_1(q)\big) \to \mathcal{A}_\kappa(\Gamma(q))$ is the map $\iota_{q^{-1}} f(z) = f(q^{-1}z)$. As $\Gamma_0(q^2) \cap \Gamma_1(q)$ is a normal subgroup of $\Gamma_0(q^2)$ with quotient group isomorphic to $(\mathbb{Z}/q\mathbb{Z})^\times$, we obtain the decomposition

$$\mathcal{A}_\kappa(\Gamma(q)) = \bigoplus_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi(-1)=(-1)^\kappa}} \iota_{q^{-1}} \mathcal{A}_\kappa(q^2, \chi),$$

thereby allowing us to choose an explicit basis $\mathcal{B}_\kappa(\Gamma(q))$ of $\mathcal{A}_\kappa(\Gamma(q))$ of the form

$$\bigsqcup_{\substack{\chi \ (\mathrm{mod}\ q) \\ \chi(-1)=(-1)^\kappa}} \bigsqcup_{\substack{q_1 q_2 = q^2 \\ q_1 \equiv 0 \ (\mathrm{mod}\ q_\chi)}} \bigsqcup_{f \in \mathcal{B}_\kappa^*(q_1, \chi)} \bigsqcup_{d \mid q_2} \Big\{ \iota_{q^{-1}} f_d = \sum_{\ell \mid d} \xi_f(\ell, d) \iota_{q^{-1}} \iota_{\ell, q_1, q} f \Big\}. \tag{3.21}$$

Once again, this allows us to make use of the pre-Kuznetsov and Kuznetsov formulæ (3.18) and (3.19) for $\mathcal{B}_\kappa(\Gamma(q))$ and $\mathcal{B}_0(\Gamma(q))$.

4. Bounds for Fourier coefficients of newforms

In the Kuznetsov formula (3.19), the Fourier coefficients $|\rho_f(1)|^2$ and the normalisation factor ξ_f both appear naturally. To remove these weights, we obtain lower bounds for $|\rho_f(1)|^2$ and ξ_f. For the former, such bounds are well-known, appearing in some generality in [Duke et al. 2002, Equation (7.16)]; nevertheless, we take this opportunity to correct some of the minor numerical errors in this proof, as well as greatly streamline the proof via the recent work of Li [2010] on obtaining upper bounds for L-functions at the edge of the critical strip.

Lemma 4.1. *For $f \in \mathcal{B}_\kappa^*(q_1, \chi)$, we have that*

$$\xi_f = \sum_{n \mid q_2^\infty} \frac{|\lambda_f(n)|^2}{n} \prod_{p \| q_2} \left(1 - \frac{\chi_{0(q_1)}(p)}{p^2}\right).$$

In particular, $\xi_f \gg 1$.

Proof. By multiplicativity,

$$\xi_f := \sum_{d \mid q_2} |\xi_f(1, d)|^2 = \prod_{p^t \| q_2} \sum_{r=0}^{t} |\xi_f(1, p^r)|^2.$$

We have that

$$\sum_{r=0}^{t} |\xi_f(1, p^r)|^2 = \begin{cases} 1 & \text{if } t = 0, \\ \dfrac{1}{1 - |A_f(p)|^2} & \text{if } t = 1, \\ \dfrac{1}{(1 - \chi_{0(q_1)}(p) p^{-2})(1 - |A_f(p)|^2)} & \text{if } t \geq 2. \end{cases}$$

The result then follows from the fact that

$$\frac{1}{1 - |A_f(p)|^2} = \left(1 - \frac{\chi_{0(q_1)}(p)}{p^2}\right) \sum_{k=0}^{\infty} \frac{|\lambda_f(p^k)|^2}{p^k}. \qquad \square$$

For $f \in \mathcal{B}_\kappa(q, \chi)$, we define

$$\nu_f := \Gamma\Big(\frac{1+\kappa}{2} + it_f\Big)\Gamma\Big(\frac{1+\kappa}{2} - it_f\Big)|\rho_f(1)|^2.$$

Note that

$$\Gamma\Big(\frac{1+\kappa}{2} + it\Big)\Gamma\Big(\frac{1+\kappa}{2} - it\Big) = \begin{cases} \dfrac{\pi}{\cosh \pi t} & \text{if } \kappa = 0, \\ \dfrac{\pi t}{\sinh \pi t} & \text{if } \kappa = 1. \end{cases}$$

Lemma 4.2. *Suppose that $f \in \mathcal{B}^*_\kappa(q_1, \chi)$ for some $q_1 \mid q$. Then*

$$\frac{\langle \iota_{1,q_1,q} f, \iota_{1,q_1,q} f\rangle_q}{\operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})} = \nu_f \operatorname*{Res}_{s=1} \sum_{n=1}^{\infty} \frac{|\lambda_f(n)|^2}{n^s}.$$

Proof. We let $\ell_1 = \ell_2 = 1$ in (3.14) and take the residue at $s = 1$, yielding

$$\frac{\langle \iota_{1,q_1,q} f, \iota_{1,q_1,q} f\rangle_q}{\operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})} = |\rho_f(1)|^2 \operatorname*{Res}_{s=1} \sum_{n=1}^{\infty} \frac{|\lambda_f(n)|^2}{n^s} \int_0^\infty \bigg(W_{\kappa/2, it_f}(y)^2 + \bigg|\frac{\Gamma((1+\kappa)/2 + it_f)}{\Gamma((1-\kappa)/2 + it_f)}\bigg|^2 W_{-\kappa/2, it_f}(y)^2 \bigg) \frac{dy}{y},$$

since the residue of $E(z, s)$ at $s = 1$ is $1/\operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})$. We have by [Gradshteyn and Ryzhik 2007, 7.611.4] that for $\kappa \in \mathbb{C}$ and $-\frac{1}{2} < \Re(it) < \frac{1}{2}$,

$$\int_0^\infty W_{\kappa/2, it}(y)^2 \frac{dy}{y} = \frac{\pi}{\sin 2\pi it} \frac{\psi((1-\kappa)/2 + it) - \psi((1-\kappa)/2 - it)}{\Gamma((1-\kappa)/2 + it)\Gamma((1-\kappa)/2 - it)},$$

where ψ is the digamma function; note that a slightly erroneous version of this appears in [Duke et al. 2002, Equation (19.6)]. By the gamma and digamma reflection formulæ, we find that

$$\int_0^\infty \bigg(W_{\kappa/2, it_f}(y)^2 + \bigg|\frac{\Gamma((1+\kappa)/2 + it_f)}{\Gamma((1-\kappa)/2 + it_f)}\bigg|^2 W_{-\kappa/2, it_f}(y)^2 \bigg) \frac{dy}{y} = \Gamma\Big(\frac{1+\kappa}{2} + it_f\Big)\Gamma\Big(\frac{1+\kappa}{2} - it_f\Big) \tag{4.3}$$

assuming that $t_f \in [0, \infty)$ if $\kappa = 1$ and $t_f \in [0, \infty)$ or $it_f \in \big(0, \frac{1}{2}\big)$ if $\kappa = 0$. $\square$

Corollary 4.4. *Suppose that $f \in \mathcal{B}^*_\kappa(q_1, \chi)$ for some $q_1 \mid q$. Then*

$$\nu_f \gg_\varepsilon \frac{\langle \iota_{1,q_1,q} f, \iota_{1,q_1,q} f\rangle_q}{\operatorname{vol}(\Gamma_0(q)\backslash\mathbb{H})} (q(3 + t_f^2))^{-\varepsilon}. \tag{4.5}$$

Proof. It is known that

$$\sum_{n=1}^{\infty} \frac{|\lambda_f(n)|^2}{n^s} = \frac{\zeta(s)L(s, \operatorname{ad} f)}{\zeta(2s)} \prod_{p \mid q} P_{f,p}(p^{-s}),$$

where for each prime p dividing q, $P_{f,p}(z)$ is a rational function satisfying $p^{-\varepsilon} \ll_\varepsilon P_{f,p}(p^{-1}) \le 1$. The work of Li [Li 2010, Theorem 2] then shows that

$$L(1, \operatorname{ad} f) \ll \exp\bigg(C \frac{\log(q(3 + t_f^2))}{\log\log(q(3 + t_f^2))} \bigg)$$

for some absolute constant $C > 0$, thereby yielding the result. $\square$

5. Bounds for sums of Kloosterman sums

We denote by

$$S(m,n;c) := \sum_{d\in(\mathbb{Z}/c\mathbb{Z})^\times} e\Big(\frac{md+n\bar{d}}{c}\Big)$$

the usual Kloosterman sum with trivial character, for which the Weil bound holds:

$$|S(m,n;c)| \le \tau(c)\sqrt{(m,n,c)c}. \tag{5.1}$$

We also require bounds for Kloosterman sums with nontrivial character. For $c \equiv 0 \pmod q$, $m, n \ge 1$, and $(a,q) = 1$, we have that

$$\sum_{\substack{\chi \pmod q \\ \chi(-1)=(-1)^\kappa}} \bar{\chi}(a) S_\chi(m,n;c) = \frac{1}{2}\sum_{d\in(\mathbb{Z}/c\mathbb{Z})^\times}\sum_{\chi \pmod q} \bar{\chi}(a)\big(\chi(d)+(-1)^\kappa\chi(-d)\big)e\Big(\frac{md+n\bar{d}}{c}\Big).$$

We break this up into two sums. In the second sum, we can replace d with $-d$ and χ with $\bar{\chi}$ and use character orthogonality to see that

$$\sum_{\substack{\chi \pmod q \\ \chi(-1)=(-1)^\kappa}} \bar{\chi}(a) S_\chi(m,n;c) = \begin{cases} \varphi(q)\Re\big(S_{a(q)}(m,n;c)\big) & \text{if } \kappa = 0, \\ i\varphi(q)\Im\big(S_{a(q)}(m,n;c)\big) & \text{if } \kappa = 1, \end{cases} \tag{5.2}$$

where we set

$$S_{a(q)}(m,n;c) := \sum_{\substack{d\in(\mathbb{Z}/c\mathbb{Z})^\times \\ d\equiv a \pmod q}} e\Big(\frac{md+n\bar{d}}{c}\Big).$$

If $c = c_1c_2$ with $(c_1,c_2) = 1$ and $c_1c_2 \equiv 0 \pmod q$, then we let $d = c_2\overline{c_2}d_1 + c_1\overline{c_1}d_2$, where $d_1 \in (\mathbb{Z}/c_1\mathbb{Z})^\times$, $d_2 \in (\mathbb{Z}/c_2\mathbb{Z})^\times$, and $c_2\overline{c_2} \equiv 1 \pmod{c_1}$, $c_1\overline{c_1} \equiv 1 \pmod{c_2}$. By the Chinese remainder theorem,

$$S_{a(q)}(m,n;c) = S_{a((q,c_1))}(m\overline{c_2}, n\overline{c_2}; c_1) S_{a((q,c_2))}(m\overline{c_1}, n\overline{c_1}; c_2).$$

To bound $S_{a(q)}(m,n;c)$, it therefore suffices to find bounds for $S_{a(p^\alpha)}(m,n;p^\beta)$ for any prime p and any $\beta \ge \alpha \ge 1$. The trivial bound is merely

$$|S_{a(p^\alpha)}(m,n;p^\beta)| \le p^{\beta-\alpha}. \tag{5.3}$$

Somewhat surprisingly, this is sufficient for our needs. Indeed, we cannot do better than this when $\beta = \alpha$, and in our applications, this will be the dominant contribution.

We also require bounds for $S_\chi(m,n;c)$. Unfortunately, it is not necessarily the case that this is bounded by $\tau(c)\sqrt{(m,n,c)c}$, which can be observed numerically at [LMFDB 2013]; see also [Knightly and Li 2013, Example 9.9].

Lemma 5.4. *Let p be an odd prime, let χ_{p^γ} be a Dirichlet character of conductor p^γ, and suppose that $(mn, p) = 1$. Then for $\beta \ge \gamma \ge 0$, we have that*

$$|S_{\chi_{p^\gamma}}(m, n;\, p^\beta)| \le 2p^{\beta/2}$$

unless $\beta = \gamma \ge 3$, in which case we only have that

$$|S_{\chi_{p^\gamma}}(m, n;\, p^\beta)| \le 2p^{\lfloor (3\beta+1)/4 \rfloor}.$$

Similarly, let χ_{2^γ} be a Dirichlet character of conductor 2^γ, and suppose that $(mn, 2) = 1$. Then for $\beta \ge \gamma \ge 0$, we have that

$$|S_{\chi_{2^\gamma}}(m, n;\, 2^\beta)| \le 8 \cdot 2^{\beta/2}$$

unless $\gamma + 1 \ge \beta \ge 3$, in which case we only have that

$$|S_{\chi_{2^\gamma}}(m, n;\, 2^\beta)| \le 4 \cdot 2^{\lfloor (3\beta+1)/4 \rfloor}.$$

Proof. This follows from [Knightly and Li 2013, Propositions 9.4, 9.7, 9.8, and Lemmata 9.6]. □

Lemma 5.5. *When $(m, n) = 1$, we have that*

$$\sum_{\substack{c \le 4\pi\sqrt{mn} \\ c \equiv 0 \ (\mathrm{mod}\ q)}} \frac{|S_{a(q)}(m, n;\, c)|}{c^{\frac{3}{2}}} \ll \frac{(\log(mn+1))^2}{q^{\frac{3}{2}}} \prod_{p \mid q} \frac{1}{1 - p^{-\frac{1}{2}}}, \tag{5.6}$$

$$\sum_{\substack{c \le 4\pi\sqrt{mn} \\ c \equiv 0 \ (\mathrm{mod}\ q^2)}} \frac{|S_{a(q)}(m, n;\, c)|}{c^{\frac{3}{2}}} \ll \frac{(\log(mn+1))^2}{q^2} \prod_{p \mid q} \frac{1}{1 - p^{-\frac{1}{2}}}. \tag{5.7}$$

If we additionally assume that $(mn, q) = 1$, then given a Dirichlet character χ modulo q, we have that

$$\sum_{\substack{c \le 4\pi\sqrt{mn} \\ c \equiv 0 \ (\mathrm{mod}\ q)}} \frac{|S_\chi(m, n;\, c)|}{c^{\frac{3}{2}}} \ll (\log(mn+1))^2 \frac{2^{\omega(q)} \dot{Q}}{\varphi(q)}. \tag{5.8}$$

Proof. We write $q = p_1^{\alpha_1} \cdots p_\ell^{\alpha_\ell}$, so that the left-hand side of (5.6) is

$$\sum_{\beta_1=\alpha_1}^{\infty} \cdots \sum_{\beta_\ell=\alpha_\ell}^{\infty} \frac{1}{(p_1^{\beta_1} \cdots p_\ell^{\beta_\ell})^{\frac{3}{2}}} \sum_{\substack{c \le 4\pi\sqrt{mn}\, p_1^{-\beta_1} \cdots p_\ell^{-\beta_\ell} \\ (c,q)=1}} \frac{1}{c^{\frac{3}{2}}} \times |S(m\overline{p_1^{\beta_1} \cdots p_\ell^{\beta_\ell}}, n\overline{p_1^{\beta_1} \cdots p_\ell^{\beta_\ell}};\, c)|\,|S_{a(q)}(m\bar{c}, n\bar{c};\, p_1^{\beta_1} \cdots p_\ell^{\beta_\ell})|.$$

Using the Weil bound (5.1) for the first Kloosterman sum and the trivial bound (5.3) for the second, we find that this is bounded by

$$\frac{1}{q} \sum_{\beta_1=\alpha_1}^{\infty} \cdots \sum_{\beta_\ell=\alpha_\ell}^{\infty} \frac{1}{\sqrt{p_1^{\beta_1} \cdots p_\ell^{\beta_\ell}}} \sum_{\substack{c \le 4\pi\sqrt{mn} \\ (c,q)=1}} \frac{\tau(c)\sqrt{(m, n, c)}}{c}.$$

If $(m,n)=1$, the inner sum is bounded by a constant multiple of $(\log(mn+1))^2$, and so the sum is bounded by a constant multiple of

$$\frac{(\log(mn+1))^2}{q}\sum_{\beta_1=\alpha_1}^{\infty}\cdots\sum_{\beta_\ell=\alpha_\ell}^{\infty}\frac{1}{\sqrt{p_1^{\beta_1}\cdots p_\ell^{\beta_\ell}}},$$

which yields (5.6) upon evaluating these geometric series. (5.7) follows similarly. Finally, (5.8) follows via the same method but using Lemma 5.4 to bound the Kloosterman sums, yielding the bound

$$8\cdot 2^{\omega(q)}\dot{Q}\sum_{\beta_1=\alpha_1}^{\infty}\cdots\sum_{\beta_\ell=\alpha_\ell}^{\infty}\frac{1}{p_1^{\beta_1}\cdots p_\ell^{\beta_\ell}}\sum_{\substack{c\le 4\pi\sqrt{mn}\\(c,q)=1}}\frac{\tau(c)}{c}$$

for the left-hand side of (5.8), from which the result easily follows. □

Lemma 5.9. *When $(m,n)=1$, we have that*

$$\sum_{\substack{c>4\pi\sqrt{mn}\\ c\equiv 0 \pmod q}}\frac{|S_{\mathfrak{a}(q)}(m,n;c)|}{c^2}\left(1+\log\frac{c}{4\pi\sqrt{mn}}\right)\ll\frac{(\log(mn+1))^2}{(mn)^{\frac14}}\frac{1}{q^{\frac32}}\prod_{p\mid q}\frac{1}{1-p^{-\frac12}},\tag{5.10}$$

$$\sum_{\substack{c>4\pi\sqrt{mn}\\ c\equiv 0 \pmod{q^2}}}\frac{|S_{\mathfrak{a}(q)}(m,n;c)|}{c^2}\left(1+\log\frac{c}{4\pi\sqrt{mn}}\right)\ll\frac{(\log(mn+1))^2}{(mn)^{\frac14}}\frac{1}{q^{2}}\prod_{p\mid q}\frac{1}{1-p^{-\frac12}}.\tag{5.11}$$

If we additionally assume that $(mn,q)=1$, then given a Dirichlet character χ modulo q, we have that

$$\sum_{\substack{c>4\pi\sqrt{mn}\\ c\equiv 0 \pmod q}}\frac{|S_{\chi}(m,n;c)|}{c^2}\left(1+\log\frac{c}{4\pi\sqrt{mn}}\right)\ll\frac{(\log(mn+1))^2}{(mn)^{\frac14}}\frac{2^{\omega(q)}\dot{Q}}{\varphi(q)}.\tag{5.12}$$

Proof. As before, with $q=p_1^{\alpha_1}\cdots p_\ell^{\alpha_\ell}$, the left-hand side of (5.10) is bounded by

$$\frac{1}{q}\sum_{\beta_1=\alpha_1}^{\infty}\cdots\sum_{\beta_\ell=\alpha_\ell}^{\infty}\frac{1}{p_1^{\beta_1}\cdots p_\ell^{\beta_\ell}}\sum_{\substack{c>4\pi\sqrt{mn}\,p_1^{-\beta_1}\cdots p_\ell^{-\beta_\ell}\\(c,q)=1}}\frac{\tau(c)\sqrt{(m,n,c)}\log c}{c^{\frac32}}.$$

If $(m,n)=1$, then the inner sum is bounded by a constant multiple of

$$\frac{(\log(mn+1))^2}{(mn)^{\frac14}}\sqrt{p_1^{\beta_1}\cdots p_\ell^{\beta_\ell}}.$$

It follows that the sum is bounded by a constant multiple of

$$\frac{(\log(mn+1))^2}{(mn)^{\frac14}}\frac{1}{q}\sum_{\beta_1=\alpha_1}^{\infty}\cdots\sum_{\beta_\ell=\alpha_\ell}^{\infty}\frac{1}{\sqrt{p_1^{\beta_1}\cdots p_\ell^{\beta_\ell}}},$$

which gives (5.10). The proof of (5.11) is analogous, while (5.12) again follows upon using Lemma 5.4 to bound the Kloosterman sums. □

Lemma 5.13 (cf. [Iwaniec and Kowalski 2004, Equation (16.50)]). *For all* $\frac{1}{2} < \sigma < 1$,

$$\sum_{\substack{c=1 \\ c\equiv 0 \pmod q}}^{\infty} \frac{|S_{a(q)}(m,n;c)|}{c^{1+\sigma}} \leq \frac{18\tau((m,n))}{(2\sigma-1)^2} \frac{1}{q^{1+\sigma}} \prod_{p|q} \frac{1}{1-p^{-\sigma}}, \tag{5.14}$$

$$\sum_{\substack{c=1 \\ c\equiv 0 \pmod{q^2}}}^{\infty} \frac{|S_{a(q)}(m,n;c)|}{c^{1+\sigma}} \leq \frac{18\tau((m,n))}{(2\sigma-1)^2} \frac{1}{q^{1+2\sigma}} \prod_{p|q} \frac{1}{1-p^{-\sigma}}. \tag{5.15}$$

If we additionally assume that $(m,n) = (mn,q) = 1$, *then given a Dirichlet character* χ *modulo* q, *we have that*

$$\sum_{\substack{c=1 \\ c\equiv 0 \pmod q}}^{\infty} \frac{|S_{\chi}(m,n;c)|}{c^{1+\sigma}} \leq \frac{72}{(2\sigma-1)^2} \frac{2^{\omega(q)}\dot{Q}}{\varphi(q)q^{\sigma-\frac{1}{2}}}. \tag{5.16}$$

Proof. Once again writing $q = p_1^{\alpha_1}\cdots p_\ell^{\alpha_\ell}$ and bounding the Kloosterman sums, we have that

$$\begin{aligned}
\sum_{\substack{c=1 \\ c\equiv 0 \pmod q}}^{\infty} \frac{|S_{a(q)}(m,n;c)|}{c^{1+\sigma}} &\leq \sum_{\substack{c=1 \\ (c,q)=1}}^{\infty} \frac{\tau(c)\sqrt{(m,n,c)}}{c^{\frac{1}{2}+\sigma}} \frac{1}{q} \sum_{\beta_1=\alpha_1}^{\infty} \cdots \sum_{\beta_\ell=\alpha_\ell}^{\infty} \frac{1}{(p_1^{\beta_1}\cdots p_\ell^{\beta_\ell})^{\sigma}} \\
&= \sum_{\substack{c=1 \\ (c,q)=1}}^{\infty} \frac{\tau(c)\sqrt{(m,n,c)}}{c^{\frac{1}{2}+\sigma}} \frac{1}{q^{1+\sigma}} \prod_{p|q} \frac{1}{1-p^{-\sigma}} \\
&\leq \zeta\left(\sigma+\tfrac{1}{2}\right)^2 \sum_{d|(m,n)} \frac{\tau(d)}{d^{\sigma}} \frac{1}{q^{1+\sigma}} \prod_{p|q} \frac{1}{1-p^{-\sigma}} \\
&\leq \frac{18\tau((m,n))}{(2\sigma-1)^2} \frac{1}{q^{1+\sigma}} \prod_{p|q} \frac{1}{1-p^{-\sigma}}.
\end{aligned}$$

This proves (5.14). The inequality (5.15) follows by a similar argument, as does (5.16) once the Kloosterman sums are bounded via Lemma 5.4. □

6. Bounds for test functions

We require bounds for the test function that we will obtain by multiplying the pre-Kuznetsov formula (3.18) by a function dependent on r and then integrating both sides over $r \in [0,T]$.

Lemma 6.1. *For* $T \geq 1$, *let*

$$\begin{aligned}
h_{\kappa,T}(t) &:= \frac{\pi^2}{\Gamma((1+\kappa)/2+it)\Gamma((1+\kappa)/2-it)} \int_0^T \frac{r|\Gamma(1-\kappa/2+ir)|^{-2}}{\cosh\pi(r-t)\cosh\pi(r+t)}\,dr \\
&= \begin{cases} \cosh\pi t \displaystyle\int_0^T \frac{\sinh\pi r}{\cosh\pi(r-t)\cosh\pi(r+t)}\,dr & \text{if } \kappa = 0, \\ \dfrac{\sinh\pi t}{t} \displaystyle\int_0^T \frac{r\cosh\pi r}{\cosh\pi(r-t)\cosh\pi(r+t)}\,dr & \text{if } \kappa = 1. \end{cases}
\end{aligned}$$

Then $h_{\kappa,T}(t)$ is positive for all $t \in \mathbb{R}$ and additionally, should κ be equal to 0, *for $it \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. Furthermore, $h_{\kappa,T}(t) \gg 1$ for $t \in [0, T]$.*

Proof. Using the fact that

$$\cosh \pi(r-t) \cosh \pi(r+t) = \cosh^2 \pi t + \sinh^2 \pi r = \sinh^2 \pi t + \cosh^2 \pi r,$$

it is clear that $h_{\kappa,T}(t)$ is positive for all $t \in \mathbb{R}$ and additionally, should κ be equal to 0, if $it \in \left(-\frac{1}{2}, \frac{1}{2}\right)$.

For $\kappa = 0$, we have that

$$\begin{aligned} h_{0,T}(t) &= \frac{\cosh \pi t}{\pi} \int_1^{\cosh \pi T} \frac{1}{x^2+\sinh^2 \pi t}\, dx \\ &= \frac{\coth \pi t}{\pi} \arctan \frac{\sinh \pi t(\cosh \pi T - 1)}{\sinh^2 \pi t + \cosh \pi T}, \end{aligned}$$

where the second line follows from the arctangent subtraction formula. The first expression shows that $h_{0,T}(t) \gg 1$ when t is small, while when t is large, the argument of arctan is essentially

$$\frac{e^{\pi(T+t)} - e^{\pi t}}{e^{2\pi t} + e^{\pi T}},$$

and this is bounded from below provided that $t \le T$, so that again $h_{0,T}(t) \gg 1$.

For $\kappa = 1$, we can similarly show via integration by parts that

$$\begin{aligned} h_{1,T}(t) &= \frac{\sinh \pi t}{\pi^2 t} \int_0^{\sinh \pi T} \frac{\operatorname{arsinh} x}{x^2+\cosh^2 \pi t}\, dx \\ &= \frac{\tanh \pi t}{\pi^2 t} \int_0^{\sinh \pi T} \frac{\arctan(\sinh \pi T/\cosh \pi t) - \arctan(x/\cosh \pi t)}{\sqrt{x^2+1}}\, dx. \end{aligned}$$

The first expression shows that $h_{1,T}(t) \gg 1$ when t is small, while when t is large, we break up the second expression into two integrals: one from 0 to $\sinh \frac{\pi t}{2}$ and one from $\sinh \frac{\pi t}{2}$ to $\sinh \pi T$. Trivially bounding the numerator in each integral, we find that

$$\begin{aligned} h_{1,T}(t) &\ge \frac{\tanh \pi t}{2\pi} \bigl(\arctan(\sinh \pi T/\cosh \pi t) - \arctan(\sinh(\pi t/2)/\cosh \pi t)\bigr) \\ &= \frac{\tanh \pi t}{2\pi} \arctan \frac{\cosh \pi t\bigl(\sinh \pi T - \sinh(\pi t/2)\bigr)}{\cosh^2 \pi t + \sinh \pi T \sinh(\pi t/2)}. \end{aligned}$$

The argument of arctan is essentially

$$\frac{e^{\pi(T+t)} - e^{3\pi t/2}}{e^{2\pi t} + e^{\pi(T+t/2)}},$$

and this is bounded from below provided that $t \le T$, while $\tanh \pi t$ is bounded from below provided that t is larger than some fixed constant. It follows again that $h_{1,T}(t) \gg 1$. □

We also require the following bound, which arises from the Kloosterman term in the pre-Kuznetsov formula (3.18).

Lemma 6.2. *For $\kappa \in \{0, 1\}$ and $T > 0$, we have the bound*

$$\int_0^T r I_\kappa(a, r)\, dr \ll \begin{cases} \sqrt{a} & \text{if } a \geq 1, \\ a(1 + \log(1/a)) & \text{if } 0 < a < 1 \end{cases} \tag{6.3}$$

uniformly in T.

Proof. From [Kuznetsov 1980, Equation (5.13)], we have that

$$\int_0^T r I_0(a, r)\, dr = a \int_0^\infty \frac{\tanh \xi}{\xi} (1 - \cos 2T\xi) \sin(a \cosh \xi)\, d\xi.$$

Similarly, using the fact that

$$K_{2ir}(\zeta) = \int_0^\infty e^{-\zeta \cosh \xi} \cos 2r\xi \, d\xi$$

for $r \in \mathbb{R}$ and $\Re(\zeta) > 0$ from [Gradshteyn and Ryzhik 2007, 8.432.1], we have that

$$\int_0^T r I_1(a, r)\, dr = -2a \int_0^\infty \int_0^T r \cos 2r\xi \, dr \int_{-i}^{i} e^{-\zeta a \cosh \xi}\, d\zeta \, d\xi.$$

Evaluating each of the inner integrals and then integrating by parts, we find that

$$\int_0^T r I_1(a, r)\, dr \\ = ia \int_0^\infty \frac{\tanh \xi}{\xi} (1 - \cos 2T\xi) \cos(a \cosh \xi)\, d\xi - i \int_0^\infty \frac{\tanh \xi}{\xi} (1 - \cos 2T\xi) \frac{\sin(a \cosh \xi)}{\cosh \xi}\, d\xi.$$

From here, one can show via stationary phase on subintervals of $(0, \infty)$ that $\int_0^T r I_0(a, r)\, dr$ and the first term in the above expression for $\int_0^T r I_1(a, r)\, dr$ both are bounded by a constant multiple of

$$\begin{cases} \sqrt{a} & \text{if } a \geq 1, \\ a(1 + \log(1/a)) & \text{if } 0 < a < 1; \end{cases}$$

see [Kuznetsov 1980, Equation (5.14)]. The second term in the expression for $\int_0^T r I_1(a, r)\, dr$ is uniformly bounded for $a \geq 1$, so we need only consider when $0 < a < 1$. In this case, the fact that $|\sin x| \leq \min\{1, |x|\}$ for $x \in \mathbb{R}$ implies that this is bounded by

$$2a \int_0^{\log(1/a)} \frac{\tanh \xi}{\xi}\, d\xi + 2 \int_{\log(1/a)}^\infty \frac{\tanh \xi}{\xi} \frac{1}{\cosh \xi}\, d\xi \ll a(1 + \log(1/a)). \qquad \square$$

7. Sarnak's density theorem for exceptional Hecke eigenvalues

We are now in a position to prove Theorem 1.1.

Proof of (1.2). By Rankin's trick,

$$\#\{f \in \mathcal{B}_\kappa(\Gamma_1(q)) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\} \leq \prod_{p \in \mathcal{P}} \alpha_p^{-2\ell_p} \sum_{\substack{f \in \mathcal{B}_\kappa(\Gamma_1(q)) \\ t_f \in [0, T]}} \prod_{p \in \mathcal{P}} |\lambda_f(p)|^{2\ell_p}$$

for any nonnegative integers ℓ_p to be chosen. Using the explicit basis (3.20) of $\mathcal{A}_\kappa(\Gamma_1(q))$ together with the lower bound (4.5) for ν_f,

$$\sum_{\substack{f\in\mathcal{B}_\kappa(\Gamma_1(q))\\ t_f\in[0,T]}}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}=\sum_{\substack{\chi \pmod q\\ \chi(-1)=(-1)^\kappa}}\sum_{\substack{q_1q_2=q\\ q_1\equiv 0 \pmod{q_\chi}}}\sum_{\substack{f\in\mathcal{B}_\kappa^*(q_1,\chi)\\ t_f\in[0,T]}}\tau(q_2)\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}$$
$$\ll_\varepsilon q^{1+\varepsilon}T^\varepsilon\sum_{\substack{\chi \pmod q\\ \chi(-1)=(-1)^\kappa}}\sum_{\substack{q_1q_2=q\\ q_1\equiv 0 \pmod{q_\chi}}}\sum_{\substack{f\in\mathcal{B}_\kappa^*(q_1,\chi)\\ t_f\in[0,T]}}\xi_f\nu_f\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}.$$

We take $m=1$ and $n=\prod_{p\in\mathcal{P}}p^{2j_p}$ in the pre-Kuznetsov formula (3.18), multiply both sides by $\prod_{p\in\mathcal{P}}\alpha_{2j_p,2\ell_p}\overline{\chi}(p)^{j_p}$, and sum over all $0\le j_p\le\ell_p$, over all $p\in\mathcal{P}$, and over all Dirichlet characters χ modulo q satisfying $\chi(-1)=(-1)^\kappa$. We then multiply both sides by $\pi^2 r|\Gamma(1-\kappa/2+ir)|^{-2}$ and integrate both sides with respect to r from 0 to T.

On the spectral side, (3.1), (3.5), and Lemma 6.1 allow us to use positivity to discard the contribution from the continuous spectrum, while we may discard the contribution of the discrete spectrum with $t\notin[0,T]$ via (3.8), (3.12), and Lemma 6.1, so that the spectral side is bounded from below by a constant multiple of

$$\sum_{\substack{\chi \pmod q\\ \chi(-1)=(-1)^\kappa}}\sum_{\substack{q_1q_2=q\\ q_1\equiv 0 \pmod{q_\chi}}}\sum_{\substack{f\in\mathcal{B}_\kappa^*(q_1,\chi)\\ t_f\in[0,T]}}\xi_f\nu_f\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}.$$

On the geometric side, we only pick up the delta term when $j_p=0$ for all $p\in\mathcal{P}$, in which case the term is bounded by a constant multiple of $qT^2\prod_{p\in\mathcal{P}}\alpha_{0,2\ell_p}$. For $\kappa=0$, we use (5.2) to write the Kloosterman term in the form

$$\frac{\varphi(q)}{\pi}\sum_{\substack{j_p=0\\ p\in\mathcal{P}}}^{\ell_p}\prod_{p\in\mathcal{P}}\alpha_{2j_p,2\ell_p}\sum_{\substack{c=1\\ c\equiv 0 \pmod q}}^{\infty}\frac{\Re\big(S_{\prod_{p\in\mathcal{P}}p^{j_p}(q)}\big(1,\prod_{p\in\mathcal{P}}p^{2j_p};c\big)\big)}{c}\int_0^T r I_0\bigg(\frac{4\pi\prod_{p\in\mathcal{P}}p^{j_p}}{c},r\bigg)\,dr.$$

For $\kappa=1$, the Kloosterman term is the same except with $i\Im$ in place of $\Re$ and I_1 in place of I_0. In either case, we bound the integral via (6.3), which allows us to use (5.6) and (5.10) to bound the summation over c, so that the Kloosterman term is bounded by a constant multiple of

$$\frac{1}{\sqrt{q}}\prod_{p'\mid q}\frac{1}{1-p'^{-\frac12}}\sum_{\substack{j_p=0\\ p\in\mathcal{P}}}^{\ell_p}\prod_{p\in\mathcal{P}}\alpha_{2j_p,2\ell_p}p^{j_p/2}\bigg(\log\bigg(\prod_{p\in\mathcal{P}}p^{2j_p}+1\bigg)\bigg)^2.$$

We bound the summation over j_p and over $p\in\mathcal{P}$ via (3.7), thereby obtaining

$$\#\big\{f\in\mathcal{B}_\kappa(\Gamma_1(q)):t_f\in[0,T],\ |\lambda_f(p)|\ge\alpha_p \text{ for all } p\in\mathcal{P}\big\}$$
$$\ll_\varepsilon q^{1+\varepsilon}T^\varepsilon\prod_{p\in\mathcal{P}}\bigg(\frac{\alpha_p}{2}\bigg)^{-2\ell_p}\bigg(qT^2+\frac{\prod_{p\in\mathcal{P}}p^{\ell_p/2}\big(\log\prod_{p\in\mathcal{P}}p^{\ell_p/2}\big)^2}{\sqrt{q}}\prod_{p'\mid q}\frac{1}{1-p'^{-\frac12}}\bigg).$$

It remains to take

$$\ell_p = \left\lfloor \frac{\mu_p \log(\mathrm{vol}(\Gamma_1(q)\backslash\mathbb{H})^{\frac{3}{2}} T^4)}{\log p} \right\rfloor. \qquad \square$$

Proof of (1.3). We use (3.21), (5.7), and (5.11) in place of (3.20), (5.6), and (5.10), thereby finding that

$$\#\big\{f \in \mathcal{B}_\kappa(\Gamma(q)) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\big\} \leq \prod_{p\in\mathcal{P}} \alpha_p^{-2\ell_p} \sum_{\substack{f\in\mathcal{B}_\kappa(\Gamma(q))\\ t_f\in[0,T]}} \prod_{p\in\mathcal{P}} |\lambda_f(p)|^{2\ell_p},$$

with

$$\begin{aligned}\sum_{\substack{f\in\mathcal{B}_\kappa(\Gamma(q))\\ t_f\in[0,T]}} \prod_{p\in\mathcal{P}} |\lambda_f(p)|^{2\ell_p} &= \sum_{\substack{\chi \pmod q\\ \chi(-1)=(-1)^\kappa}} \sum_{\substack{q_1q_2=q^2\\ q_1\equiv 0 \pmod{q_\chi}}} \sum_{\substack{f\in\mathcal{B}_\kappa^*(q_1,\chi)\\ t_f\in[0,T]}} \tau(q_2) \prod_{p\in\mathcal{P}} |\lambda_f(p)|^{2\ell_p} \\ &\ll_\varepsilon q^{2+\varepsilon} T^\varepsilon \prod_{p\in\mathcal{P}} 2^{2\ell_p} \left(qT^2 + \frac{\prod_{p\in\mathcal{P}} p^{\ell_p/2} \big(\log \prod_{p\in\mathcal{P}} p^{\ell_p/2}\big)^2}{q} \prod_{p'|q} \frac{1}{1-p'^{-\frac{1}{2}}} \right).\end{aligned}$$

Taking

$$\ell_p = \left\lfloor \frac{\mu_p \log(\mathrm{vol}(\Gamma(q)\backslash\mathbb{H})^{\frac{4}{3}} T^4)}{\log p} \right\rfloor$$

completes the proof. $\square$

Proof of (1.4). Using (3.16), (5.8), and (5.12) in place of (3.20), (5.6), and (5.10),

$$\begin{aligned}&\#\big\{f \in \mathcal{B}_\kappa(q,\chi) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\big\} \\ &\qquad\ll_\varepsilon q^{1+\varepsilon} T^\varepsilon \prod_{p\in\mathcal{P}} \left(\frac{\alpha_p}{2}\right)^{-2\ell_p} \left(T^2 + \prod_{p\in\mathcal{P}} p^{\ell_p/2} \Big(\log \prod_{p\in\mathcal{P}} p^{\ell_p/2}\Big)^2 \frac{2^{\omega(q)} \dot{Q}}{\varphi(q)} \right).\end{aligned}$$

Upon taking

$$\ell_p = \left\lfloor \frac{\mu_p \log(\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})^2 T^4 \dot{Q}^{-2})}{\log p} \right\rfloor,$$

we conclude that

$$\begin{aligned}&\#\big\{f \in \mathcal{B}_\kappa(q,\chi) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\big\} \\ &\qquad\ll_\varepsilon \big(\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H}) T^2\big)^{1-4\sum_{p\in\mathcal{P}} \mu_p (\log \alpha_p/2)/\log p + \varepsilon} \dot{Q}^{4\sum_{p\in\mathcal{P}} \mu_p (\log\alpha_p/2)/\log p}. \qquad (7.1)\end{aligned}$$

On the other hand, by the inclusion $\mathcal{A}_\kappa(q,\chi) \subset \mathcal{A}_\kappa(q\ddot{Q},\chi)$,

$$\begin{aligned}&\#\big\{f \in \mathcal{B}_\kappa(q,\chi) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\big\} \\ &\qquad\leq \#\big\{f \in \mathcal{B}_\kappa(q\ddot{Q},\chi) : t_f \in [0, T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\big\}.\end{aligned}$$

Since $q_{\chi\psi^2} \mid q_\chi$, we have that $\dot{Q}(q\ddot{Q}, q_{\chi\psi^2}) = 1$. Consequently, (7.1) yields the bound

$$\begin{aligned}\#\{f \in \mathcal{B}_\kappa(q,\chi) : t_f \in [0,T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\}& \\ \ll_\varepsilon \big(\mathrm{vol}(\Gamma_0(q\ddot{Q})\backslash\mathbb{H})T^2\big)^{1-4\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p+\varepsilon}& \\ \ll_\varepsilon \big(\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})T^2\big)^{1-4\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p+\varepsilon}\ddot{Q}^{1-4\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p}&. \qquad \square\end{aligned}$$

Remark 7.2. Should we wish to improve (1.4) to be uniform in $\mathcal{P}$, then one needs to take into account the fact that

$$\begin{aligned}\prod_{p\in\mathcal{P}}\left(\frac{\alpha_p}{2}\right)^{-2\ell_p} = \big(\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})T^2\big)^{-4\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p+\varepsilon}\dot{Q}^{4\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p}& \\ \times\prod_{p\in\mathcal{P}}\left(\frac{\alpha_p}{2}\right)^{2\left(\mu_p\log(\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})^2T^4\dot{Q}^{-2})\right)/\log p}&,\end{aligned}$$

where $\{x\}$ denotes the fractional part of x, and the last term need not necessarily be $\ll_\varepsilon (\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})T^2)^\varepsilon$. For this reason, [Blomer et al. 2014, Proposition 1] is not correct in the generality in which it is stated, namely the claim that the result is uniform for $T > p$. Instead, one requires that $p \ll_\varepsilon T^\varepsilon$.

8. Huxley's density theorem for exceptional laplacian eigenvalues

Theorem 1.5 is proved similarly to Theorem 1.1, though we use the Kuznetsov formula (3.19) with a carefully chosen test function in place of the pre-Kuznetsov formula (3.18), and we require different methods to bound the Kloosterman term.

Proof of (1.6). We again use Rankin's trick with nonnegative integers ℓ_p and a positive real number $X \geq 1$ to be chosen:

$$\begin{aligned}\#\big\{f \in \mathcal{B}_0(\Gamma_1(q)) : it_f \in \big(\alpha_0, \tfrac{1}{2}\big),\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\big\}& \\ \leq X^{-2\alpha_0}\prod_{p\in\mathcal{P}}\alpha_p^{-2\ell_p}\sum_{\substack{f\in\mathcal{B}_0(\Gamma_1(q)) \\ it_f\in(0,\frac{1}{2})}} X^{2it_f}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}&.\end{aligned}$$

Again using (3.20) and (4.5),

$$\begin{aligned}\sum_{\substack{f\in\mathcal{B}_0(\Gamma_1(q)) \\ it_f\in(0,\frac{1}{2})}} X^{2it_f}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p} &= \sum_{\substack{\chi \pmod q \\ \chi(-1)=1}}\ \sum_{\substack{q_1q_2=q \\ q_1\equiv 0 \pmod{q_\chi}}}\ \sum_{\substack{f\in\mathcal{B}_0^*(q_1,\chi) \\ it_f\in(0,\frac{1}{2})}} \tau(q_2)X^{2it_f}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p} \\ &\ll_\varepsilon q^{1+\varepsilon}\sum_{\substack{\chi \pmod q \\ \chi(-1)=1}}\ \sum_{\substack{q_1q_2=q \\ q_1\equiv 0 \pmod{q_\chi}}}\ \sum_{\substack{f\in\mathcal{B}_0^*(q_1,\chi) \\ it_f\in(0,\frac{1}{2})}} \xi_f\nu_f X^{2it_f}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}\end{aligned}$$

We take $m = 1$, $n = \prod_{p\in\mathcal{P}} p^{2j_p}$, and

$$h(t) = h_X(t) = \left(\frac{X^{it}+X^{-it}}{t^2+1}\right)^2$$

in the Kuznetsov formula (3.19), multiply both sides by $\prod_{p\in\mathcal{P}} \alpha_{2j_p,2\ell_p}\overline{\chi}(p)^{j_p}$, and sum over all $0\le j_p\le \ell_p$, over all $p\in\mathcal{P}$, and over all even Dirichlet characters modulo q. On the spectral side, we discard all but the discrete spectrum for which $it_f\in\left(0,\frac{1}{2}\right)$ via positivity, so that the spectral side is bounded from below by a constant multiple of

$$\sum_{\substack{\chi \pmod q\\ \chi(-1)=1}} \sum_{\substack{q_1q_2=q\\ q_1\equiv 0 \pmod{q_\chi}}} \sum_{\substack{f\in\mathcal{B}_0^*(q_1,\chi)\\ it_f\in(0,\frac{1}{2})}} \xi_f\nu_f X^{2it_f}\prod_{p\in\mathcal{P}}|\lambda_f(p)|^{2\ell_p}.$$

We only pick up the delta term on the geometric side when $j_p=0$ for all $p\in\mathcal{P}$, in which case the term is bounded by a constant multiple of $q\prod_{p\in\mathcal{P}}2^{2\ell_p}$. We write the Kloosterman term in the form

$$\frac{\varphi(q)}{2\pi i}\sum_{\substack{j_p=0\\ p\in\mathcal{P}}}^{\ell_p}\prod_{p\in\mathcal{P}}\alpha_{2j_p,2\ell_p}\int_{\sigma-i\infty}^{\sigma+i\infty}\sum_{\substack{c=1\\ c\equiv 0 \pmod q}}^{\infty}\frac{\Re\left(S_{\prod_{p\in\mathcal{P}}p^{j_p}(q)}\left(1,\prod_{p\in\mathcal{P}}p^{2j_p};c\right)\right)}{c}J_s\left(\frac{4\pi\prod_{p\in\mathcal{P}}p^{j_p}}{c}\right)\frac{sh_X(is/2)}{\cos(\pi s/2)}\,ds$$

for any $\frac{1}{2}<\sigma<1$. We have, via [Gradshteyn and Ryzhik 2007, 8.411.4], the bound

$$J_s(x)\ll\frac{x^\sigma}{\left|\Gamma\left(s+\frac{1}{2}\right)\right|}\ll e^{\pi|s|/2}\left(\frac{x}{|s|}\right)^{\sigma},$$

and so the integral in the Kloosterman term is bounded by a constant multiple of

$$\prod_{p\in\mathcal{P}}p^{j_p\sigma}\sum_{\substack{c=1\\ c\equiv 0 \pmod q}}^{\infty}\frac{\left|S_{\prod_{p\in\mathcal{P}}p^{j_p}(q)}\left(1,\prod_{p\in\mathcal{P}}p^{2j_p};c\right)\right|}{c^{1+\sigma}}\int_{\sigma/2-i\infty}^{\sigma/2+i\infty}|r^{\frac{3}{4}}h_X(ir)|\,dr.$$

We take

$$\sigma=\frac{1}{2}+\frac{1}{\log\left(X\prod_{p\in\mathcal{P}}p^{\ell_p}\right)},$$

so that the integral is bounded by a constant multiple of $\sqrt{X}$, and use (5.14) to bound the summation over c and (3.7) to bound the summation over j_p and $p\in\mathcal{P}$ in order to find that

$$\begin{aligned}&\#\left\{f\in\mathcal{B}_0(\Gamma_1(q)):it_f\in\left(\alpha_0,\tfrac{1}{2}\right),\ |\lambda_f(p)|\ge\alpha_p \text{ for all } p\in\mathcal{P}\right\}\\ &\qquad\ll_\varepsilon q^{1+\varepsilon}X^{-2\alpha_0}\prod_{p\in\mathcal{P}}\left(\frac{\alpha_p}{2}\right)^{-2\ell_p}\left(q+\sqrt{X}\prod_{p\in\mathcal{P}}p^{\ell_p/2}\left(\log\left(X\prod_{p\in\mathcal{P}}p^{\ell_p}\right)\right)^2\frac{1}{\sqrt{q}}\prod_{p'\mid q}\frac{1}{1-p'^{-\frac{1}{2}}}\right).\end{aligned}$$

The result follows upon taking

$$X=\operatorname{vol}(\Gamma_1(q)\backslash\mathbb{H})^{3\mu_0/2},\qquad \ell_p=\left\lfloor\frac{\mu_p\log\operatorname{vol}(\Gamma_1(q)\backslash\mathbb{H})^{\frac{3}{2}}}{\log p}\right\rfloor. \qquad\square$$

Proof of (1.7). By using (3.21) and (5.15) in place of (3.20) and (5.14), we obtain

$$\#\{f \in \mathcal{B}_0(\Gamma(q)) : it_f \in (\alpha_0, \tfrac{1}{2}),\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\} \\ \ll_\varepsilon q^{2+\varepsilon} X^{-2\alpha_0} \prod_{p\in\mathcal{P}} \left(\frac{\alpha_p}{2}\right)^{-2\ell_p} \left(q + \sqrt{X} \prod_{p\in\mathcal{P}} p^{\ell_p/2} \left(\log\left(X \prod_{p\in\mathcal{P}} p^{\ell_p}\right)\right)^2 \frac{1}{q} \prod_{p'\mid q} \frac{1}{1-p'^{-\frac{1}{2}}}\right),$$

and it remains to take

$$X = \mathrm{vol}(\Gamma(q)\backslash\mathbb{H})^{4\mu_0/3}, \quad \ell_p = \left\lfloor \frac{\mu_p \log \mathrm{vol}(\Gamma(q)\backslash\mathbb{H})^{\frac{4}{3}}}{\log p} \right\rfloor. \qquad \square$$

Proof of (1.8). We use (3.16) and (5.16) in place of (3.20) and (5.14), so that

$$\#\{f \in \mathcal{B}_0(q,\chi) : it_f \in (\alpha_0, \tfrac{1}{2}),\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\} \\ \ll_\varepsilon q^{1+\varepsilon} X^{-2\alpha_0} \prod_{p\in\mathcal{P}} \left(\frac{\alpha_p}{2}\right)^{-2\ell_p} \left(1 + \sqrt{X} \prod_{p\in\mathcal{P}} p^{\ell_p/2} \left(\log\left(X \prod_{p\in\mathcal{P}} p^{\ell_p}\right)\right)^2 \frac{2^{\omega(q)}\dot{Q}}{\varphi(q)}\right).$$

We find that

$$\#\{f \in \mathcal{B}_0(q,\chi) : it_f \in (\alpha_0, \tfrac{1}{2}),\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\} \\ \ll_\varepsilon \mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})^{1-4(\mu_0\alpha_0+\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p)+\varepsilon} \dot{Q}^{4(\mu_0\alpha_0+\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p)}.$$

by taking

$$X = \mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})^{2\mu_0} \dot{Q}^{-2\mu_0}, \quad \ell_p = \left\lfloor \frac{\mu_p \log(\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})^2 \dot{Q}^{-2})}{\log p} \right\rfloor.$$

Again, we also have that

$$\#\{f \in \mathcal{B}_0(q,\chi) : it_f \in (\alpha_0, \tfrac{1}{2}),\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\} \\ \leq \#\{f \in \mathcal{B}_0(q\ddot{Q}, \chi\psi^2) : it_f \in (\alpha_0, \tfrac{1}{2}),\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\}$$

for any primitive character ψ modulo $\ddot{Q}$, which implies that

$$\#\{f \in \mathcal{B}_0(q,\chi) : it_f \in (\alpha_0, \tfrac{1}{2}) \in [0,T],\ |\lambda_f(p)| \geq \alpha_p \text{ for all } p \in \mathcal{P}\} \\ \ll_\varepsilon \mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H})^{1-4(\mu_0\alpha_0+\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p)+\varepsilon} \ddot{Q}^{1-4(\mu_0\alpha_0+\sum_{p\in\mathcal{P}}\mu_p(\log\alpha_p/2)/\log p)}. \qquad \square$$

9. Improving Theorems 1.1 and 1.5 for $\Gamma_1(q)$ via twisting

In this section, we prove Theorem 1.9 Let $f \in \mathcal{B}_\kappa^*(q,\chi)$ be a newform, and for a primitive character ψ modulo q_ψ with $q_\psi \mid q$, we let $f \otimes \psi$ denote the twist of f by ψ; this is the *newform* whose Hecke eigenvalues $\lambda_{f\otimes\psi}(n)$ are equal to $\lambda_f(n)\psi(n)$ whenever $(n,q)=1$. By [Atkin and Li 1978, Proposition 3.1], the weight of $f \otimes \psi$ is κ, the level of $f \otimes \psi$ divides q^2, and the nebentypus is the primitive character that induces $\chi\psi^2$. We make crucial use of the fact that twisting by a Dirichlet character preserves the Laplacian eigenvalue $\lambda_f = \frac{1}{4} + t_f^2$ and the absolute value $|\lambda_f(n)|$ of the Hecke eigenvalues of f for all

$(n, q) = 1$. Moreover, if $f_1 \in \mathcal{B}^*_\kappa(q_1, \chi_1)$, $f_2 \in \mathcal{B}^*_\kappa(q_2, \chi_2)$ are such that there exist primitive Dirichlet characters ψ_1 modulo q_{ψ_1} and ψ_2 modulo q_{ψ_2} with $q_{\psi_1}, q_{\psi_2} \mid q$ such that

$$f_1 \otimes \psi_1 = f_2 \otimes \psi_2,$$

then $f_2 = f_1 \otimes \psi_1 \overline{\psi_2}$.

Lemma 9.1. *If q is squarefree, ψ is a primitive Dirichlet modulo q_ψ, where $q_\psi \mid q$, and $f \in \mathcal{B}^*_\kappa(q, \chi)$, then the level of $f \otimes \psi$ divides q if and only if $\overline{\psi}$ divides χ, in the sense that $\psi\chi$ has conductor dividing q_χ.*

Proof. This follows via the methods of [Humphries 2017]. For $p \mid q$, let π_p be the local component of the cuspidal automorphic representation π of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ associated to the newform f, so that the central character ω_p of π_p is the local component of the Hecke character ω that is the idèlic lift of χ. As q is squarefree, π_p is either a principal series representation or a special representation.

In the former case, $\pi_p = \omega_{p,1} \boxplus \omega_{p,2}$ with central character $\omega_p = \omega_{p,1}\omega_{p,2}$, where $\omega_{p,1}, \omega_{p,2}$ are characters of $\mathbb{Q}_p^\times$ with conductor exponents $c(\omega_{p,1}), c(\omega_{p,2}) \in \{0, 1\}$ such that the conductor exponent $c(\pi_p)$ of π_p is $c(\omega_{p,1}) + c(\omega_{p,2}) = 1$. The twist $\pi_p \otimes \omega'_p$ of π_p by a character ω'_p of $\mathbb{Q}_p^\times$ of conductor exponent $c(\omega'_p) \in \{0, 1\}$ is $\omega_{p,1}\omega'_p \boxplus \omega_{p,2}\omega'_p$ with corresponding conductor exponent $c(\pi_p \otimes \omega'_p) = c(\omega_{p,1}\omega'_p) + c(\omega_{p,2}\omega'_p)$. For this to be at most 1, either ω'_p is unramified, or one of $c(\omega_{p,1}\omega'_p), c(\omega_{p,2}\omega'_p)$ must be equal to 0, so that $\overline{\omega'_p}$ is equal to $\omega_{p,1}$ or $\omega_{p,2}$ up to multiplication by an unramified character.

In the latter case, $\pi_p = \omega_{p,1}\,\mathrm{St}$ with central character $\omega_p = \omega_{p,1}^2$ such that $c(\omega_{p,1}) = 0$, so that $c(\pi_p) = 1$. The twist of π_p by ω'_p is $\omega_{p,1}\omega'_p\,\mathrm{St}$, with corresponding conductor exponent $c(\pi_p \otimes \omega'_p) = \max\{1, 2c(\omega_{p,1}\omega'_p)\}$. For this to be at most 1, ω'_p must be unramified.

It follows that if the Hecke character ω' is the idèlic lift of ψ, then the conductor of $\pi \otimes \omega'$ divides q if and only if the conductor of $\omega'\omega$ divides the conductor of ω. ☐

From this, we have the following.

Corollary 9.2. *Let q be squarefree. Given a newform g of level dividing q^2, there exist at most $\tau(q)$ newforms f of level dividing q that can be twisted by a Dirichlet character of conductor dividing q to give g.*

Proof. Suppose that $f_1 \in \mathcal{B}^*_\kappa(q_1, \chi_1)$ and $f_2 \in \mathcal{B}^*_\kappa(q_2, \chi_2)$ with q_1 and q_2 dividing q are such that there exist Dirichlet characters ψ_1 and ψ_2 of conductors dividing q for which $f_1 \otimes \psi_1 = f_2 \otimes \psi_2 = g$. Then $f_2 = f_1 \otimes \psi_1\overline{\psi_2}$, and Lemma 9.1 implies that $\overline{\psi_1}\psi_2$ divides χ_1. Since the conductor of χ_1 divides q_1, the level of f_1, the proof is complete by noting that the number of Dirichlet characters ψ_2 modulo q for which this may occur is bounded by the number of divisors of q. ☐

Lemma 9.3. *Let q be squarefree, let $\mathcal{P}$ be a finite collection of primes not dividing q, let E_0 be a measurable subset of $[0, \infty) \cup i\big(0, \frac{1}{2}\big)$, and let E_p be a measurable subset of $[0, \infty)$ for each $p \in \mathcal{P}$. Then*

$$\#\big\{f \in \mathcal{B}_\kappa(\Gamma_1(q)) : t_f \in E_0,\ |\lambda_f(p)| \in E_p \text{ for all } p \in \mathcal{P}\big\} \\ \leq \frac{\tau(q)^2}{\varphi(q)} \#\big\{f \in \mathcal{B}_\kappa(\Gamma(q)) : t_f \in E_0,\ |\lambda_f(p)| \in E_p \text{ for all } p \in \mathcal{P}\big\}.$$

Proof. From (3.20),

$$\#\bigl\{f \in \mathcal{B}_\kappa(\Gamma_1(q)) : t_f \in E_0,\ |\lambda_f(p)| \in E_p \text{ for all } p \in \mathcal{P}\bigr\}$$

is equal to

$$\sum_{\substack{\chi \pmod q \\ \chi(-1)=(-1)^\kappa}} \sum_{\substack{q_1q_2=q \\ q_1 \equiv 0 \pmod{q_\chi}}} \tau(q_2)\#\bigl\{f \in \mathcal{B}_\kappa^*(q_1,\chi) : t_f \in E_0,\ |\lambda_f(p)| \in E_p \text{ for all } p \in \mathcal{P}\bigr\},$$

which, in turn, is equal to

$$\frac{1}{\varphi(q)} \sum_{\psi \pmod q} \sum_{\substack{\chi \pmod q \\ \chi(-1)=(-1)^\kappa}} \sum_{\substack{q_1q_2=q \\ q_1 \equiv 0 \pmod{q_\chi}}} \tau(q_2) \\ \times \#\bigl\{f \otimes \psi : f \in \mathcal{B}_\kappa^*(q_1,\chi),\ t_f \in E_0,\ |\lambda_f(p)| \in E_p \text{ for all } p \in \mathcal{P}\bigr\},$$

as twisting preserves Laplacian eigenvalues and the absolute value of Hecke eigenvalues. Each twist $g = f \otimes \psi$ of some $f \in \mathcal{B}_\kappa^*(q_1,\chi)$ is a newform of weight κ, level dividing q^2, and nebentypus of conductor dividing q, and Corollary 9.2 implies that there are at most $\tau(q)$ newforms of level dividing q that can be twisted by a Dirichlet character of conductor dividing q to yield g. Since $\tau(q_2) \le \tau(q)$, the above quantity is bounded by

$$\frac{\tau(q)^2}{\varphi(q)} \sum_{\substack{\chi \pmod q \\ \chi(-1)=(-1)^\kappa}} \sum_{\substack{q_1q_2=q^2 \\ q_1 \equiv 0 \pmod{q_\chi}}} \#\bigl\{g \in \mathcal{B}_\kappa^*(q_1,\chi) : t_g \in E_0,\ |\lambda_g(p)| \in E_p \text{ for all } p \in \mathcal{P}\bigr\},$$

while the explicit basis (3.21) of $\mathcal{B}_\kappa(\Gamma(q))$ implies that

$$\#\{g \in \mathcal{B}_\kappa(\Gamma(q)) : t_g \in E_0,\ |\lambda_g(p)| \in E_p \text{ for all } p \in \mathcal{P}\}$$

is equal to

$$\sum_{\substack{\chi \pmod q \\ \chi(-1)=(-1)^\kappa}} \sum_{\substack{q_1q_2=q^2 \\ q_1 \equiv 0 \pmod{q_\chi}}} \tau(q_2)\#\{g \in \mathcal{B}_\kappa^*(q_1,\chi) : t_g \in E_0,\ |\lambda_g(p)| \in E_p \text{ for all } p \in \mathcal{P}\}.$$

This yields the result. □

Combining this with the fact that $\mathrm{vol}(\Gamma(q)\backslash\mathbb{H}) = q\,\mathrm{vol}(\Gamma_1(q)\backslash\mathbb{H})$, we deduce Theorem 1.9. It is likely that a more careful analysis could obtain this same result even when q is not squarefree via the methods in [Humphries 2017].

Acknowledgements

The author thanks Peter Sarnak for many helpful discussions on this topic, as well as the referee for correcting several mistakes in an earlier version of this paper.

References

[Atkin and Li 1978] A. O. L. Atkin and W. C. W. Li, "Twists of newforms and pseudo-eigenvalues of W-operators", *Invent. Math.* **48**:3 (1978), 221–243. MR Zbl

[Blomer and Milićević 2015] V. Blomer and D. Milićević, "The second moment of twisted modular L-functions", *Geom. Funct. Anal.* **25**:2 (2015), 453–516. MR Zbl

[Blomer et al. 2007] V. Blomer, G. Harcos, and P. Michel, "Bounds for modular L-functions in the level aspect", *Ann. Sci. École Norm. Sup.* (4) **40**:5 (2007), 697–740. MR Zbl

[Blomer et al. 2014] V. Blomer, J. Buttcane, and N. Raulf, "A Sato–Tate law for GL(3)", *Comment. Math. Helv.* **89**:4 (2014), 895–919. MR Zbl

[Booker and Strömbergsson 2007] A. R. Booker and A. Strömbergsson, "Numerical computations with the trace formula and the Selberg eigenvalue conjecture", *J. Reine Angew. Math.* **607** (2007), 113–161. MR Zbl

[Bruggeman and Miatello 2009] R. W. Bruggeman and R. J. Miatello, *Sum formula for* SL_2 *over a totally real number field*, Mem. Amer. Math. Soc. **919**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[Conrey et al. 1997] J. B. Conrey, W. Duke, and D. W. Farmer, "The distribution of the eigenvalues of Hecke operators", *Acta Arith.* **78**:4 (1997), 405–409. MR Zbl

[Duke et al. 2002] W. Duke, J. B. Friedlander, and H. Iwaniec, "The subconvexity problem for Artin L-functions", *Invent. Math.* **149**:3 (2002), 489–577. MR Zbl

[Gradshteyn and Ryzhik 2007] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed., Elsevier, Amsterdam, 2007. MR Zbl

[Hughes and Miller 2007] C. P. Hughes and S. J. Miller, "Low-lying zeros of L-functions with orthogonal symmetry", *Duke Math. J.* **136**:1 (2007), 115–172. MR Zbl

[Humphries 2017] P. Humphries, "Spectral multiplicity for Maaß newforms of non-squarefree level", *Int. Mat. Res. Not.* (online publication December 2017).

[Huxley 1986] M. N. Huxley, "Exceptional eigenvalues and congruence subgroups", pp. 341–349 in *The Selberg trace formula and related topics* (Brunswick, ME, 1984), edited by D. A. Hejhal et al., Contemp. Math. **53**, Amer. Math. Soc., Providence, RI, 1986. MR Zbl

[Iwaniec 2002] H. Iwaniec, *Spectral methods of automorphic forms*, 2nd ed., Graduate Studies in Math. **53**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Iwaniec and Kowalski 2004] H. Iwaniec and E. Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloquium Publications **53**, Amer. Math. Soc., Providence, RI, 2004. MR Zbl

[Iwaniec et al. 2000] H. Iwaniec, W. Luo, and P. Sarnak, "Low lying zeros of families of L-functions", *Inst. Hautes Études Sci. Publ. Math.* **91** (2000), 55–131. MR Zbl

[Kim and Sarnak 2003] H. H. Kim and P. Sarnak, "Refined estimates towards the Ramanujan and Selberg conjectures", (2003). Appendix 2 in H. H. Kim, "Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2", *J. Amer. Math. Soc.* **16**:1 (2003), 139–183. MR Zbl

[Knightly and Li 2013] A. Knightly and C. Li, *Kuznetsov's trace formula and the Hecke eigenvalues of Maass forms*, Mem. Amer. Math. Soc. **1055**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

[Kuznetsov 1980] N. V. Kuznetsov, "The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture: sums of Kloosterman sums", *Mat. Sb.* (*N.S.*) **111(153)**:3 (1980), 334–383. In Russian; translated in *Math. USSR-Sb* **39**:3 (1981), 299–342.

[Li 2010] X. Li, "Upper bounds on L-functions at the edge of the critical strip", *Int. Math. Res. Not.* **2010**:4 (2010), 727–755. MR Zbl

[LMFDB 2013] The LMFDB Collaboration, "The L-functions and modular forms database", electronic reference, 2013, Available at http://www.lmfdb.org.

[Maga 2013] P. Maga, "A semi-adelic Kuznetsov formula over number fields", *Int. J. Number Theory* **9**:7 (2013), 1649–1681. MR Zbl

[Petrow and Young 2018] I. Petrow and M. P. Young, "A generalized cubic moment and the Petersson formula for newforms", *Math. Ann.* (online publication August 2018).

[Sarnak 1987] P. Sarnak, "Statistical properties of eigenvalues of the Hecke operators", pp. 321–331 in *Analytic number theory and Diophantine problems* (Stillwater, OK, 1984), edited by A. C. Adolphson et al., Progr. Math. **70**, Birkhäuser, Boston, 1987. MR Zbl

[Schulze-Pillot and Yenirce 2018] R. Schulze-Pillot and A. Yenirce, "Petersson products of bases of spaces of cusp forms and estimates for Fourier coefficients", *Int. J. Number Theory* **14**:8 (2018), 2277–2290. MR

Communicated by Philippe Michel
Received 2017-01-30 Revised 2018-04-02 Accepted 2018-06-02

pclhumphries@gmail.com *Department of Mathematics, University College London, United Kingdom*

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1611

Irreducible components of minuscule affine Deligne–Lusztig varieties

Paul Hamacher and Eva Viehmann

We examine the set of $J_b(F)$-orbits in the set of irreducible components of affine Deligne–Lusztig varieties for a hyperspecial subgroup and minuscule coweight μ. Our description implies in particular that its number of elements is bounded by the dimension of a suitable weight space in the Weyl module associated with μ of the dual group.

1. Introduction

Let F be a finite extension of $\mathbb{Q}_p$ or $\mathbb{F}_p((t))$ and Γ its absolute Galois group. We denote by $\mathcal{O}_F$ and $k_F \cong \mathbb{F}_q$ its ring of integers and its residue field, and by ϵ a fixed uniformizer. Let L denote the completion of the maximal unramified extension of F, and $\mathcal{O}_L$ its ring of integers. Its residue field is an algebraic closure k of k_F. We denote by σ the Frobenius of L over F and of k over k_F.

Let G be a reductive group scheme over $\mathcal{O}_F$, and denote $K = G(\mathcal{O}_L)$. Then G_F is automatically unramified. We fix $S \subset T \subset B \subset G$, where S is a maximal split torus, T a maximal torus, and B a Borel subgroup of G. Let W be the absolute Weyl group of G. There exist k_F-ind schemes called the loop group LG, the positive loop group L^+G, and the affine Grassmannian $\mathcal{G}r_G := LG/L^+G$ of G whose k-valued points are canonically identified with $G(L)$, $K = G(\mathcal{O}_L)$, and $G(L)/G(\mathcal{O}_L)$, respectively (compare [Pappas and Rapoport 2008; Zhu 2017; Bhatt and Scholze 2017]).

The authors were partially supported by European Research Council starting grant 277889 “Moduli spaces of local G-shtukas” and thank Miaofen Chen and Xinwen Zhu for helpful conversations and in particular for sharing their conjecture describing the $J_b(F)$-orbits of irreducible components in terms of $V_\mu(\lambda)$.
MSC2010: primary 14G35; secondary 20G25.
Keywords: affine Deligne–Lusztig variety, Rapoport–Zink spaces, affine Grassmannian.

Let $\mu \in X_*(T)_{\mathrm{dom}}$, and let $b \in G(L)$. Then the affine Deligne–Lusztig variety associated with b and μ is the reduced subscheme $X_\mu(b)$ of $\mathcal{G}r_G$ whose k-valued points are

$$X_\mu(b)(k) = \{g \in G(L)/K \mid g^{-1}b\sigma(g) \in K\mu(\epsilon)K\}.$$

Let $X_{\preceq\mu}(b) = \bigcup_{\mu' \preceq \mu} X_{\mu'}(b)$ where $\mu' \preceq \mu$ if $\mu - \mu'$ is a nonnegative integral linear combination of positive coroots. It is closed in the affine Grassmannian and called the closed affine Deligne–Lusztig variety. For minuscule μ (the case we are mainly interested in for this paper) it agrees with $X_\mu(b)$.

Notice that up to isomorphism, both affine Deligne–Lusztig varieties depend only on the $G(L)$-σ-conjugacy class $[b] \in B(G)$ of b. An affine Deligne–Lusztig variety $X_\mu(b)$ or $X_{\preceq\mu}(b)$ is nonempty if and only if $[b] \in B(G, \mu)$, a finite subset of $B(G)$. The following basic assertion seems to be well known, but we could not find a reference in the literature.

Lemma 1.1. *The scheme $X_\mu(b)$ is locally of finite type in the equal characteristic case and locally of perfectly finite type in the case of unequal characteristic.*

Proof. The proof of this is the same as the corresponding part of the analogous statement for moduli spaces of local G-shtukas; compare the proof of Theorem 6.3 in [Hartl and Viehmann 2011] (where only the first half of page 113 is needed). In that proof, the case of equal characteristic and split G is considered. However, the general statement follows from the same proof. □

Notice that in general $X_\mu(b)$ is not quasicompact since it may have infinitely many irreducible components. It is conjectured to be equidimensional, but this has not been proven in full generality yet. In Section 3 we give an overview of the cases where equidimensionality has been proven. In the case of μ minuscule, which we are primarily interested in here, there are only a few exceptional cases where this is not yet known.

Definition 1.2. For a finite-dimensional k-scheme X we denote by $\Sigma(X)$ the set of irreducible components of X and by $\Sigma^{\mathrm{top}}(X) \subset \Sigma(X)$ the subset of those irreducible components which are top-dimensional.

The affine Deligne–Lusztig varieties $X_\mu(b)$ and $X_{\preceq\mu}(b)$ carry a natural action (by left multiplication) by the group

$$J_b(F) = \{g \in G(L) \mid g^{-1}b\sigma(g) = b\}.$$

This action induces an action of $J_b(F)$ on the set of irreducible components.

A complete description of the set of orbits was previously only known for the groups GL_n and GSp_{2n} and minuscule μ where the action is transitive [Viehmann 2008a; 2008b], and for some other particular cases; see for example [Vollaard and Wedhorn 2011] for a particular family of unitary groups and minuscule μ.

To describe the (conjectured) number of orbits, denote by $\widehat{G}$ the dual group of G in the sense of Deligne and Lusztig. That is, $\widehat{G}$ is the reductive group scheme over $\mathcal{O}_F$ that contains a Borel subgroup $\widehat{B}$ with maximal torus $\widehat{T}$ and maximal split torus $\hat{S}$ such that there exists a Galois equivariant isomorphism

$X^*(\widehat{T}) \cong X_*(T)$ identifying simple coroots of $\widehat{T}$ with simple roots of T. For any $\mu \in X_*(T)_{\mathrm{dom}} = X^*(\widehat{T})_{\mathrm{dom}}$ we denote by V_μ the associated Weyl module of $\widehat{G}_{\mathbb{O}_L}$.

In the following we use an element $\lambda_G(b) \in X^*(\widehat{T}^\Gamma)$ that we define in Section 2. Its restriction λ to $\hat{S}$ can be seen as a "best integral approximation" of the Newton point ν_b of $[b]$, while its precise value in $X^*(\widehat{T}^\Gamma)$ will depend on the Kottwitz point $\kappa_G(b)$. We choose a lift $\tilde{\lambda} \in X_*(T)$.

Conjecture 1.3 (Chen and Zhu). *There exists a canonical bijection between $J_b(F) \setminus \Sigma(X_\mu(b))$ and the basis of $V_\mu(\lambda_G(b))$ constructed by Mirković and Vilonen* [2007], *where $V_\mu(\lambda_G(b))$ denotes the $\lambda_G(b)$-weight space* (*for the action of $\widehat{T}^\Gamma$*) *of V_μ.*

In this paper, we describe the set $J_b(F) \setminus \Sigma^{\mathrm{top}}(X_\mu(b))$ for minuscule μ. Our main result, Theorem 5.12, implies in particular the following theorem.

Theorem 1.4. *Let $\mu \in X_*(T)_{\mathrm{dom}}$ be minuscule, $b \in [b] \in B(G,\mu)$, and $\tilde{\lambda} \in X_*(T)$ be an associated element as in Section 2. There exists a canonical surjective map*

$$\phi : W.\mu \cap [\tilde{\lambda} + (1-\sigma)X_*(T)] \twoheadrightarrow J_b(F) \setminus \Sigma^{\mathrm{top}}(X_\mu(b)).$$

Moreover, this map is a bijection in the following cases:

(1) *G is split and*

(2) *$[b] \cap \mathrm{Cent}_G(\nu_b)$ is a union of superbasic σ-conjugacy classes in* $\mathrm{Cent}_G(\nu_b)$.

Remark 1.5. (a) Let us explain how the theorem is a special case of the conjecture. Since μ is minuscule, we have for any $\tilde{\mu} \in X_*(T)$

$$\dim V_\mu(\tilde{\mu}) = \begin{cases} 1 & \text{if } \tilde{\mu} \in W.\mu, \\ 0 & \text{otherwise,} \end{cases}$$

where now $V_\mu(\tilde{\mu})$ denotes the $\tilde{\mu}$-weight space for the action of $\widehat{T}$. Thus, indeed we obtain a bijection between the Mirković–Vilonen basis of $V_\mu(\lambda)$ and $W.\mu \cap [\tilde{\lambda} + (1-\sigma)X_*(T)]$.

(b) We can replace the weight space $V_\mu(\lambda_G(b))$ by the weight space $V_\mu(\lambda)$ for the action of $\hat{S}$ in Conjecture 1.3. A priori one might expect the second space to be bigger; the equality is a consequence of the relation between λ and the Kottwitz point $\kappa_G(b)$ (see Remark 2.6 for details).

(c) An analogous formula has first been shown by Xiao and Zhu [2017] for $[b]$ such that the F-ranks of J_b and G coincide. In this case one can simply choose $\lambda = \nu_b$, the Newton point of $[b]$. It was then observed by Chen and Zhu (in oral communication) that an expression similar to the above should give $|J_b(F) \setminus \Sigma(X_\mu(b))|$ also for general $[b]$, and all μ.

(d) In particular, Theorems 1.4 and 5.12 apply to all cases that correspond to Newton strata in Shimura varieties of Hodge type.

In the case where b is superbasic, we prove the following stronger result, which was conjectured in [Hamacher 2015a]. For the ordering $\leq$ compare the definition at the top of page 1615.

Proposition 1.6. *Assume $b \in G(L)$ is superbasic. There exists a decomposition into disjoint $J_b(F)$-stable locally closed subschemes*

$$X_\mu(b) = \bigcup_{\substack{\tilde{\mu} \in W.\mu \\ \tilde{\mu}|_{\hat{S}} \preceq \nu_b}} C_{\tilde{\mu}}$$

such that $C_{\tilde{\mu}}$ intersected with any connected component of $\mathcal{G}r_G$ is universally homeomorphic to an affine space. These affine spaces are of dimension $d(\tilde{\mu}) := \sum \lfloor \langle \tilde{\mu} - \mu_{\mathrm{adom}}, \hat{\omega}_F \rangle \rfloor$ where we take the sum over all relative fundamental coweights $\hat{\omega}_F$ of $\widehat{G}$ and where μ_{adom} denotes the antidominant representative in the Weyl group orbit of μ.

Note that varying b within $[b]$ only changes $X_\mu(b)$ by an isomorphism. For suitably chosen $b \in [b]$, the connected components of $C_{\tilde{\mu}}$ are precisely the intersections of $X_\mu(b)$ with some Iwahori-orbit on $\mathcal{G}r_G$ [Chen and Viehmann 2018, §3]. Since the latter form a stratification on $\mathcal{G}r_G$, we can apply the localization long exact sequence to calculate the cohomology of $X_\mu(b)$. For example for the constant sheaf one obtains the following result.

Corollary 1.7. *Assume $b \in G(L)$ is superbasic, and denote by $J_b(F)^0$ the (unique) parahoric subgroup of $J_b(F)$. Then the $J_b(F)$-equivariant cohomology of $X_\mu(b)$ (for $\ell \neq p$) is given by*

$$H_c^{2i+1}(X_\mu(b), \mathbb{Q}_\ell) = 0,$$
$$H_c^{2i}(X_\mu(b), \mathbb{Q}_\ell) = \text{c-ind}_{J_b(F)^0}^{J_b(F)} V_i,$$

where V_i is a diagonalizable $J_b(F)^0$-representation with coefficients in $\mathbb{Q}_\ell$ and of dimension

$$\#\{\tilde{\mu} \in W.\mu \mid d(\tilde{\mu}) = i\}.$$

2. Definition of λ

We associate with every σ-conjugacy class $[b]$ a not necessarily dominant coinvariant $\lambda_G(b) \in X^*(\widehat{T})_\Gamma$ which lifts the Kottwitz point of b and at the same time is a "best approximation" of the Newton point (in a sense to be made precise below). In the split case it is closely connected to the notion of σ-straight elements in the extended affine Weyl group of G.

Invariants of σ-conjugacy classes. By work of Kottwitz [1985], a σ-conjugacy class $[b] \in B(G)$ is uniquely determined by two invariants: the Newton point $\nu_G(b) \in X_*(S)_{\mathbb{Q},\mathrm{dom}}$ and the Kottwitz point $\kappa_G(b) \in \pi_1(G)_\Gamma$. Here $\pi_1(G)$ denotes Borovoi's fundamental group, i.e., the quotient of $X_*(T)$ by its coroot lattice. We also consider the Kottwitz homomorphism w_G as in [Kottwitz 1985]. Let $w : X_*(T) \twoheadrightarrow \pi_1(G)$ denote the canonical projection. By the Cartan decomposition $G(L) = \coprod_{\mu \in X_*(T)_{\mathrm{dom}}} K\mu(\epsilon)K$, and we extend w to a map $w_G : G(L) \to \pi_1(G)$ mapping $K\mu(\epsilon)K$ to $w(\mu)$. Then for every $b \in G(L)$ the projection of $w_G(b)$ to $\pi_1(G)_\Gamma$ coincides with $\kappa_G(b)$.

We define a partial order $\preceq$ on $X^*(\widehat{T})$ such that $\mu' \preceq \mu$ holds if and only if $\mu - \mu'$ is a linear combination of positive roots with nonnegative, integral coefficients. Since the set of positive roots is preserved by the Galois action, this descends to a partial order on $X^*(\widehat{T})_\Gamma$. Similarly, we define its rational analogue $\leq$ on $X^*(T)_{\mathbb{Q}}$ such that $\mu \leq \mu'$ holds if and only if $\mu - \mu'$ is a linear combination of positive roots with nonnegative, rational coefficients. By the same argument as above this order descends to $X^*(\widehat{T})_{\mathbb{Q},\Gamma} = X^*(\hat{S})$.

Lemma/Definition 2.1. *Let $b \in G(L)$. Then the set*

$$\{\tilde{\lambda} \in X^*(\widehat{T})_\Gamma \mid w(\tilde{\lambda}) = \kappa_G(b),\ \tilde{\lambda}|_{\hat{S}} \leq \nu_G(b)\}$$

has a unique maximum $\lambda_G(b)$ characterized by the property that $w(\lambda_G(b)) = \kappa_G(b)$ and that for every relative fundamental coweight $\omega^\vee_{\widehat{G},F}$ of $\widehat{G}$, one has

$$\langle \lambda_G(b) - \nu_G(b), \omega^\vee_{\widehat{G},F}\rangle \in (-1, 0]. \tag{2.2}$$

Proof. Denote by $\widehat{Q} \subset X^*(\widehat{T})$ the root lattice. Then the restriction $X^*(\widehat{T}) \twoheadrightarrow X^*(\hat{S})$ canonically identifies the relative root lattice with $\widehat{Q}_\Gamma$. Note that the preimage $w^{-1}(\kappa_G(b))_\Gamma$ in $X^*(\widehat{T})_\Gamma$ is a $\widehat{Q}_\Gamma$-coset. Thus, one has $\lambda' \succeq \lambda$ for two elements in $w^{-1}(\kappa_G(b))_\Gamma$ if and only if

$$\langle \lambda', \omega^\vee_{\widehat{G},F}\rangle - \langle \lambda, \omega^\vee_{\widehat{G},F}\rangle \geq 0$$

for all relative fundamental coweights $\omega^\vee_{\widehat{G},F}$ of $\widehat{G}$ and moreover the left-hand side always has integral value. Thus, if a $\lambda_G(b)$ as in (2.2) exists, it is the unique maximum. One easily constructs such a $\lambda_G(b)$ by choosing some $\lambda' \in w^{-1}(\kappa_G(b))_\Gamma$ and defining

$$\lambda_G(b) := \lambda' - \sum_{\hat{\beta}} \lceil \langle \lambda' - \nu_G(b), \omega^\vee_{\hat{\beta}}\rangle \rceil \cdot \hat{\beta},$$

where the sum runs over all positive simple roots $\hat{\beta} \in \widehat{Q}_\Gamma$ and $\omega^\vee_{\hat{\beta}}$ denotes the corresponding fundamental coweight. $\square$

Example 2.3. Assume that $G = \mathrm{GL}_n$, B is the upper-triangular Borel subgroup, and that $S = T$ is the diagonal torus. Then $\lambda_G(b)$ has the following geometric interpretation. To an element $\nu \in \mathbb{Q}^n \cong X_*(\hat{S})_{\mathbb{Q}}$, we associate a polygon $P(\nu)$ which is defined over $[0, n]$ with starting point $(0, 0)$ and slope ν_i over $(i-1, i)$. We denote by f_ν the corresponding piecewise linear function. Then $P(\nu_G(b))$ is the (concave) Newton polygon of b and $P(\lambda_G(b))$ is the largest polygon below $P(\nu_G(b))$ with integral slopes and break points. Indeed, the fundamental coweights of GL_n are given by $\omega_i = (\underbrace{1, \dots, 1}_{i \text{ times}}, \underbrace{0, \dots, 0}_{n-i \text{ times}})$; thus,

$$\langle \lambda_G(b) - \nu_G(b), \omega_i\rangle = f_{\lambda_G(b)}(i) - f_{\nu_G(b)}(i),$$

which implies $f_{\lambda_G(b)}(i) = \lfloor f_{\nu_G(b)}(i) \rfloor$ by (2.2). An example is illustrated in Figure 1.

Lemma 2.4. *Let $f : H \to G$ be a morphism of reductive groups over $\mathcal{O}_F$. Then we have $\lambda_G(f(b)) = f(\lambda_H(b))$ in the following cases*:

(1) *f is a central isogeny and*

(2) *f is the embedding of a standard Levi subgroup, such that $\nu_H(b)$ is G-dominant.*

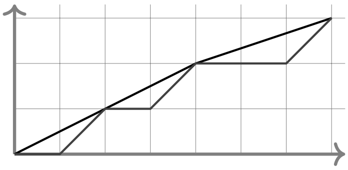

Figure 1. The polygons associated to $\nu_G(b)$ and $\lambda_G(b)$ for $[b] \in B(\mathrm{GL}_7)$ given by $\nu_G(b) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Proof. If f is a central isogeny, we have $X^*(\widehat{T}_H) = X^*(\widehat{T}_G) \times_{\pi_1(G)} \pi_1(H)$ compatibly with the obvious Galois action and partial order on the right-hand side. Thus, f and λ commute.

Now assume that H is a standard Levi subgroup of G and $\nu_H(b)$ is dominant, i.e., $\nu_H(b) = \nu_G(b)$. By (2.2) we have $-1 < \langle f(\lambda_H(b)) - \nu_G(b), \omega^\vee_{\widehat{H},F}\rangle \le 0$ for every relative fundamental coweight of H. Let $\omega^\vee_{\widehat{G},F}$ be a relative fundamental coweight of G, but not of H. Then $\omega^\vee_{\widehat{G},F}$ factorizes through the center of H; thus, for every quasicharacter $\nu' \in X_*(\widehat{T})_{\mathbb{Q}}$ the value of $\langle \nu', \omega^\vee_{\widehat{G},F}\rangle$ is determined by the image of ν' in $\pi_1(H)_{\Gamma,\mathbb{Q}}$. In $\pi_1(H)_{\Gamma,\mathbb{Q}}$ we have equalities

$$(\text{image of } \nu_H(b)) = (\text{image of } \kappa_H(b)) = (\text{image of } \lambda_H(b));$$

thus, $\langle \nu_H(b) - \lambda_H(b), \omega^\vee_{\widehat{G},F}\rangle = 0$. □

Notation 2.5. For fixed $b \in G(L)$ we denote by $\tilde{\lambda} \in X^*(\widehat{T})$ an arbitrary but fixed lift of $\lambda_G(b)$ and by λ its image in $X^*(\hat{S})$.

Remark 2.6. Since G is quasisplit, the maximal torus of the derived group T^{der} is induced and hence $\widehat{T^{\mathrm{der}}}^{\Gamma} \subseteq \hat{S}$. Thus, any two elements in $X^*(\widehat{T}^{\Gamma})$ with the same image in $X^*(\hat{S})$ differ by a central cocharacter and thus have a different image in $\pi_1(G)_\Gamma$. In particular

$$\{\tilde{\mu} \in X^*(\widehat{T}) \mid \tilde{\mu}|_{\widehat{T}^\Gamma} = \lambda_G(b),\ w_G(\tilde{\mu}) = w_G(\mu)\} = \{\tilde{\mu} \in X^*(\widehat{T}) \mid \tilde{\mu}|_{\hat{S}} = \lambda,\ w_G(\tilde{\mu}) = w_G(\mu)\}.$$

Since $V_\mu(\tilde{\mu}) = 0$ unless $\tilde{\mu} \le \mu$, this implies $V_\mu(\lambda_G(b)) = V_\mu(\lambda)$.

A group-theoretic definition of λ_G in the split case. We denote by $\widetilde{W} = \widetilde{W}_G := (\mathrm{Norm}_G(T))(L)/T(O_L)$ the extended affine Weyl group of G. Recall that we have canonical isomorphisms $\widetilde{W}_G \cong X_*(T) \rtimes W \cong W_a \rtimes \Omega_G$ where W_a denotes the affine Weyl group of G and $\Omega_G \subset \widetilde{W}_G$ the set of elements stabilizing the base alcove, which we choose as the unique alcove in the dominant Weyl chamber whose closure contains 0. In particular, we can lift the length function ℓ on W_a to $\widetilde{W}_G$.

The embedding $\mathrm{Norm}_G(T) \hookrightarrow G$ induces a natural map $B(\widetilde{W}_G) \to B(G)$, where $B(\widetilde{W}_G)$ denotes the set of $\widetilde{W}_G$-σ-conjugacy classes in $\widetilde{W}_G$. In general the notion of $\widetilde{W}_G$-conjugacy is finer than the notion of $G(L)$-conjugacy. Hence, we consider only a certain subset of $B(\widetilde{W}_G)$.

Definition 2.7. (1) We call $x \in \widetilde{W}_G$ basic if it is contained in Ω_G. A σ-conjugacy class $O \in B(\widetilde{W}_G)$ is called basic if it contains a basic element.

(2) An element $x \in \widetilde{W}_G$ is called σ-straight if it satisfies

$$\ell(x\sigma(x)\cdots\sigma^{n-1}(x)) = \ell(x) + \ell(\sigma(x)) + \cdots + \ell(\sigma^{n-1}(x)).$$

for any nonnegative integer n. Note that the right-hand side might also be written as $n \cdot \ell(x)$. A σ-conjugacy class $O \in B(\widetilde{W}_G)$ is called straight if it contains a σ-straight element.

He and Nie gave a characterization of the set of straight σ-conjugacy classes which is analogous to Kottwitz's description of $B(G)$ [1985, §6].

Proposition 2.8 [He and Nie 2014, Proposition 3.2]. *A σ-conjugacy class $O \in B(\widetilde{W}_G)$ is straight if and only if it contains a basic σ-conjugacy class $O' \in B(\widetilde{W}_M)$ for some standard Levi subgroup $M \subset G$.*

Finally, by [He and Nie 2014, Theorem 3.3] each $[b] \in B(G)$ contains a unique straight $O_{[b]} \in B(\widetilde{W}_G)$. We obtain the following description of λ_G in the split case.

Proposition 2.9. *Let G be a split group over $\mathcal{O}_F$, let $b \in G(L)$, and let $x \in O_{[b]}$ be a σ-straight element. Denote by λ' its image under the canonical projection $\widetilde{W}_G \to X_*(T)$. Then $\lambda'_{\mathrm{dom}} = \lambda_G(b)_{\mathrm{dom}}$.*

Proof. By Proposition 2.8 there exists a standard Levi subgroup $M \subset G$ and an M-basic element $x_M \in \Omega_M$ such that x and x_M are $\widetilde{W}_G$-conjugate. By [He and Nie 2015, Proposition 4.5] any two such elements are even W-conjugate and thus correspond to the same element in $X_*(T)_{\mathrm{dom}}$. Since the same holds true for $\lambda_G(b)_{\mathrm{dom}}$ by Lemma 2.4, it suffices to prove the proposition in the basic case, i.e., when $\nu_G(b)$ is central.

If $[b]$ is basic, then x is basic; thus, λ' is the (unique) dominant minuscule character with $w(\lambda') = \kappa_G(b)$; compare [Bourbaki 1968, §2, Proposition 6]. Hence, it suffices to show that $\lambda_G(b)$ is minuscule. By Lemma 2.4(2) we may assume that G is of adjoint type. This leaves finitely many cases, which can easily be checked using the explicit description of root systems in [Bourbaki 1968]. ☐

3. Equidimensionality

While it is conjectured that $X_\mu(b)$ is equidimensional [Rapoport 2005, Conjecture 5.10], this has not yet been proven in all cases. We give a partial result after reviewing the necessary geometry of $X_\mu(b)$ first.

Connected components. Let $w_G : G(L) \to \pi_1(G)$ be the Kottwitz homomorphism, as considered in [Kottwitz 1985]; compare the bottom of page 1614. It induces a map $\mathcal{G}r_G(k) \to \pi_1(G)$. After base change to $\operatorname{Spec} k$, this induces isomorphisms $\pi_0(LG_k) \cong \pi_0(\mathcal{G}r_{G,k}) \cong \pi_1(G)$; compare [Pappas and Rapoport 2008, Theorem 0.1] in the equal characteristic case and [Zhu 2017, Proposition 1.21] in the mixed characteristic case. Here we used that as G is unramified, the action of the inertia subgroup of the absolute Galois group of F on $\pi_1(G)$ is trivial.

For $\omega \in \pi_1(G)$, we let LG^ω and $\mathcal{G}r_G^\omega$ be the corresponding connected components. Denote for any subgroup $H \subset LG_k$ and subscheme $X \subset \mathcal{G}r_{G,k}$ the intersection $H^\omega := H \cap LG^\omega$ and $X^\omega := X \cap \mathcal{G}r_G^\omega$.

In particular, $X_\mu(b)^\omega$ is a union of connected components, and the $J_b(F)$-orbit of $X_\mu(b)^\omega$ equals $X_\mu(b)$ by [Nie 2015, Theorem 1.2] (see also [Chen et al. 2015, Theorem 1.2]) whenever $X_\mu(b)^\omega$ is nonempty. One can even show that under some mild condition on the triple $(G, [b], \mu)$ every connected component of $X_\mu(b)$ is of the form $X_\mu(b)^\omega$ (see [Nie 2015, Theorem 1.1] and also [Chen et al. 2015, Theorem 1.1]), but we will not need this result.

The following general result on affine flag varieties is formulated in greater generality than needed in this paper. We will only apply it in the case where $H = G$ is a reductive group scheme. For consistency we denote affine flag varieties by the same symbol $\mathcal{G}r$ as affine Grassmannians.

Proposition 3.1. *Let $f : H' \to H$ be a morphism of parahoric group schemes over $\mathbb{O}_F$ such that the induced homomorphism on their adjoint groups is an isomorphism. Then the induced morphism on connected components of affine flag varieties*

$$f^\omega_{\mathcal{G}r} : \mathcal{G}r^\omega_{H'} \to \mathcal{G}r^{f(\omega)}_H$$

is a universal homeomorphism.

Proof. This is proven in [Pappas and Rapoport 2008, §6] if char $F = p$ and p does not divide the order of $\pi_1(H'_{\mathrm{der}})$ or $\pi_1(H_{\mathrm{der}})$ (see also [He and Zhou 2016, Proposition 4.3] for the statement if char $F = 0$). We briefly recall the proof in [Pappas and Rapoport 2008] and explain how to generalize it.

Note that it suffices to show that $f^\omega_{\mathcal{G}r}$ is bijective on geometric points. Indeed, it is a morphism of ind-proper ind-schemes (see [Richarz 2016, Corollary 2.3] if char $F = p$ and [Zhu 2017, §1.5.2] if char $F = 0$) and thus universally closed.

By homogeneity under the actions of $H'(L)$ and $H(L)$, respectively, we may assume $\omega = 0$. Denote by H_{der} the derived group of H and by $\widetilde{H}$ the simply connected cover of H_{der}. Since we have a commutative diagram

$$\begin{tikzcd} H' \arrow[rrrr, bend left, "f"] & H'_{\mathrm{der}} \arrow[l, hook'] & \widetilde{H}' = \widetilde{H} \arrow[l, two heads] \arrow[r, two heads] & H_{\mathrm{der}} \arrow[r, hook] & H \end{tikzcd}$$

it suffices to prove the theorem in the following two special cases.

Case 1: $H' = H_{\mathrm{der}}$. One can show that $f^0_{\mathcal{G}r}$ is universally bijective using the argument in [Pappas and Rapoport 2008, p. 144].

Case 2: *H is semisimple and $H' = \widetilde{H}$*. The following argument can be found in [Pappas and Rapoport 2008, p. 140–141]. Fix an algebraically closed field $l \supset k$, and let $M \supset L$ be the corresponding field extension of ramification index 1. We denote by Z the kernel of $\widetilde{H} \to H$ and let T and $\widetilde{T}$ denote the Néron models of fixed maximal tori in H_F and $\widetilde{H}_F$ satisfying $\widetilde{T}_F = f^{-1}(T_F)$. Since $\widetilde{H}_F$ is simply connected, $\widetilde{T}_F$ is an induced torus, i.e., there exist finite field extensions F_i/F such that

$$\widetilde{T}^0 \cong \prod_i \mathrm{Res}_{\mathbb{O}_{F_i}/\mathbb{O}_F} \mathbb{G}_m;$$

thus, there exists an $n \in \mathbb{N}$ such that

$$Z_F \subset \prod_i \mathrm{Res}_{F_i/F} \mu_n.$$

In particular, we have $Z(M) \subset \widetilde{T}^0(\mathbb{O}_M)$. Since $\widetilde{T}^0 \subset \widetilde{H}$, $f^0_{\mathscr{G}r}$ is injective on geometric points. The surjectivity is a direct consequence of [Pappas and Rapoport 2008, Appendix, Lemma 14]. □

Remark 3.2. If char $F = p$ and p does not divide the order of $\pi_1(H'_{\mathrm{der}})$ or $\pi_1(H_{\mathrm{der}})$, it is shown in [Pappas and Rapoport 2008, §6] that $f^\omega_{\mathscr{G}r}$ even induces an isomorphism of the underlying reduced ind-schemes. However, Pappas and Rapoport [2008, Example 6.4] show that this is not necessarily the case when we drop the condition on p. On the other hand $f^\omega_{\mathscr{G}r}$ is always an isomorphism in the case char $F = 0$, since universal homeomorphisms of perfect schemes are isomorphisms by [Bhatt and Scholze 2017, Lemma 3.8].

Let G^{ad} be the adjoint group of G. We denote by a subscript "ad" the image of an element of $G(L)$, $X_*(T)$, or $\pi_1(G)$ in $G^{\mathrm{ad}}(L)$, $X_*(T^{\mathrm{ad}})$, or $\pi_1(G^{\mathrm{ad}})$, respectively. By [Chen et al. 2015, Corollary 2.4.2], the homeomorphism of Proposition 3.1 induces a universal homeomorphism

$$X_\mu(b)^\omega \to X_{\mu_{\mathrm{ad}}}(b_{\mathrm{ad}})^{\omega_{\mathrm{ad}}} \tag{3.3}$$

whenever $X_\mu(b)^\omega$ is nonempty.

Equidimensionality for some affine Deligne–Lusztig varieties. Equidimensionality is known to hold in the following cases.

Theorem 3.4. *Let G, b, and μ be as above.*

(1) *If* char $F = p$, *then $X_\mu(b)$ and $X_{\preceq\mu}(b)$ are equidimensional. Furthermore, $X_{\preceq\mu}(b)$ is the closure of $X_\mu(b)$.*

(2) *Let F be an unramified extension of $\mathbb{Q}_p$, and let G be classical, μ be minuscule, and either $p \neq 2$ or all simple factors of G^{ad} be of type A or C. Then $X_\mu(b)$ is equidimensional.*

Proof. Assume first that char $F = p$. In the case where G is split the assertion is proven in [Hartl and Viehmann 2012, Corollary 6.8] by identifying the formal neighborhood of a closed point in the affine Deligne–Lusztig variety with a certain closed subscheme in the deformation space of a local G-shtuka. We briefly explain how to generalize the arguments in the proof of [Hartl and Viehmann 2012, Corollary 6.8] to arbitrary reductive group schemes over $\mathbb{O}_F$.

The main ingredient is the following result in [Viehmann and Wu 2018], generalizing [Hartl and Viehmann 2012, Theorem 6.6]. Let $x = gK \in X_{\preceq\mu}(b)(k)$, and denote $b' := gb\sigma(g)^{-1}$. Consider the deformation functor

$$\begin{aligned} \mathscr{D}ef_{b',0} &: (\mathrm{Art}/k) \to (\mathrm{Sets}), \\ A &\mapsto \{\tilde{b} \in (\overline{K\mu(\epsilon)K})(A) \mid \tilde{b}_k = b'\}/_{\cong} \end{aligned}$$

where $\tilde{b} \cong \tilde{b}'$ if there is an $h \in G(A[\![\epsilon]\!])$ with $h_k = 1$ and $h^{-1}\tilde{b}\sigma(h) = \tilde{b}'$. By [Viehmann and Wu 2018, Proposition 2.6] this functor is prorepresented by the formal completion of $K \setminus K\mu(\epsilon)K$ at b'. Moreover, the universal object has a unique algebraization by [Viehmann and Wu 2018, Lemma 2.8]. We denote by $D_{b',0}$ the algebraization of $(K \setminus K\mu(\epsilon)K)^\wedge_{b'}$ and by $\tilde{b} \in LG(D_{b',0})$ a lift of the universal object. We denote by $N_{[b],0} \subset D_{b',0}$ the minimal Newton stratum, that is, the set of all geometric points $\bar{s} : \operatorname{Spec} k_{\bar{s}} \to D_{b',0}$ such that $\tilde{b}_{\bar{s}}$ is $G(k_{\bar{s}}((\epsilon)))$-$\sigma$-conjugate to b (or b'). Since $N_{[b],0}$ is closed, we may equip it with the structure of a reduced subscheme. By [Viehmann and Wu 2018, Theorems 2.9 and 2.11] there exists a surjective finite morphism

$$\operatorname{Spec} k[\![x_1, \dots, x_{2\langle \rho_G, \nu_G(b)\rangle}]\!] \widehat{\times} X_{\preceq\mu}(b)^{\wedge}_x \to N_{[b],0}$$

where ρ_G denotes the half-sum of all absolute positive roots in G and $X_{\preceq\mu}(b)^{\wedge}_x$ the algebraization of the completion of $X_{\preceq\mu}(b)$ in x. In particular, we get

$$\begin{aligned} \dim N_{[b],0} &= 2\langle \rho_G, \nu_G(b)\rangle + \dim X_{\preceq\mu}(b)^{\wedge}_x \\ &\leq \langle \rho_G, \mu + \nu_G(b)\rangle - \tfrac{1}{2} \operatorname{def}_G(b). \end{aligned}$$

Here the last inequality follows from the dimension formula of $X_{\preceq\mu}(b)$ in [Hamacher 2015a, Theorem 1.1] and equality holds if and only if $\dim X_{\preceq\mu}(b)^{\wedge}_x = \dim X_{\preceq\mu}(b)$. The Newton stratification on $D_{b',0}$ satisfies strong purity in the sense of [Viehmann 2015, Definition 5.8]. Indeed, this is shown for $G = \mathrm{GL}_n$ in [Viehmann 2013, Theorem 7] and the general case follows by [Hamacher 2017, Proposition 2.2]. Thus, the conditions of [Viehmann 2015, Lemma 5.12] are satisfied and we get the dimension formula and closure relations of all Newton strata in $D_{b',0}$. In particular,

$$\dim N_{[b],0} = \langle \rho_G, \mu + \nu_G(b)\rangle - \tfrac{1}{2} \operatorname{def}_G(b).$$

Thus, $\dim X_{\preceq\mu}(b)^{\wedge}_x = \dim X_{\preceq\mu}(b)$ and since x was an arbitrary closed geometric point of $X_{\preceq\mu}(b)$, this implies equidimensionality. Since $\dim X_{\preceq\mu'}(b) < \dim X_{\preceq\mu}(b)$ for every $\mu' \prec \mu$ by [Hamacher 2015a, Theorem 1.1] this also implies the equidimensionality of $X_{\preceq\mu}(b)$ and that $X_\mu(b)$ is dense in $X_{\preceq\mu}(b)$.

Now consider $F = \mathbb{Q}_p$, $p \neq 2$, and assume first that there exists a faithful representation $\rho : G \hookrightarrow \mathrm{GL}_n$ such that the action of $\mathbb{G}_m$ via $\rho(\mu)$ has weights 0 and 1. Then we can associate a Rapoport–Zink space of Hodge type $\mathscr{M}_{G,\mu}(b)$ to the triple (G, μ, b), whose perfection equals $X_\mu(b)$ by [Zhu 2017, Theorem 3.10]. Since $\mathscr{M}_{G,\mu}(b)$ is equidimensional by [Hamacher 2017, Theorem 1.3], so is $X_\mu(b)$.

Now the morphism $X_\mu(b) \to X_{\mu_{\mathrm{ad}}}(b_{\mathrm{ad}})$ induced by the canonical projection $G \twoheadrightarrow G^{\mathrm{ad}}$ is an isomorphism on connected components by (3.3). Thus, all connected components of $X_{\mu_{\mathrm{ad}}}(b_{\mathrm{ad}})$ which are contained in the image of $X_\mu(b)$ are equidimensional. Since all connected components are isomorphic to each other by [Chen et al. 2015, Theorem 1.2], this implies that $X_{\mu_{\mathrm{ad}}}(b_{\mathrm{ad}})$ is equidimensional. Thus, any affine Deligne–Lusztig variety with G classical and adjoint and μ minuscule is equidimensional. Applying (3.3) once more, the claim follows for $p \neq 2$. If $p = 2$, the spaces $\mathscr{M}_{G,\mu}(b)$ are only defined if (G, μ, b) is of PEL type, but in this case the rest of the proof is identical.

If F is an unramified field extension of $\mathbb{Q}_p$, let $G' = \operatorname{Res}_{O_F/\mathbb{Z}_p} G$ and $\mu' = (\mu, 0, \dots, 0)$ and $b' = (b, 1, \dots, 1)$ with respect to the identification $G'_L \cong \prod_{F \hookrightarrow L} G$. By [Zhu 2017, Lemma 3.6] and the Cartesian diagram below it, we have $X_{\mu'}(b') \cong X_\mu(b)$. Thus, $X_\mu(b)$ is equidimensional. □

4. Irreducible components in the superbasic case

In this section we prove Theorem 1.4 for superbasic σ-conjugacy classes. In [Hamacher 2015a, §8] this has been reduced to a purely combinatorial statement, which we prove using the bijectivity of sweep maps on rational Dyck paths.

Superbasic σ-conjugacy classes. An element $b \in G(L)$ or the corresponding σ-conjugacy class $[b] \in B(G)$ is called superbasic if no element of $[b]$ is contained in a proper Levi subgroup of G defined over F.

Remark 4.1 [Chen et al. 2015, §3.1]. (1) If b is superbasic in $G(L)$, then the simple factors of the adjoint group G^{ad} are of the form $\mathrm{Res}_{F_d|F}\,\mathrm{PGL}_n$ for unramified extensions F_d of F (of degree d) and $n \geq 2$. In particular, $X_\mu(b)$ is equidimensional if $\mathrm{char}\,F = p$ or F is an unramified extension of $\mathbb{Q}_p$.

(2) For every $[b] \in B(G)$ there is a standard parabolic subgroup $P \subset G$ defined over F and with the following property. Let T be a fixed maximal torus of G and M the Levi factor of P containing T. Then there is a $b \in [b] \cap M(L)$ which is superbasic in M.

We first consider the special case where $[b]$ is superbasic and where G is of the form $\mathrm{Res}_{F_d|F}\,\mathrm{GL}_n$ for some d and n. In this case we give a proof using EL-charts as in [Hamacher 2015a] (see also [de Jong and Oort 2000] for the split case). We then reduce the general superbasic case to this particular case.

For G as above $L \otimes_F F_d \cong \prod_{\tau : F_d \hookrightarrow L} L$ yields an identification

$$G(L) = \prod_{\tau \in I} \mathrm{GL}_n(L)$$

mapping $g \in G(L)$ to a tuple $(g_\tau)_{\tau \in I}$ where $I := \mathrm{Gal}(F_d, F) \cong \mathbb{Z}/d\mathbb{Z}$. Let $S \subset T \subset B \subset G$ be the split diagonal torus, the diagonal torus, and the upper-triangular Borel, respectively. We have a canonical identification $X_*(T) \cong (\mathbb{Z}^n)^{|I|}$. Then the dominant elements in $X_*(T)$ are precisely the $\mu = (\mu_\tau)_{\tau \in I} \in X_*(T)$ such that the components of μ_τ are weakly decreasing for each τ.

We identify $X_*(S)$ with the invariants $X_*(T)^I = \mathbb{Z}^n$; thus,

$$\mu|_{\hat{S}} = \sum_{\tau \in I} \mu_\tau .$$

Moreover, this identifies the partial order $\leq$ on $X_*(S)_{\mathbb{Q}}$ with the dominance order on $\mathbb{Q}^n$.

A combinatorial identity. An important tool when considering the combinatorics of EL-charts is the sweep map defined by Armstrong, Loehr, and Warrington [Armstrong et al. 2015]. We need a multiple-component version of it, which turns out to be easily realized as a special case of the classical sweep map.

Notation 4.2. By a word $\boldsymbol{w}$ we mean a finite sequence of integers $w_1 \cdots w_r$. For $1 \leq k \leq r$ we define the level of $\boldsymbol{w}$ at k by $l(\boldsymbol{w})_k := \sum_{i=1}^{k} w_i$. We consider the following sets for fixed sequences of integers $a_{\tau,1}, \ldots, a_{\tau,n}$ where $1 \leq \tau \leq d$.

(1) Let $\mathscr{A}_{\mathbb{Z}}^{(d)}$ denote the set of words $\boldsymbol{w} = w_1 \cdots w_{d\cdot n}$ such that the subword $\boldsymbol{w}_{(\tau)} := w_\tau w_{\tau+d} \cdots w_{\tau+(n-1)\cdot d}$ is a rearrangement of $a_{\tau,1}, \ldots, a_{\tau,n}$ for any $\tau \in \{1, \ldots, d\}$.

(2) Denote by $\mathscr{A}_{\mathbb{N}}^{(d)} \subset \mathscr{A}_{\mathbb{Z}}^{(d)}$ the subset of words whose level at multiples of d is nonnegative. Following [Thomas and Williams 2017; Armstrong et al. 2015], we call its elements (d-component) Dyck words.

Definition 4.3. The sweep map $\mathrm{sw}^{(d)} : \mathscr{A}_{\mathbb{Z}}^{(d)} \to \mathscr{A}_{\mathbb{Z}}^{(d)}$ is the map that sorts $\boldsymbol{w}$ according to its level by permuting $\boldsymbol{w}_{(\tau)}$ using the following algorithm. Initialize $\mathrm{sw}^{(d)}(\boldsymbol{w})_{(\tau)} = \varnothing$ for any $1 \le \tau \le d$. For each a down from -1 to $-\infty$ and then down from ∞ to 0 read $\boldsymbol{w}_{(\tau)}$ from right to left and append to $\mathrm{sw}^{(d)}(\boldsymbol{w})_{(\tau)}$ all letters w_k such that $l(\boldsymbol{w})_k = a$.

We deduce the bijectivity of $\mathrm{sw}^{(d)}$ from Williams' result for the classical sweep map in [Thomas and Williams 2017].

Proposition 4.4. $\mathrm{sw}^{(d)}$ *is bijective and preserves* $\mathscr{A}_{\mathbb{N}}^{(d)}$.

Proof. If $d = 1$, the map $\mathrm{sw}^{(1)}$ is precisely the sweep map defined in [Thomas and Williams 2017] and the proposition is proven in [Thomas and Williams 2017, Theorems 6.1 and 6.3]. In order to reduce to this case, we need to construct an injection $\mathscr{A}_{\mathbb{Z}}^{(d)} \hookrightarrow \mathscr{A}_{\mathbb{Z}}^{(1)}$ which identifies Dyck words and preserves the sweep map, i.e., such that the diagram

$$\begin{tikzcd}
\mathscr{A}_{\mathbb{Z}}^{(d)} \arrow[r, hook] \arrow[d, "\mathrm{sw}^{(d)}"] & \mathscr{A}_{\mathbb{Z}}^{(1)} \arrow[d, "\mathrm{sw}^{(1)}"] \\
\mathscr{A}_{\mathbb{Z}}^{(d)} \arrow[r, hook] & \mathscr{A}_{\mathbb{Z}}^{(1)}
\end{tikzcd} \tag{4.5}$$

commutes. Note that part of this construction is also the choice of a sequence $\{a'_1, \dots, a'_{n\cdot d}\}$ for $\mathscr{A}_{\mathbb{Z}}^{(1)}$.

In preparation, fix an integer N big enough such that for any $\boldsymbol{w} \in \mathscr{A}_{\mathbb{Z}}^{(d)}$ and $1 \le \tau \le d$ as above the inequalities

$$\min\{l(\boldsymbol{w})_k + N \mid k \equiv \tau \pmod d\} > \max\{l(\boldsymbol{w})_k \mid k \equiv \tau - 1 \pmod d\}, \tag{4.6}$$

$$\min\{l(\boldsymbol{w})_k + \tau \cdot N \mid k \equiv \tau \pmod d\} \ge 0 \tag{4.7}$$

hold. We now construct a map $\mathscr{A}_{\mathbb{Z}}^{(d)} \to \mathscr{A}_{\mathbb{Z}}^{(1)}$, $\boldsymbol{w} \mapsto \boldsymbol{w}^{+N}$ satisfying the conditions above as follows. For given $\boldsymbol{w}$, let $\boldsymbol{w}^{+N}$ be the word which one obtains by replacing $w_{\tau+(i-1)\cdot d}$ by

$$w'_{\tau+(i-1)\cdot d} := \begin{cases} w_{\tau+(i-1)\cdot d} + N & \text{if } \tau \neq d, \\ w_{\tau+(i-1)\cdot d} - N \cdot (d-1) & \text{if } \tau = d \end{cases}$$

for $1 \le i \le n$ and $1 \le \tau \le d$. Then $\boldsymbol{w}^{+N} \in \mathscr{A}_{\mathbb{Z}}^{(1)}$ for the choice $\{a'_1, \dots, a'_{n\cdot d}\}$, where

$$a'_{\tau+(i-1)\cdot d} := \begin{cases} a_{\tau,i} + N & \text{if } \tau \neq d, \\ a_{\tau,i} - N \cdot (d-1) & \text{if } \tau = d. \end{cases}$$

The map $\boldsymbol{w} \mapsto \boldsymbol{w}^{+N}$ is obviously injective. Note that for any k we have $l(\boldsymbol{w}^{+N})_k = l(\boldsymbol{w})_k + \bar{k} \cdot N$ where $0 \le \bar{k} \le d-1$ denotes the residue of k modulo d. Thus, $l(\boldsymbol{w}^{+N})_k = l(\boldsymbol{w})_k$ if k is a multiple of d, and $l(\boldsymbol{w}^{+n}) \ge 0$ by (4.7) otherwise. Hence, $\boldsymbol{w}^{+N} \in \mathscr{A}_{\mathbb{N}}^{(1)}$ if any only if $\boldsymbol{w} \in \mathscr{A}_{\mathbb{N}}^{(d)}$.

By (4.6), we have $\min_i l(\boldsymbol{w}^{+N})_{\tau+(i-1)\cdot d} > \max_i l(\boldsymbol{w}^{+N})_{\varsigma+(i-1)\cdot d}$ for all $1 \leq \tau < \varsigma \leq d-1$ or $\varsigma < \tau = d$. Thus, the permutation of letters of $\boldsymbol{w}^{+N}$ induced by the classical sweep map decomposes into a product of permutations of the subsets $\{\tau, \tau+d, \tau+2d, \ldots\}$. Since moreover $l(\boldsymbol{w}^{+N})_{\tau+(i-1)\cdot d} \geq l(\boldsymbol{w}^{+N})_{\tau+(j-1)\cdot d}$ if and only if $l(\boldsymbol{w})_{\tau+(i-1)\cdot d} \geq l(\boldsymbol{w})_{\tau+(j-1)\cdot d}$, the permutations induced by the classical sweep map $\mathrm{sw}^{(1)}$ applied to $\boldsymbol{w}^{+N}$ and $\mathrm{sw}^{(d)}$ applied to $\boldsymbol{w}$ coincide. In other words, the diagram (4.5) commutes. □

Characterization of EL-charts. Throughout the section, we fix a positive integer m coprime to n and denote $\nu|_{\hat{S}} = (m/n, \ldots, m/n) \in X^*(\hat{S}) = X_*(T)^I$. Let $m_1, \ldots, m_d$ be arbitrary integers such that $m_1 + \cdots + m_d = m$. We shall later make convenient choices of them depending on μ. We recall the notion of EL-charts as they were presented in [Hamacher 2015a, §5].

Let $\mathbb{Z}^{(d)} := \coprod_{\tau \in I} \mathbb{Z}_{(\tau)}$ be the disjoint union of d copies of $\mathbb{Z}$. We impose the notation that for any subset $A \subset \mathbb{Z}^{(d)}$ we write $A_{(\tau)} := A \cap \mathbb{Z}_{(\tau)}$. For $x \in \mathbb{Z}$ we denote by $x_{(\tau)}$ the corresponding element of $\mathbb{Z}_{(\tau)}$ and write $|a_{(\tau)}| := a$. We equip $\mathbb{Z}^{(d)}$ with a partial order $\leq$ defined by

$$x_{(\tau)} \leq y_{(\varsigma)} \;:\Longleftrightarrow\; \tau = \varsigma \text{ and } x \leq y$$

and a $\mathbb{Z}$-action given by

$$x_{(\tau)} + z := (x+z)_{(\tau)}.$$

Furthermore, we consider a $\mathbb{Z}$-equivariant function $f : \mathbb{Z}^{(d)} \to \mathbb{Z}^{(d)}$ with

$$f(a_{(\tau)}) = a_{(\tau+1)} + m_\tau.$$

In particular, $f(\mathbb{Z}_{(\tau)}) = \mathbb{Z}_{(\tau+1)}$ and $f^d(a) = a + m$.

Definition 4.8. (1) An EL-chart is a nonempty subset $A \subset \mathbb{Z}^{(d)}$ which is bounded from below and satisfies $f(A) \subset A$ and $A + n \subset A$.

(2) Two EL-charts A and A' are called equivalent if there exists an integer z such that $A + z = A'$. We write $A \sim A'$.

Let A be an EL-chart and $B = A \setminus (A+n)$. It is easy to see that $\#B_{(\tau)} = n$ for all $\tau \in I$. We define a sequence $b_0, \ldots, b_{d\cdot n}$ as follows. Let $b_0 = b_{n\cdot d} = \min B_{(0)}$, and for given b_i let $b_{i+1} \in B$ be the unique element of the form

$$b_{i+1} = f(b_i) - \mu'_{i+1} \cdot n$$

for a nonnegative integer μ'_i. These elements are indeed distinct: if $b_i = b_j$, then obviously $i \equiv j \pmod d$ and then $b_{i+k\cdot d} \equiv b_i + k \cdot m \pmod n$ implies that $i = j$ as m and n are coprime.

It will later be helpful to distinguish the b_i and μ'_i of different components. For this we change the index set to $I \times \{1, \ldots, n\}$ via

$$b_{\tau,i} := b_{\tau+(i-1)\cdot d},$$
$$\mu'_{\tau,i} := \mu'_{\tau+(i-1)\cdot d}.$$

Here we choose the set of representatives $\{1, \ldots, d\} \subset \mathbb{Z}$ of I.

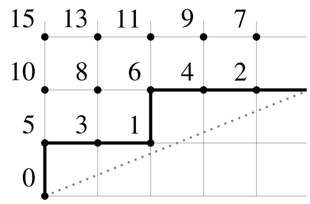

Figure 2. The associated Dyck path and $(5, -2)$-levels for $m = 5$, $n = 7$, and $A = \mathbb{N}_0$.

Definition 4.9. With the notation above, μ' is called the type of A.

Remark 4.10. This definition differs slightly from the definition of the type in [Hamacher 2015a, p. 12822]. In this article we choose the indices such that $\mu'_{\tau,i}$ measures the difference between $b_{\tau,i}$ and $b_{\tau-1,i}$ while in [Hamacher 2015a] it yields the difference between $b_{\tau,i}$ and $b_{\tau+1,i}$. Since one can alternate between those two notions by replacing f by $\bar{f} := f^{-1}$ and μ by $(-\mu)_{\mathrm{dom}}$, we can still use the combinatorial results of [Hamacher 2015a]. Moreover, we consider the Borel of *upper*-triangular matrices instead of lower-triangular matrices in [Hamacher 2015a], thus inverting the order on $X_*(S)$ and $X_*(T)$.

The type characterizes an EL-chart up to equivalence.

Lemma 4.11 [Hamacher 2015a, Lemma 5.3]. *Let*

$$P_{m,n,d} := \{\mu' \in (\mathbb{Z}^n_{\geq 0})^{|I|} \mid \mu'|_{\hat{S}} \leq \nu|_{\hat{S}}\}.$$

Then the type of any EL-chart A lies in $P_{m,n,d}$, and the type defines a bijection

$$\{\textit{EL-charts}\}/_{\sim} \leftrightarrow P_{m,n,d}.$$

Example 4.12. There are two important special cases of EL-charts.

(a) An EL-chart is called small if $A + n \subset f(A)$, in other words if its type only has entries 0 and 1. They correspond to the affine Deligne–Lusztig varieties with minuscule Hodge point.

(b) A semimodule is an EL-chart $A \subset \mathbb{Z}$. These are the invariants that occur in the split case.

There is a bijection between small semimodules up to equivalence and rational Dyck paths from $(0, 0)$ to $(n - m, m)$, that is, lattice paths allowing only steps in the north and east directions which stay above the diagonal. This gives a purely combinatorial motivation for the definitions below.

The bijection is given as follows (see [Gorsky and Mazin 2013] for more details). With a given equivalence class $[A]$ of small semimodules, we associate the path which goes east at the i-th step if $\mathrm{type}(A)_i = 0$ and north if $\mathrm{type}(A)_i = 1$. By the above lemma, this map is well defined and a bijection. Moreover, if we choose $\min A = 0$, then one can recover A from the Dyck path as the set of $(m, m - n)$-levels in the sense of [Armstrong et al. 2015] of points on or above the path, giving the inverse to the bijection. An example is illustrated in Figure 2.

There is another invariant of EL-charts which is more important for the application of this theory, as it allows us to calculate the dimension of strata inside the affine Deligne–Lusztig variety.

Definition 4.13. Let A be an EL-chart of type μ', and let $b_{\tau,i}$ be defined as above. For each $\tau \in I$ let $\tilde{b}_{\tau,1} > \cdots > \tilde{b}_{\tau,n}$ be the elements of $B_{(\tau)}$ arranged in decreasing order. Define

$$\tilde{\mu}_{\tau,i} = \mu'_{\tau,i'}$$

where i is the unique number such that $\tilde{b}_{\tau,i} = b_{\tau,i'}$. We call $\tilde{\mu}$ the cotype of A.

It is shown in [Hamacher 2015a, p. 12831] that cotype$(A) \in P_{m,n,d}$. Since the cotype is obviously invariant under equivalence, we obtain a map

$$\zeta : P_{m,n,d} \to P_{m,n,d}, \quad \text{type}(A) \mapsto \text{cotype}(A).$$

We claim that ζ is bijective. For this we note that ζ is the composition of

$$\mu' \mapsto (w_k := m_{k \pmod d} - \mu'_k \cdot n)_{k=1,\dots,n\cdot d} \overset{\mathrm{sw}^{(d)}}{\mapsto} (\widetilde{w}_k) \mapsto \Big(\frac{m_\tau - \widetilde{w}_{\tau+i\cdot d}}{n}\Big)_{\tau,i}. \tag{4.14}$$

Thus, its bijectivity follows from Proposition 4.4.

Example 4.15. For $d=1$, $n=7$, $m=5$, and $A=\mathbb{N}_0$, we can describe (4.14) as follows. In Figure 2 one sees that $\mu' := \text{type}(A) = (0, 1, 1, 0, 1, 1, 1)$. This is mapped to the word $\boldsymbol{w} = (5, -2, -2, 5, -2, -2, -2)$, whose levels $l(\boldsymbol{w}) = (5, 3, 1, 6, 4, 2, 0)$ are the corresponding elements of $B := A \setminus (A+n)$. Thus, applying the sweep map, which sorts the letters of $\boldsymbol{w}$ according to their levels, is nothing else than permuting the letters such that the corresponding elements of B get arranged in decreasing order. Now $\text{sw}(\boldsymbol{w}) = (5, 5, -2, -2, -2, -2, -2)$, which yields $\zeta(\mu') = (0, 0, 1, 1, 1, 1, 1)$.

Altogether, we obtain the following theorem, which generalizes the result of [Thomas and Williams 2017, Corollary 6.4]. It was conjectured in [Hamacher 2015a, Conjecture 8.3] and in the split case by de Jong and Oort [2000, Remark 6.16].

Theorem 4.16. *The cotype induces a bijection*

$$\{\textit{EL-charts}\}/_\sim \leftrightarrow P_{m,n,d}.$$

The superbasic case. Proposition 1.6 is a direct consequence of Theorem 4.16 together with the relation between orbits of irreducible components and EL-charts in [Hamacher 2015a, §8]. We briefly recall this relation for the reader's convenience before proving Proposition 1.6.

When applying the results of the previous subsection to affine Deligne–Lusztig varieties, we consider EL-charts satisfying certain additional criteria.

Definition 4.17. Let A be an EL-chart.

(1) A is called normalized if $\sum_{b\in B_{(0)}} |b| = \binom{n}{2}$ where $B_{(0)} = A_{(0)} \setminus (A_{(0)} + n)$.

(2) The Hodge point of A is defined as $\text{type}(A)_{\text{dom}}$.

Note that every EL-chart is equivalent to a unique normalized EL-chart. Let $P_\mu := \{\mu' \in P_{m,n,d} \mid \mu'_{\mathrm{dom}} = \mu\}$. Then by Lemma 4.11 $A \mapsto \mathrm{type}(A)$ induces a bijection

$$\{\text{normalized EL-charts with Hodge point } \mu\} \leftrightarrow P_\mu.$$

It is easy to see that ζ stabilizes P_μ. Thus, Theorem 4.16 says that $A \mapsto \mathrm{cotype}(A)$ induces a bijection between the set of normalized EL-charts with Hodge point μ and P_μ.

For every minuscule $\mu \in X_*(T)_{\mathrm{dom}}$ there exists a unique basic σ-conjugacy class in $B(G,\mu)$. We choose a representative of this σ-conjugacy class as follows. Let $m_\tau = \mathrm{val}\det \mu(\epsilon)_\tau$, and choose $b = ((b_{\tau,i,j})_{i,j=1}^n)_{\tau\in I}$ with

$$b_{\tau,i,j} = \begin{cases} \epsilon^{\lfloor (i+m_\tau)/n \rfloor} & \text{if } j-i \equiv m_\tau \pmod n, \\ 0 & \text{otherwise.} \end{cases}$$

Then the invariants $\lambda, \nu \in X^*(\hat{S}) = X^*(\widehat{T})_\Gamma \cong \mathbb{Z}^n$ are given by $\nu = (m/(d\cdot n), \dots, m/(d\cdot n))$ with $m = \sum_{\tau\in I} m_\tau$ and $\lambda = (\lambda_1, \dots, \lambda_n)$ with $\lambda_i = \lfloor i\cdot m/n \rfloor - \lfloor (i-1)\cdot m/n \rfloor$. The requirement that b is in fact superbasic corresponds to the assertion that m and n are coprime.

By our choice of b, the variety $X_\mu(b)^0 := X_\mu(b) \cap \mathscr{G}r_G^0$ is nonempty. In [Hamacher 2015a; 2015b] we constructed a $J_b(F)^0$-invariant cellular decomposition

$$X_\mu(b)^0 = \bigcup_A S_A$$

where the union runs over all normalized EL-charts with Hodge-point μ. We denote

$$V_A := \{(i,j) \mid b_i < b_j,\ \mu'_i = \mu'_j + 1\}.$$

In [Hamacher 2015a, Proposition 6.5; 2015b, Proposition 13.9] we show that $\mathbb{A}^{V_A} \xrightarrow{\sim} S_A$ by constructing an element $g_A \in LG(\mathbb{A}^{V_A})$ and a corresponding basis $(v_{\tau,i})$ of the universal G-lattice over S_A. In particular $\dim S_A = \#V_A$.

Following the calculations of the term S_1 in [Hamacher 2015a, p. 12831], one obtains $\#V_A$ from $\tilde{\mu}$ using the formula

$$\#V_A = \sum \lfloor \langle \tilde{\mu}|_{\hat{S}} - \mu_{\mathrm{adom}}, \hat{\omega}_F^\vee \rangle \rfloor,$$

where the sum runs over all relative fundamental coweights $\hat{\omega}_F^\vee$ of $\widehat{G}$ and μ_{adom} denotes the antidominant element in the W-orbit of μ. In particular, S_A is top-dimensional if and only if $\mathrm{cotype}(A)|_{\hat{S}} = \lambda$.

Proof of Proposition 1.6. Let G be arbitrary. We assume without loss of generality that $b \in K\mu(\epsilon)K$; thus, $X_\mu(b)^0 \neq \varnothing$. Since $J_b(F)$ acts transitively on $\pi_0(X_\mu(b))$ by [Chen et al. 2015, Theorem 1.2], it suffices to construct $C_{\tilde{\mu}}^0 := C_{\tilde{\mu}} \cap X_\mu(b)^0$, which have to be $J_b(F)^0$-stable and universally homeomorphic to affine spaces of the correct dimension. In particular, we may take $C^0_{\mathrm{cotype}(A)} = S_A$ if $G = \mathrm{Res}_{F_d/F}\mathrm{GL}_n$.

By Remark 4.1 we have $G^{\mathrm{ad}} \cong \prod_{i=1}^n \mathrm{Res}_{F_{d_i}/F}\mathrm{PGL}_{n_i}$. Let $G' = \prod_{i=1}^n \mathrm{Res}_{F_{d_i}/F}\mathrm{GL}_{n_i}$ and b' and μ' be lifts of b_{ad} and μ_{ad} to G', such that $X_{\mu'}(b')^0 \neq \varnothing$. We identify the underlying topological spaces $X_\mu(b)^0 = X_{\mu_{\mathrm{ad}}}(b_{\mathrm{ad}})^0 = X_{\mu'}(b')^0$ via the homeomorphism (3.3). Thus, we get a cellular decomposition of $X_\mu(b)^0$ per transport of structure from $X_{\mu'}(b')^0$. Since it is $J_{b'}(F)^0$-stable, we consider the canonical projections $J_b(F)^0 \xrightarrow{p} J_{b_{\mathrm{ad}}}(F)^0 \xleftarrow{q} J_{b'}(F)^0$. It suffices to show that q is surjective (implying that the decomposition is $J_{b_{\mathrm{ad}}}(F)^0$-stable) and that the $J_b(F)^0$-action factors through $J_{b_{\mathrm{ad}}}(F)^0$.

To prove the surjectivity, let $j \in J_{b_{\mathrm{ad}}}(F)^0$ and choose a preimage $g \in G(L)^0$ of j. The element g satisfies $g^{-1}b\sigma(g) = zb$ for some $z \in Z'(L) \cap G(L)^0 = Z(\mathbb{O}_L)$, where Z' denotes the center of G'. We choose $z' \in Z'(\mathbb{O}_L)$ with $(z')^{-1}\sigma(z') = z^{-1}$. Then $gz' \in J_{b'}(F)^0$ maps to j, as claimed.

Now an elementary calculation of the kernel shows that we have an exact sequence

$$1 \to Z(\mathbb{O}_F) \to J_b(F)^0 \to J_{b_{\mathrm{ad}}}(F)^0,$$

where Z denotes the center of G. Since $Z(\mathbb{O}_F)$ acts trivially on $\mathscr{G}r_G$, the $J_b(F)^0$-action factors through $J_{b_{\mathrm{ad}}}(F)^0$, as claimed. □

Corollary 4.18. *Conjecture 1.3 is true if b is superbasic and μ minuscule.*

Proof. We have

$$J_b(F) \setminus \Sigma(X_\mu(b)) \cong \{\tilde{\mu} \in P_\mu \mid C_{\tilde{\mu}} \text{ top-dimensional}\} \cong \{\tilde{\mu} \in W.\mu \mid \tilde{\mu}|_{\hat{S}} = \lambda\}.$$ □

5. Reduction to the superbasic case

In this section we consider the general case of Theorem 1.4; i.e., G is an unramified reductive group over F, μ is minuscule, and b is an arbitrary element of $G(L)$. The goal is to use a reduction method, first introduced in [Görtz et al. 2006], to relate to the superbasic case.

Let $P \subset G$ be a smallest standard parabolic subgroup of G, defined over F and with the following property. Let M be the Levi factor of P containing T. Then we want that $M(L)$ contains a σ-conjugate of b which is superbasic in M. Fix a representative $b \in M(L)$ of $[b]_G = [b]$. Then we furthermore want that the M-dominant Newton point of b is already G-dominant. For existence of such P, M, and b compare Remark 4.1. We write $P = M \cdot N$ where N denotes the unipotent radical of P. Since $b \in M(L)$, this induces a decomposition

$$J_b(F) \cap P(L) = (J_b(F) \cap M(L)) \cdot (J_b(F) \cap N(L)).$$

Throughout the section, we may refer to subschemes of the loop group or Grassmannian by their k-valued points to improve readability, e.g., write K instead of L^+G or $N(L)$ instead of LN. We denote $K_M = M(\mathbb{O}_L)$, $K_N = N(\mathbb{O}_L)$, and $K_P = P(\mathbb{O}_L)$.

We consider the variety

$$X_\mu^{M\subset G}(b) = \{gK_M \in \mathscr{G}r_M \mid g^{-1}b\sigma(g) \in K\mu K\}.$$

Then we have $X_\mu^{M\subset G}(b) = \coprod_{\mu' \in I_{\mu,b}} X_{\mu'}^M(b)$ where $I_{\mu,b}$ is the set of M-conjugacy classes of cocharacters μ' in the G-conjugacy class of μ with $[b]_M \in B(M,\mu')$. As $[b]_M$ is basic in M, this latter condition is equivalent to $\kappa_M(b) = \kappa_M(\mu')$ in $\pi_1(M)_\Gamma$. We identify an element of $I_{\mu,b}$ with its M-dominant representative in $X_*(T)$. Note that $I_{\mu,b}$ is nonempty and finite, but may have more than one element if G is not split.

Notation 5.1. Note that $X_\mu^{M\subset G}(b)$ is in general not equidimensional, although the individual summands are conjectured to be. We define

$$\Sigma'(X_\mu^{M\subset G}(b)) := \bigcup_{\mu'\in I_{\mu,b}} \Sigma^{\mathrm{top}}(X_{\mu'}^M(b)).$$

Using Corollary 4.18 we can show that $X_\mu^{M\subset G}(b)$ has the same number of orbits of irreducible components as given by the right-hand side of Theorem 1.4.

Lemma 5.2. $(J_b(F)\cap M(L))\setminus \Sigma'(X_\mu^{M\subset G}(b)) = W.\mu \cap [\tilde{\lambda} + (1-\sigma)X_*(T)]$

Proof. By Corollary 4.18 we have

$$(J_b(F)\cap M(L))\setminus \bigcup_{\mu'\in I_{\mu,b}} \Sigma^{\mathrm{top}}(X_{\mu'}^M(b)) = \bigcup_{\mu'\in I_{\mu,b}} (W_M.\mu' \cap [\tilde{\lambda}_M + (1-\sigma)X_*(T)]).$$

Here the unions on both sides are disjoint, and $\tilde{\lambda}_M = \tilde{\lambda}_M(b)$ denotes the element associated with $[b]\in B(M)$ whereas $\tilde{\lambda} = \tilde{\lambda}_G(b)$. By Lemma 2.4, the above union is equal to $\bigcup_{\mu'\in I_{\mu,b}}(W_M.\mu' \cap [\tilde{\lambda} + (1-\sigma)X_*(T)])$. As $\tilde{\lambda}$ is minuscule, the set $W_M.\mu' \cap [\tilde{\lambda} + (1-\sigma)X_*(T)]$ is nonempty for a given $\mu' \in W.\mu$ if and only if $\kappa_M(\mu') = \kappa_M(\tilde{\lambda})$ $(=\kappa_M(b))$, i.e., if and only if $\mu' \in I_{\mu,b}$. Hence, $\bigcup_{\mu'\in I_{\mu,b}}(W_M.\mu' \cap [\tilde{\lambda} + (1-\sigma)X_*(T)]) = W.\mu \cap [\tilde{\lambda} + (1-\sigma)X_*(T)]$. □

In order to relate the irreducible components of $X_\mu^{M\subset G}(b)$ to those of $X_\mu(b)$, we consider the variety

$$X_\mu^{P\subset G}(b) := \{gK_P \in \mathscr{G}r_P \mid g^{-1}b\sigma(g) \in K\mu(\epsilon)K\}$$

as an intermediate object. The inclusion $P \hookrightarrow G$ induces a natural map $X_\mu^{P\subset G}(b) \to X_\mu(b)$. Using the Iwasawa decomposition $G(L) = P(L)K$ we see that this map is surjective, and in fact $X_\mu^{P\subset G}(b)$ is nothing but a decomposition of $X_\mu^G(b)$ into locally closed subsets (see, e.g., [Hamacher 2015a, Lemma 2.2]). Thus, we obtain a natural bijection

$$\Sigma^{\mathrm{top}}(X_\mu^{P\subset G}(b)) \to \Sigma^{\mathrm{top}}(X_\mu^G(b))$$

which induces a surjection

$$\alpha_\Sigma : (J_b(F)\cap P(L))\setminus \Sigma^{\mathrm{top}}(X_\mu^{P\subset G}(b)) \twoheadrightarrow J_b(F)\setminus \Sigma^{\mathrm{top}}(X_\mu^G(b)). \tag{5.3}$$

Furthermore, $\dim X_\mu^{P\subset G}(b) = \dim X_\mu^G(b)$.

On the other hand, the restriction of the canonical projection $\mathscr{G}r_P \twoheadrightarrow \mathscr{G}r_M$ induces a surjective morphism

$$\beta : X_\mu^{P\subset G}(b) \to X_\mu^{M\subset G}(b)$$

by [Hamacher 2015a, Proposition 2.9]. Moreover the fiber dimension for $x \in X_{\mu'}^M(b)$ is given by

$$\dim \beta^{-1}(x) = \dim X_\mu^{P\subset G}(b) - \dim X_{\mu'}^M(b) \tag{5.4}$$

[Hamacher 2015a, Lemma 2.8 and Proposition 2.9(2)], using that for minuscule μ, equality in [Hamacher 2015a, Lemma 2.8] always holds, and using the dimension formula [Hamacher 2015a, Theorem 1.1]. Note that this only depends on μ' (but indeed depends on the choice of $\mu' \in I_{\mu,b}$), but not on the point x.

Lemma 5.5. *β induces a well defined surjective map*

$$\beta_\Sigma : \Sigma^{\mathrm{top}}(X_\mu^{P\subset G}(b)) \to \Sigma'(X_\mu^{M\subset G}(b)).$$

It is $J_b(F) \cap P(L)$-equivariant for the natural action on the left-hand side, and the action through the natural projection $J_b(F) \cap P(L) \twoheadrightarrow J_b(F) \cap M(L)$ on the right-hand side.

Recall that a subset of $G(L)$ is called bounded if it is contained in a finite union of K-double cosets.

Proof. Let $\mathscr{C}$ be a top-dimensional irreducible component of $X_\mu^{P\subset G}(b)$. Then $\beta(\mathscr{C})$ is irreducible and thus contained in one of the open and closed subschemes $X_{\mu'}^M(b)$. By (5.4), its dimension is equal to $\dim(X_{\mu'}^M(b))$; hence, $\beta(\mathscr{C})$ is a dense subscheme of one of the irreducible components of $X_{\mu'}^M(b)$. In this way we obtain the claimed map β_Σ. It is surjective and $J_b(F) \cap P(L)$-equivariant because the same holds for β. □

Proposition 5.6. *Let $Z \subset X_\mu^{M\subset G}(b)$ be an irreducible subscheme. Then $J_b(F) \cap N(L)$ acts transitively on $\Sigma(\beta^{-1}(Z))$.*

In the proof we need the following remark.

Remark 5.7. For $x \in \widetilde{W}$ let IxI be the locally closed subscheme of LG whose k-valued points are $I(k)xI(k)$. Let Y be a scheme and $g \in (IxI)(Y)$. Then we claim that there are elements $i_1, i_2 \in I(Y)$ with $g = i_1 x i_2$. In equal characteristic, this is [Hartl and Viehmann 2012, Lemma 2.4] (whose proof shows the above statement, although the lemma only claims the assertion étale locally on Y). Let us explain how to modify the proof to deduce the above statement in general: we consider the morphism $I/(I \cap xIx^{-1}) \to LG/I$ to the affine flag variety given by $g \mapsto gx$. By writing down the obvious inverse one sees that it is an immersion with image IxI/I.

Let $g \in (IxI)(Y)$ and $\bar{g}$ be its image in the affine flag variety. Then the above shows that $\bar{g}$ is the image of some $\bar{i} \in I/(I \cap xIx^{-1})(Y)$. Note that $I/(I \cap xIx^{-1}) = I_0/(I_0 \cap xI_0x^{-1})$ where I_0 is the unipotent radical of I. By [Hartl and Viehmann 2012, Lemma 2.1] we can thus lift $\bar{i} \in I_0/(I_0 \cap xI_0x^{-1})(Y)$ to an element $i_1 \in I_0(Y)$ which is as claimed.

Proof of Proposition 5.6. As we have to take an inverse image of an element under σ later in this proof, we replace all occurring ind-schemes by their perfections. Note that this does not change the underlying topological spaces of the schemes. Moreover, since we may check the assertion on an open covering of Z, we may replace Z by an open subscheme $Y \subset Z$ containing one fixed but arbitrary point $z \in Z(k)$.

Étale locally there is a lifting of the inclusion $Z \hookrightarrow X_{\mu'}^M(b)$ to LM [Pappas and Rapoport 2008, Lemma 1.4] (the proof also works for char $F = 0$; compare [Zhu 2017, Proposition 1.20]). Thus, there exists $Y' \to Z$ étale with $z \in \mathrm{im}(Y' \to Z)$ such that there exists a lift $\iota : Y' \to LM$. By replacing Y' by an irreducible component if necessary, we may assume that Y' is again irreducible. We denote by Y the image of Y' in Z, and by $y \in Y'$ a point mapping to z.

We denote

$$\Phi = \{(m, n) \in \iota(Y') \times N(L) \mid mnK_P \in X_\mu^{P \subset G}(b)\}$$

and $b_m := m^{-1} b\sigma(m)$ for any $m \in M(L)$. For $g = mn \in P(L)$ we have

$$g^{-1} b\sigma(g) = b_m \cdot [b_m^{-1} n^{-1} b_m \sigma(n)] \tag{5.8}$$

where the bracket is in $N(L)$ and where $b_m \in M(L)$. The condition $gK_P \in \beta^{-1}(Y)$ is then equivalent to the condition that we may choose $m \cdot n \in gK_P$ with $m \in \iota(Y') \subset LM$ and $n \in N(L)$ such that the last bracket is in $N(L) \cap b_m^{-1} K\mu(\epsilon)K$. Thus, we have a morphism

$$\gamma : \Phi \to \mathscr{E} := \{(m, c) \mid m \in \iota(Y'),\ c \in N(L) \cap b_m^{-1} K\mu(\epsilon)K\}, \quad (m, n) \mapsto (m, b_m^{-1} n^{-1} b_m \sigma(n)).$$

In order to get an easier description of $\mathscr{E}$, we show that one can assume $b_m \in K_M \cdot \mu'$ after further shrinking Y and replacing ι if necessary. Let $x \in \widetilde{W}$ such that $I_M x I_M \subset K_M \mu'(\epsilon) K_M$ is the open cell, where I_M denotes the standard Iwahori subgroup of M. Then $K\mu(\epsilon)K = KxK$, and we fix $k_0, k_0' \in K$ such that $b_{\iota(y)} = k_0 x k_0'$. We replace Y' (and thus Y) by the open neighborhood of y such that $b_m \in k_0 \cdot I_M x I_M \cdot k_0'$ for all $m \in \iota(Y')$. By Remark 5.7 we have a global decomposition $b_m = k_0 i_1 x i_2 k_0'$ with $i_j \in I_M(\iota(Y'))$. As $Y \subseteq X_{\mu'}^M(b)$ we have $x = w_1 \mu' w_2 \in W_M \mu W_M$; thus, $b_m = k_0 i_1 w_1 \mu'(\epsilon) w_2 i_2 k_0'$. We now replace m by $m\sigma^{-1}(w_2 i_2 k_0')^{-1} \in mK_M$ and modify ι accordingly. With respect to this new choice we obtain a decomposition of b_m of the form $k_1 \mu'(\epsilon)$ with $k_1 = \sigma^{-1}(w_2 i_2 k_0') k_0 i_1 w_1 \in L^+M(\iota(Y'))$. Now

$$\begin{aligned} N(L) \cap b_m^{-1} K\mu(\epsilon)K &= N(L) \cap \mu'(\epsilon)^{-1} K\mu(\epsilon)K \\ &= \mu'(\epsilon)^{-1}(N(L) \cdot \mu'(\epsilon) \cap K\mu(\epsilon)K). \end{aligned}$$

Note that this only depends on the constant element μ'. Hence,

$$\mathscr{E} = \iota(Y') \times (N(L) \cap \mu'(\epsilon)^{-1} K\mu(\epsilon)K).$$

Claim 1: *$\mathscr{E}$ is irreducible.* As $\iota(Y')$ is irreducible, we have to show that $N(L) \cap \mu'(\epsilon)^{-1} K\mu(\epsilon)K$ is irreducible. For this we consider the morphism $\mathrm{pr}_{\mu'} : N(L) \to N(L)\mu'(\epsilon)K \subset \mathscr{G}r_G$, $n \mapsto \mu'(\epsilon)n$. Then $N(L) \cap \mu'(\epsilon)^{-1} K\mu(\epsilon)K$ is the preimage of $N(L)\mu'(\epsilon)K \cap K\mu(\epsilon)K$, which is irreducible by [Mirković and Vilonen 2007, Corollary 13.2]. On the other hand $\mathrm{pr}_{\mu'}$ is a K_N-torsor, since it is surjective and factorizes as

$$N(L) \to \mathscr{G}r_N \hookrightarrow \mathscr{G}r_G \xrightarrow{\mu'(\epsilon)\cdot} \mathscr{G}r_G.$$

Here the first map is the projection, a K_N-torsor. The second is the natural closed embedding, and the third the isomorphism obtained by left multiplication by $\mu'(\epsilon)$. As K_N is also irreducible, this completes the proof of Claim 1.

Claim 2: *Let $\mathscr{F} \subseteq \Phi$ be a nonempty open subscheme with $\mathscr{F} = \mathscr{F} K_N$ where K_N acts by right multiplication on the second component. Then its image under γ contains an open subscheme of $\mathscr{E}$. In particular, it is dense by Claim 1.*

Fix an irreducible component C of Φ such that its intersection with $\mathscr{F}$ is nonempty. We may replace $\mathscr{F}$ by an open and dense subscheme of points only contained in the one irreducible component C. As $\mathscr{F}$ is invariant under right multiplication by K_N and m is contained in a bounded subscheme of LM, its image under γ is invariant under right multiplication by some (sufficiently small) open subgroup K'_N of K_N (this follows from the same proof as [Görtz et al. 2006, Proposition 5.3.1], which carries over literally to the unramified case and the case $\operatorname{char} F = 0$). Thus, it is enough to show that the image of $\gamma(\mathscr{F})$ in $\iota(Y') \times (N(L) \cap \mu'(\epsilon)^{-1} K \mu K)/K'_N$ is open. Let $g_0 \in \mathscr{F}$, and let $U = \operatorname{Spec} R$ be an affine open neighborhood of $\gamma(g_0)$ in $\mathscr{E}$. After possibly replacing K'_N by a smaller open subgroup we may assume that U is K'_N-invariant. Let (m_U, n_U) be the universal element. Then m_U and n_U are contained in bounded subsets of LM and LN, respectively. By Corollary 5.11 there is an étale covering R' of R and a morphism $\operatorname{Spec} R' \to \Phi$ such that the composite with γ and the quotient modulo K'_N maps $\operatorname{Spec} R'$ surjectively to U/K'_N. Intersecting $\operatorname{Spec} R'$ with the inverse image of the open subscheme $\mathscr{F}$ of Φ and using that $R \to R'$ is finite étale, we obtain an open subscheme of $\operatorname{Spec} R'$, or of $\mathscr{F}$ mapping surjectively to an open neighborhood of $g_0 K'_N$. This implies the claim.

Finally, we show that all irreducible components of $\beta^{-1}(Y)$ are contained in one $J_b(F) \cap N(L)$-orbit of irreducible components of $X_\mu^{P \subset G}(b)$. Let $\mathscr{D}$ and $\mathscr{D}'$ be irreducible components of $\beta^{-1}(Y)$. We have to show that all dense open subsets D and D' of the two components contain points p and p' which are in the same $J_b(F)$-orbit. Consider the K_N-torsor

$$\phi : \Phi \to \beta^{-1}(Y), \quad (m, n) \mapsto mnK_P.$$

Then it is enough to show that for all nonempty open subsets C_1 and C_2 of Φ with $C_i K_N = C_i$ there are points $q_i \in C_i$ and a $j \in J$ with $\phi(q_1) = j\phi(q_2)$. This latter condition follows if we can show that $\gamma(q_1) = \gamma(q_2)$. But by Claim 2, $\gamma(C_1)$ and $\gamma(C_2)$ are both open and dense in $\mathscr{E}$, which implies the existence of such q_1 and q_2. ☐

Corollary 5.9. *β_Σ induces a bijection*

$$(J_b(F) \cap P(L)) \setminus \Sigma(X_\mu^{P \subset G}(b)) \xrightarrow{1:1} (J_b(F) \cap M(L)) \setminus \Sigma(X_\mu^{M \subset G}(b))$$

which restricts to

$$(J_b(F) \cap P(L)) \setminus \Sigma^{\mathrm{top}}(X_\mu^{P \subset G}(b)) \xrightarrow{1:1} (J_b(F) \cap M(L)) \setminus \Sigma'(X_\mu^{M \subset G}(b)).$$

In particular $X_\mu^{P \subset G}(b)$ is equidimensional if and only if the $X_{\mu'}^{M}(b)$ are for all $\mu' \in I_{\mu,b}$.

We use the following notation. Let R be an integral k-algebra. In the arithmetic case we assume R to be perfect and let $\mathscr{R} = W_{\mathcal{O}_F}(R)$. In the function field case, let $\mathscr{R} = R[[\epsilon]]$. In both cases let $\mathscr{R}_L = \mathscr{R}[1/\epsilon]$.

For $m \in M(\mathcal{R}_L)$ consider the map

$$f_m : LN_R \to LN_R, \quad n \mapsto (m^{-1}n^{-1}m)\sigma(n).$$

Lemma 5.10 (Chen, Kisin, and Viehmann). *Let $b \in [b] \cap M(L)$ with $b\sigma(b)\cdots\sigma^{l_0-1}(b) = \epsilon^{l_0\nu_b}$ for some $l_0 > 0$ such that $l_0\nu_b \in X_*(T)$. Let R be an integral k-algebra, $\mathcal{R}$ and $\mathcal{R}_L$ as above, and $y \in N(\mathcal{R}_L)$ contained in a bounded subscheme. Further let $x_1 \in \operatorname{Spec} R(k)$ and $z_1 \in N(L)$ with $f_b(z_1) = y(x_1)$. Then for any bounded open subgroup $K' \subset N(L)$ there exists a finite étale covering $R \to R'$ with associated $\mathcal{R} \to \mathcal{R}'$ and $z \in N(\mathcal{R}'_L)$ such that*

(1) *for every k-valued point x of R' we have $f_b(z(x))K' = y(x)K'$ and*

(2) *there exists a point $x'_1 \in \operatorname{Spec} R'(k)$ over x_1 such that $z(x'_1) = z_1$.*

Proof. This is [Chen et al. 2015, Lemma 3.4.4], except for the fact that there R is assumed to be smooth, and only the case of mixed characteristic is considered. But actually, none of these assumptions is needed in the proof given there. □

Corollary 5.11. *Let $b \in [b] \cap M(L)$ and R and $\mathcal{R}$ be as in the previous lemma. Let $m \in M(\mathcal{R}_L)$, and $y \in N(\mathcal{R}_L)$, each contained in a bounded subscheme. Further let $x_1 \in \operatorname{Spec} R(k)$ and $z_1 \in N(L)$ with $f_b(z_1) = y(x_1)$. Let $b_m = m^{-1}b\sigma(m) \in M(\mathcal{R}_L)$. Then for any bounded open subgroup $K' \subset N(L)$ there exists a finite étale covering $R \to R'$ with associated extension $\mathcal{R} \to \mathcal{R}'$ and $z \in N(\mathcal{R}'_L)$ such that*

(1) *for every k-valued point x of R' we have $f_{b_m}(z(x))K' = y(x)K'$ and*

(2) *there exists a point $x'_1 \in \operatorname{Spec} R'(k)$ over x_1 such that $z(x'_1) = z_1$.*

Proof. For $n \in N(L)$ we have

$$\begin{aligned} f_{b_m}(n) &= (\sigma(m)^{-1}b^{-1}m)n^{-1}(m^{-1}b\sigma(m))\sigma(n) \\ &= \sigma(m)^{-1}b^{-1}(mn^{-1}m^{-1})b\sigma(mnm^{-1})\sigma(m) \\ &= \sigma(m)^{-1}f_b(mnm^{-1})\sigma(m). \end{aligned}$$

By the boundedness assumption on m, there is a bounded open subgroup K'' such that

$$\sigma(m(x))^{-1}K''\sigma(m(x)) \in K'$$

for all $\bar{k}$-valued points x of $\operatorname{Spec} R$. Applying Lemma 5.10 to $\sigma(m)y\sigma(m)^{-1}$ and K'', and conjugating the result by m, we obtain the desired lifting with respect to f_{b_m}. □

Theorem 5.12. *Let $\mu \in X_*(T)_{\mathrm{dom}}$ be minuscule, $b \in [b] \in B(G, \mu)$, and $\tilde{\lambda} \in X_*(T)$ be an associated element. Then the map*

$$\phi = \alpha_\Sigma \circ \beta_\Sigma^{-1} : W.\mu \cap [\tilde{\lambda} + (1-\sigma)X_*(T)] \to J_b(F) \setminus \Sigma^{\mathrm{top}}(X_\mu(b))$$

constructed above is surjective and it is bijective if and only if $J_b(F)$ acts trivially on

$$(J_b(F) \cap P(L)) \setminus \Sigma^{\mathrm{top}}(X_\mu(b)).$$

Proof. From Lemma 5.2, Corollary 5.9, and (5.3) we obtain the claimed maps

$$\begin{aligned} W.\mu \cap [\tilde{\lambda} + (1-\sigma)X_*(T)] &= (J_b(F) \cap M(L)) \setminus \Sigma(X_\mu^{M\subset G}(b)) \\ &\overset{\beta_\Sigma^{-1}}{\to} (J_b(F) \cap P(L)) \setminus \Sigma^{\mathrm{top}}(X_\mu^{P\subset G}(b)) \\ &\overset{\alpha_\Sigma}{\twoheadrightarrow} J_b(F) \setminus \Sigma^{\mathrm{top}}(X_\mu(b)). \end{aligned}$$

As $\Sigma^{\mathrm{top}}(X_\mu(b)) \cong \Sigma^{\mathrm{top}}(X_\mu^{P\subset G}(b))$, this description also implies the assertion about bijectivity. □

Proof of Theorem 1.4. The first assertion is a direct consequence of the previous theorem.

If G is split, then $W.\mu \cap [\tilde{\lambda} + (1-\sigma)X_*(T)] = \{\tilde{\lambda}\}$ has only one element; hence, the map is also injective.

If the second condition holds, then $J_b(F) \subset P(L)$; hence, α_Σ and also ϕ are bijective. □

References

[Armstrong et al. 2015] D. Armstrong, N. A. Loehr, and G. S. Warrington, "Sweep maps: a continuous family of sorting algorithms", *Adv. Math.* **284** (2015), 159–185. MR Zbl

[Bhatt and Scholze 2017] B. Bhatt and P. Scholze, "Projectivity of the Witt vector affine Grassmannian", *Invent. Math.* **209**:2 (2017), 329–423. MR Zbl

[Bourbaki 1968] N. Bourbaki, *Éléments de mathématique, XXXIV: Groupes et algèbres de Lie*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. MR Zbl

[Chen and Viehmann 2018] M. Chen and E. Viehmann, "Affine Deligne–Lusztig varieties and the action of J", *J. Algebraic Geom.* **27**:2 (2018), 273–304. MR Zbl

[Chen et al. 2015] M. Chen, M. Kisin, and E. Viehmann, "Connected components of affine Deligne–Lusztig varieties in mixed characteristic", *Compos. Math.* **151**:9 (2015), 1697–1762. MR Zbl

[Gorsky and Mazin 2013] E. Gorsky and M. Mazin, "Compactified Jacobians and q, t-Catalan numbers, I", *J. Combin. Theory Ser. A* **120**:1 (2013), 49–63. MR Zbl

[Görtz et al. 2006] U. Görtz, T. J. Haines, R. E. Kottwitz, and D. C. Reuman, "Dimensions of some affine Deligne–Lusztig varieties", *Ann. Sci. École Norm. Sup.* (4) **39**:3 (2006), 467–511. MR Zbl

[Hamacher 2015a] P. Hamacher, "The dimension of affine Deligne–Lusztig varieties in the affine Grassmannian", *Int. Math. Res. Not.* **2015**:23 (2015), 12804–12839. MR Zbl

[Hamacher 2015b] P. Hamacher, "The geometry of Newton strata in the reduction modulo p of Shimura varieties of PEL type", *Duke Math. J.* **164**:15 (2015), 2809–2895. MR Zbl

[Hamacher 2017] P. Hamacher, "The almost product structure of Newton strata in the deformation space of a Barsotti–Tate group with crystalline Tate tensors", *Math. Z.* **287**:3-4 (2017), 1255–1277. MR Zbl

[Hartl and Viehmann 2011] U. Hartl and E. Viehmann, "The Newton stratification on deformations of local G-shtukas", *J. Reine Angew. Math.* **656** (2011), 87–129. MR Zbl

[Hartl and Viehmann 2012] U. Hartl and E. Viehmann, "Foliations in deformation spaces of local G-shtukas", *Adv. Math.* **229**:1 (2012), 54–78. MR Zbl

[He and Nie 2014] X. He and S. Nie, "Minimal length elements of extended affine Weyl groups", *Compos. Math.* **150**:11 (2014), 1903–1927. MR Zbl

[He and Nie 2015] X. He and S. Nie, "P-alcoves, parabolic subalgebras and cocenters of affine Hecke algebras", *Selecta Math. (N.S.)* **21**:3 (2015), 995–1019. MR Zbl

[He and Zhou 2016] X. He and R. Zhou, "On the connected components of affine Deligne–Lusztig varieties", preprint, 2016. arXiv [de Jong and Oort 2000] A. J. de Jong and F. Oort, "Purity of the stratification by Newton polygons", *J. Amer. Math. Soc.* **13**:1 (2000), 209–241. MR Zbl

[Kottwitz 1985] R. E. Kottwitz, "Isocrystals with additional structure", *Compositio Math.* **56**:2 (1985), 201–220. MR Zbl

[Mirković and Vilonen 2007] I. Mirković and K. Vilonen, "Geometric Langlands duality and representations of algebraic groups over commutative rings", *Ann. of Math.* (2) **166**:1 (2007), 95–143. MR Zbl

[Nie 2015] S. Nie, "Connected components of closed affine Deligne–Lusztig varieties in affine Grassmannians", preprint, 2015. arXiv

[Pappas and Rapoport 2008] G. Pappas and M. Rapoport, "Twisted loop groups and their affine flag varieties", *Adv. Math.* **219**:1 (2008), 118–198. MR Zbl

[Rapoport 2005] M. Rapoport, "A guide to the reduction modulo p of Shimura varieties", pp. 271–318 in *Automorphic forms, I* (Paris, 2000), edited by J. Tilouine et al., Astérisque **298**, 2005. MR Zbl

[Richarz 2016] T. Richarz, "Affine Grassmannians and geometric Satake equivalences", *Int. Math. Res. Not.* **2016**:12 (2016), 3717–3767. MR

[Thomas and Williams 2017] H. Thomas and N. Williams, "Sweeping up zeta", *Sém. Lothar. Combin.* **78B** (2017), 10. MR Zbl

[Viehmann 2008a] E. Viehmann, "Moduli spaces of p-divisible groups", *J. Algebraic Geom.* **17**:2 (2008), 341–374. MR Zbl

[Viehmann 2008b] E. Viehmann, "The global structure of moduli spaces of polarized p-divisible groups", *Doc. Math.* **13** (2008), 825–852. MR Zbl

[Viehmann 2013] E. Viehmann, "Newton strata in the loop group of a reductive group", *Amer. J. Math.* **135**:2 (2013), 499–518. MR Zbl

[Viehmann 2015] E. Viehmann, "On the geometry of the Newton stratification", preprint, 2015. To appear in *Stabilization of the trace formula, Shimura varieties, and arithmetic applications, II*: *Shimura varieties and Galois representations*, edited by M. Harris and T. Haines. arXiv

[Viehmann and Wu 2018] E. Viehmann and H. Wu, "Central leaves in loop groups", *Math. Res. Lett.* **25**:3 (2018), 989–1008.

[Vollaard and Wedhorn 2011] I. Vollaard and T. Wedhorn, "The supersingular locus of the Shimura variety of $\mathrm{GU}(1, n-1)$, II", *Invent. Math.* **184**:3 (2011), 591–627. MR Zbl

[Xiao and Zhu 2017] L. Xiao and X. Zhu, "Cycles on Shimura varieties via geometric Satake", preprint, 2017. arXiv

[Zhu 2017] X. Zhu, "Affine Grassmannians and the geometric Satake in mixed characteristic", *Ann. of Math.* (2) **185**:2 (2017), 403–492. MR Zbl

Communicated by Kiran S. Kedlaya
Received 2017-03-07 Revised 2017-12-28 Accepted 2018-03-30

hamacher@ma.tum.de *Fakultät für Mathematik, Technische Universität München, Garching bei München, Germany*

viehmann@ma.tum.de *Fakultät für Mathematik, Technische Universität München, Garching bei München, Germany*

mathematical sciences publishers msp

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1635

Arithmetic degrees and dynamical degrees of endomorphisms on surfaces

Yohsuke Matsuzawa, Kaoru Sano and Takahiro Shibata

For a dominant rational self-map on a smooth projective variety defined over a number field, Kawaguchi and Silverman conjectured that the (first) dynamical degree is equal to the arithmetic degree at a rational point whose forward orbit is well-defined and Zariski dense. We prove this conjecture for surjective endomorphisms on smooth projective surfaces. For surjective endomorphisms on any smooth projective varieties, we show the existence of rational points whose arithmetic degrees are equal to the dynamical degree. Moreover, if the map is an automorphism, there exists a Zariski dense set of such points with pairwise disjoint orbits.

1. Introduction

Let k be a number field, X a smooth projective variety over $\bar{k}$, and $f : X \dashrightarrow X$ a dominant rational self-map on X over $\bar{k}$. Let $I_f \subset X$ be the indeterminacy locus of f. Let $X_f(\bar{k})$ be the set of $\bar{k}$-rational points P on X such that $f^n(P) \notin I_f$ for every $n \geq 0$. For $P \in X_f(\bar{k})$, its *forward f-orbit* is defined as $\mathcal{O}_f(P) := \{f^n(P) : n \geq 0\}$.

Let H be an ample divisor on X defined over $\bar{k}$. The (*first*) *dynamical degree* of f is defined by

$$\delta_f := \lim_{n\to\infty} ((f^n)^* H \cdot H^{\dim X - 1})^{1/n}.$$

MSC2010: primary 14G05; secondary 11G35, 11G50, 37P05, 37P15, 37P30.
Keywords: arithmetic degree, dynamical degrees, arithmetic dynamics.

The first dynamical degree of a dominant rational self-map on a smooth complex projective variety was first defined by Dinh and Sibony [2004; 2005]. Dang [2017] and Truong [2015] gave algebraic definitions of dynamical degrees.

The *arithmetic degree*, introduced by Silverman [2014], of f at a $\bar{k}$-rational point $P \in X_f(\bar{k})$ is defined by

$$\alpha_f(P) := \lim_{n\to\infty} h_H^+(f^n(P))^{1/n}$$

if the limit on the right-hand side exists. Here, $h_H : X(\bar{k}) \to [0, \infty)$ is the (absolute logarithmic) Weil height function associated with H, and we put $h_H^+ := \max\{h_H, 1\}$.

Then we have two types of quantity concerned with the iteration of the action of f. It is natural to consider the relation between dynamical degrees and arithmetic degrees. In this direction, Kawaguchi and Silverman formulated the following conjecture.

Conjecture 1.1 (The Kawaguchi–Silverman conjecture [2016b, Conjecture 6]). *For every $\bar{k}$-rational point $P \in X_f(\bar{k})$, the arithmetic degree $\alpha_f(P)$ exists. Moreover, if the forward f-orbit $\mathcal{O}_f(P)$ is Zariski dense in X, the arithmetic degree $\alpha_f(P)$ is equal to the dynamical degree δ_f, i.e., we have*

$$\alpha_f(P) = \delta_f.$$

Remark 1.2. Let X be a complex smooth projective variety with $\kappa(X) > 0$, $\Phi : X \dashrightarrow W$ the Iitaka fibration of X, and $f : X \dashrightarrow X$ a dominant rational self-map on X. Nakayama and Zhang [2009, Theorem A] proved that there exists an automorphism $g : W \to W$ of finite order such that $\Phi \circ f = g \circ \Phi$. This implies that any dominant rational self-map on a smooth projective variety of positive Kodaira dimension does not have a Zariski dense orbit. So the latter half of Conjecture 1.1 is meaningful only for smooth projective varieties of nonpositive Kodaira dimension. However, we do not use their result in this paper.

When f is a dominant *endomorphism* (i.e., f is defined everywhere), the existence of the limit defining the arithmetic degree was proved in [Kawaguchi and Silverman 2016a]. But in general, the convergence is not known. It seems difficult at the moment to prove Conjecture 1.1 in full generality.

In this paper, we prove Conjecture 1.1 for any endomorphism on any smooth projective surface.

Theorem 1.3. *Let k be a number field, X a smooth projective surface over $\bar{k}$, and $f : X \to X$ a surjective endomorphism on X. Then Conjecture 1.1 holds for f.*

As by-products of our arguments, we also obtain the following two cases for which Conjecture 1.1 holds:

Theorem 1.4 (Theorem 3.6). *Let k be a number field, X a smooth projective irrational surface over $\bar{k}$, and $f : X \dashrightarrow X$ a birational automorphism on X. Then Conjecture 1.1 holds for f.*

Theorem 1.5 (Theorem 3.7). *Let k be a number field, X a smooth projective toric variety over $\bar{k}$, and $f : X \to X$ a toric surjective endomorphism on X. Then Conjecture 1.1 holds for f.*

Lin [2018] gives a precise description of the arithmetic degrees of toric self-maps on toric varieties.

As we will see in the proof of Theorem 1.3, there does not always exist a Zariski dense orbit for a given self-map. For instance, a self-map cannot have a Zariski dense orbit if it is a self-map over a variety of positive Kodaira dimension. So it is also important to consider whether a self-map has a $\bar{k}$-rational point whose orbit has full arithmetic complexity, that is, whose arithmetic degree coincides with the dynamical degree. We prove that such a point always exists for any surjective endomorphism on any smooth projective variety.

Theorem 1.6. *Let k be a number field, X a smooth projective variety over $\bar{k}$, and $f: X \to X$ a surjective endomorphism on X. Then there exists a $\bar{k}$-rational point $P \in X(\bar{k})$ such that $\alpha_f(P) = \delta_f$.*

If f is an automorphism, we can construct a "large" collection of points whose orbits have full arithmetic complexity.

Theorem 1.7. *Let k be a number field, X a smooth projective variety over $\bar{k}$, and $f: X \to X$ an automorphism. Then there exists a subset $S \subset X(\bar{k})$ which satisfies all of the following conditions*:

(1) *For every $P \in S$, $\alpha_f(P) = \delta_f$.*

(2) *For $P, Q \in S$ with $P \neq Q$, $\mathcal{O}_f(P) \cap \mathcal{O}_f(Q) = \varnothing$.*

(3) *S is Zariski dense in X.*

Remark 1.8. Kawaguchi, Silverman, and the second author proved Conjecture 1.1 in the following cases:

(1) f is an endomorphism and the Néron–Severi group of X has rank one [Kawaguchi and Silverman 2014, Theorem 2(a)].

(2) f is the extension to $\mathbb{P}^N$ of a regular affine automorphism on $\mathbb{A}^N$ [Kawaguchi and Silverman 2014, Theorem 2(b)].

(3) X is a smooth projective surface and f is an automorphism on X [Kawaguchi 2008, Theorem A; Kawaguchi and Silverman 2014, Theorem 2(c)].

(4) f is the extension to $\mathbb{P}^N$ of a monomial endomorphism on $\mathbb{G}_m^N$ and $P \in \mathbb{G}_m^N(\bar{k})$ [Silverman 2014, Proposition 19] .

(5) X is an abelian variety. Note that any rational map between abelian varieties is automatically a morphism [Kawaguchi and Silverman 2016a, Corollary 31; Silverman 2017, Theorem 2].

(6) f is an endomorphism and X is the product $\prod_{i=1}^n X_i$ of smooth projective varieties, with the assumption that each variety X_i satisfies one of the following conditions [Sano 2016, Theorem 1.3]:
   - The first Betti number of $(X_i)_\mathbb{C}$ is zero and the Néron–Severi group of X_i has rank one.
   - X_i is an abelian variety.
   - X_i is an Enriques surface.
   - X_i is a $K3$ surface.

(7) f is an endomorphism and X is the product $X_1 \times X_2$ of positive dimensional varieties such that one of X_1 or X_2 is of general type. (In fact, there do not exist Zariski dense forward f-orbits on such $X_1 \times X_2$.) [Sano 2016, Theorem 1.4]

Notation. Throughout this paper:

- We fix a number field k.
- A *variety* always means an integral separated scheme of finite type over $\bar{k}$.
- A *divisor* on a variety X means a divisor on X defined over $\bar{k}$.
- An *endomorphism* on a variety X means a morphism from X to itself defined over $\bar{k}$. A *noninvertible endomorphism* is a surjective endomorphism which is not an automorphism.
- A *curve* or *surface* simply means a smooth projective variety of dimension 1 or 2, respectively, unless otherwise stated.
- For any curve C, the genus of C is denoted by $g(C)$.
- When we say that P is a point of X or write as $P \in X$, it means that P is a $\bar{k}$-rational point of X.
- The Néron–Severi group of a smooth projective variety X is denoted by $\mathrm{NS}(X)$. It is well-known that $\mathrm{NS}(X)$ is a finitely generated abelian group. We put $\mathrm{NS}(X)_{\mathbb{R}} := \mathrm{NS}(X) \otimes_{\mathbb{Z}} \mathbb{R}$.
- The symbols $\equiv$, $\sim$, $\sim_{\mathbb{Q}}$ and $\sim_{\mathbb{R}}$ mean algebraic equivalence, linear equivalence, $\mathbb{Q}$-linear equivalence, and $\mathbb{R}$-linear equivalence, respectively.
- Let X be a smooth projective variety and $f\colon X \dashrightarrow X$ a dominant rational self-map. A point $P \in X_f(\bar{k})$ is called *preperiodic* if the forward f-orbit $\mathcal{O}_f(P)$ of P is a finite set. This is equivalent to the condition that $f^n(P) = f^m(P)$ for some $n, m \geq 0$ with $n \neq m$.
- Let f, g and h be real-valued functions on a domain S. The equality $f = g + O(h)$ means that there is a positive constant C such that $|f(x) - g(x)| \leq C|h(x)|$ for every $x \in S$. The equality $f = g + O(1)$ means that there is a positive constant C' such that $|f(x) - g(x)| \leq C'$ for every $x \in S$.

Outline of this paper. In Section 2, we recall the definitions and some properties of dynamical and arithmetic degrees. In Section 3, at first we recall some lemmata about reduction for Conjecture 1.1, which were proved in [Sano 2016; Silverman 2017]. Then, we prove the birational invariance of arithmetic degree. As its corollary, we prove Theorem 1.4 by reducing to the automorphism case, using minimal models. We also prove Theorem 1.5. In Section 4, by using the Enriques classification of smooth projective surfaces, we reduce Theorem 1.3 to three cases, i.e., the case of $\mathbb{P}^1$-bundles, hyperelliptic surfaces, and surfaces of Kodaira dimension one. In Section 5 we recall fundamental properties of $\mathbb{P}^1$-bundles over curves. In Sections 6, 7, and 8, we prove Theorem 1.3 in each case explained in Section 4. Finally, in Section 9, we prove Theorems 1.6 and 1.7. In the proof of Theorem 1.6, we use a nef $\mathbb{R}$-divisor D that satisfies $f^*D \equiv \delta_f D$.

2. Dynamical degree and arithmetic degree

Let H be an ample divisor on a smooth projective variety X. The (*first*) *dynamical degree* of a dominant rational self-map $f\colon X \dashrightarrow X$ is defined by

$$\delta_f := \lim_{n\to\infty}((f^n)^*H\cdot H^{\dim X-1})^{1/n}.$$

The limit defining δ_f exists, and δ_f does not depend on the choice of H [Dinh and Sibony 2005, Corollary 7; Guedj 2005, Proposition 1.2]. Note that if f is an endomorphism, we have $(f^n)^* = (f^*)^n$ as a linear self-map on $\mathrm{NS}(X)$. But if f is merely a rational self-map, then $(f^n)^* \neq (f^*)^n$ in general.

Remark 2.1 [Dinh and Sibony 2005, Proposition 1.2(iii); Kawaguchi and Silverman 2016b, Remark 7]. Let $\rho((f^n)^*)$ be the spectral radius of the linear self-map $(f^n)^*\colon\ \mathrm{NS}(X)_{\mathbb{R}} \to \mathrm{NS}(X)_{\mathbb{R}}$. The dynamical degree δ_f is equal to the limit $\lim_{n\to\infty}(\rho((f^n)^*))^{1/n}$. Thus we have $\delta_{f^n} = \delta_f^n$ for every $n \geq 1$.

Let $X_f(\bar{k})$ be the set of points P on X such that f is defined at $f^n(P)$ for every $n \geq 0$. The *arithmetic degree* of f at a point $P \in X_f(\bar{k})$ is defined as follows. Let

$$h_H\colon X(\bar{k}) \to [0,\infty)$$

be the (absolute logarithmic) Weil height function associated with H [Hindry and Silverman 2000, Theorem B3.2]. We put

$$h_H^+(P) := \max\{h_H(P), 1\}.$$

We call

$$\overline{\alpha}_f(P) := \limsup_{n\to\infty} h_H^+(f^n(P))^{1/n} \quad \text{and} \quad \underline{\alpha}_f(P) := \liminf_{n\to\infty} h_H^+(f^n(P))^{1/n}$$

the upper arithmetic degree and *the lower arithmetic degree* of f at P, respectively. It is known that $\overline{\alpha}_f(P)$ and $\underline{\alpha}_f(P)$ do not depend on the choice of H [Kawaguchi and Silverman 2016b, Proposition 12]. If $\overline{\alpha}_f(P) = \underline{\alpha}_f(P)$, the limit

$$\alpha_f(P) := \lim_{n\to\infty} h_H^+(f^n(P))^{1/n}$$

is called *the arithmetic degree of f at P*.

Remark 2.2. Let D be a divisor on X and f a dominant rational self-map on X. Take $P \in X_f(\bar{k})$. Then we can easily check that

$$\overline{\alpha}_f(P) \geq \limsup_{n\to\infty} h_D^+(f^n(P))^{1/n} \quad \text{and} \quad \underline{\alpha}_f(P) \geq \liminf_{n\to\infty} h_D^+(f^n(P))^{1/n}.$$

So when these limits exist, we have

$$\alpha_f(P) \geq \lim_{n\to\infty} h_D^+(f^n(P))^{1/n}.$$

Remark 2.3. When f is an endomorphism, the existence of the limit defining the arithmetic degree $\alpha_f(P)$ was proved by Kawaguchi and Silverman [2016a, Theorem 3]. But it is not known in general.

Remark 2.4. The inequality $\overline{\alpha}_f(P) \le \delta_f$ was proved by Kawaguchi and Silverman, and the third author [Kawaguchi and Silverman 2016b, Theorem 4; Matsuzawa 2016, Theorem 1.4]. Hence, in order to prove Conjecture 1.1, it is enough to prove the opposite inequality $\underline{\alpha}_f(P) \ge \delta_f$.

3. Some reductions for Conjecture 1.1

Reductions. We recall some lemmata which are useful to reduce the proof of some cases of Conjecture 1.1 to easier cases.

Lemma 3.1. *Let X be a smooth projective variety and $f\colon X \to X$ a surjective endomorphism. Then Conjecture 1.1 holds for f if and only if Conjecture 1.1 holds for f^t for some $t \ge 1$.*

Proof. See [Sano 2016, Lemma 3.3]. □

Lemma 3.2 [Silverman 2017, Lemma 6]. *Let $\psi\colon X \to Y$ be a finite morphism between smooth projective varieties. Let $f_X\colon X \to X$ and $f_Y\colon Y \to Y$ be surjective endomorphisms on X and Y, respectively. Assume that $\psi \circ f_X = f_Y \circ \psi$.*

(i) *For any $P \in X(\bar{k})$, we have $\alpha_{f_X}(P) = \alpha_{f_Y}(\psi(P))$.*

(ii) *Assume that ψ is surjective. Then Conjecture 1.1 holds for f_X if and only if Conjecture 1.1 holds for f_Y.*

Proof. (i) Take any point $P \in X(\bar{k})$. Let H be an ample divisor on Y. Then ψ^*H is an ample divisor on X. Hence we have

$$\begin{aligned}\alpha_{f_X}(P) &= \lim_{n\to\infty} h^+_{\psi^*H}(f_X^n(P))^{1/n}\\ &= \lim_{n\to\infty} h^+_H(\psi \circ f_X^n(P))^{1/n}\\ &= \lim_{n\to\infty} h^+_H(f_Y^n \circ \psi(P))^{1/n}\\ &= \alpha_{f_Y}(\psi(P)).\end{aligned}$$

(ii) For a point $P \in X(\bar{k})$, the forward f_X-orbit $\mathcal{O}_{f_X}(P)$ is Zariski dense in X if and only if the forward f_Y-orbit $\mathcal{O}_{f_Y}(\psi(P))$ is Zariski dense in Y since ψ is a finite surjective morphism. Moreover we have $\dim X = \dim Y$. So we obtain

$$\begin{aligned}\delta_{f_X} &= \lim_{n\to\infty} ((f_X^n)^*\psi^*H \cdot (\psi^*H)^{\dim X-1})^{1/n}\\ &= \lim_{n\to\infty} (\psi^*(f_Y^n)^*H \cdot (\psi^*H)^{\dim Y-1})^{1/n}\\ &= \lim_{n\to\infty} (\deg(\psi)((f_Y^n)^*H \cdot H^{\dim Y-1}))^{1/n}\\ &= \delta_{f_Y}.\end{aligned}$$

Therefore the assertion follows. □

Birational invariance of the arithmetic degree. We show that the arithmetic degree is invariant under birational conjugacy.

Lemma 3.3. *Let $\mu\colon X \dashrightarrow Y$ be a birational map of smooth projective varieties. Take Weil height functions h_X and h_Y associated with ample divisors H_X and H_Y on X and Y, respectively. Then there are constants $M \in \mathbb{R}_{>0}$ and $M' \in \mathbb{R}$ such that*

$$h_X(P) \geq M h_Y(\mu(P)) + M'$$

for any $P \in X(\bar{k}) \setminus I_\mu(\bar{k})$.

Proof. Replacing H_Y by a positive multiple, we may assume that H_Y is very ample. Take a smooth projective variety Z and a birational morphism $p\colon Z \to X$ such that p is isomorphic over $X \setminus I_\mu$ and $q = \mu \circ p\colon Z \to Y$ is a morphism. Let $\{F_i\}_{i=1}^r$ be the collection of prime p-exceptional divisors. We take H_Y as not containing $q(F_i)$ for any i, so q^*H_Y does not contain F_i for any i. Then $E = p^*p_*q^*H_Y - q^*H_Y$ is an effective divisor contained in the exceptional locus of p. Take a sufficiently large integer N such that $NH_X - p_*q^*H_Y$ is very ample. Then, for $P \in X(\bar{k}) \setminus I_\mu$, we have

$$\begin{aligned} h_X(P) &= \tfrac{1}{N}(h_{NH_X - p_*q^*H_Y}(P) + h_{p_*q^*H_Y}(P)) + O(1) \\ &\geq \tfrac{1}{N} h_{p_*q^*H_Y}(P) + O(1) \\ &= \tfrac{1}{N} h_{p^*p_*q^*H_Y}(p^{-1}(P)) + O(1) \\ &= \tfrac{1}{N} h_{q^*H_Y}(p^{-1}(P)) + h_E(p^{-1}(P)) + O(1) \\ &= \tfrac{1}{N} h_Y(\mu(P)) + h_E(p^{-1}(P)) + O(1). \end{aligned}$$

We know that $h_E \geq O(1)$ on $Z(\bar{k}) \setminus \operatorname{Supp} E$ [Hindry and Silverman 2000, Theorem B.3.2(e)]. Since $\operatorname{Supp} E \subset p^{-1}(I_\mu)$, $h_E(p^{-1}(P)) \geq O(1)$ for $P \in X(\bar{k}) \setminus I_\mu$. Finally, we obtain that

$$h_X(P) \geq (1/N) h_Y(\mu(P)) + O(1) \quad \text{for } P \in X(\bar{k}) \setminus I_\mu. \qquad \square$$

Theorem 3.4. *Let $f\colon X \dashrightarrow X$ and $g\colon Y \dashrightarrow Y$ be dominant rational self-maps on smooth projective varieties and $\mu\colon X \dashrightarrow Y$ a birational map such that $g \circ \mu = \mu \circ f$.*

(i) *Let $U \subset X$ be a Zariski open subset such that $\mu|_U\colon U \to \mu(U)$ is an isomorphism. Then $\overline{\alpha}_f(P) = \overline{\alpha}_g(\mu(P))$ and $\underline{\alpha}_f(P) = \underline{\alpha}_g(\mu(P))$ for $P \in X_f(\bar{k}) \cap \mu^{-1}(Y_g(\bar{k}))$ such that $\mathcal{O}_f(P) \subset U(\bar{k})$.*

(ii) *Take $P \in X_f(\bar{k}) \cap \mu^{-1}(Y_g(\bar{k}))$. Assume that $\mathcal{O}_f(P)$ is Zariski dense in X and both $\alpha_f(P)$ and $\alpha_g(\mu(P))$ exist. Then $\alpha_f(P) = \alpha_g(\mu(P))$.*

Proof. (i) Using Lemma 3.3 for both μ and μ^{-1}, there are constants $M_1, L_1 \in \mathbb{R}_{>0}$ and $M_2, L_2 \in \mathbb{R}$ such that

$$M_1 h_Y(\mu(P)) + M_2 \leq h_X(P) \leq L_1 h_Y(\mu(P)) + L_2 \tag{$*$}$$

for $P \in U(\bar{k})$. The claimed equalities follow from $(*)$.

(ii) Since $\mathcal{O}_f(P)$ is Zariski dense in X, we can take a subsequence $\{f^{n_k}(P)\}_k$ of $\{f^n(P)\}_n$ contained in U. Using $(*)$ again, it follows that

$$\alpha_f(P) = \lim_{k\to\infty} h_X^+(f^{n_k}(P))^{1/n_k} = \lim_{k\to\infty} h_Y^+(g^{n_k}(\mu(P)))^{1/n_k} = \alpha_g(\mu(P)). \qquad \square$$

Remark 3.5. Silverman [2014] dealt with a height function on $\mathbb{G}_m^n$ induced by an open immersion $\mathbb{G}_m^n \hookrightarrow \mathbb{P}^n$ and proved Conjecture 1.1 for monomial maps on $\mathbb{G}_m^n$. It seems that it has not been checked in the literature that the arithmetic degrees of endomorphisms on quasiprojective varieties does not depend on the choice of open immersions to projective varieties. Now by Theorem 3.4, the arithmetic degree of a rational self-map on a quasiprojective variety at a point does not depend on the choice of an open immersion of the quasiprojective variety to a projective variety. Furthermore, by the birational invariance of dynamical degrees, we can state Conjecture 1.1 for rational self-maps on quasiprojective varieties, such as semiabelian varieties.

Applications of the birational invariance. In this subsection, we prove Theorems 1.4 and 1.5 as applications of Theorem 3.4.

Theorem 3.6 (Theorem 1.4). *Let X be an irrational surface and $f\colon X \dashrightarrow X$ a birational automorphism on X. Then Conjecture 1.1 holds for f.*

Proof. Take a point $P \in X_f(\bar{k})$. If $\mathcal{O}_f(P)$ is finite, the limit $\alpha_f(P)$ exists and is equal to 1. Next, assume that the closure $\overline{\mathcal{O}_f(P)}$ of $\mathcal{O}_f(P)$ has dimension 1. Let Z be the normalization of $\overline{\mathcal{O}_f(P)}$ and $\nu\colon Z \to X$ the induced morphism. Then an endomorphism $g\colon Z \to Z$ satisfying $\nu \circ g = f \circ \nu$ is induced. Take a point $P' \in Z$ such that $\nu(P') = P$. Then $\alpha_g(P') = \alpha_f(P)$ since ν is finite by Lemma 3.2 (i). It follows from [Kawaguchi and Silverman 2016a, Theorem 2] that $\alpha_g(P')$ exists (note that their theorem holds for possibly nonsurjective endomorphisms on possibly reducible normal varieties). Therefore $\alpha_f(P)$ exists.

Finally, assume that $\mathcal{O}_f(P)$ is Zariski dense. If $\delta_f = 1$, then $1 \le \underline{\alpha}_f(P) \le \overline{\alpha}_f(P) \le \delta_f = 1$ by Remark 2.4, so $\alpha_f(P)$ exists and $\alpha_f(P) = \delta_f = 1$. So we may assume that $\delta_f > 1$. Since X is irrational and $\delta_f > 1$, $\kappa(X)$ must be nonnegative [Diller and Favre 2001, Theorem 0.4, Proposition 7.1 and Theorem 7.2]. Take a birational morphism $\mu\colon X \to Y$ to the minimal model Y of X and let $g\colon Y \dashrightarrow Y$ be the birational automorphism on Y defined as $g = \mu \circ f \circ \mu^{-1}$. Then g is in fact an automorphism since, if g has indeterminacy, Y must have a K_Y-negative curve. It is obvious that $\mathcal{O}_g(\mu(P))$ is also Zariski dense in Y. Since $\mu(\mathrm{Exc}(\mu))$ is a finite set, there is a positive integer n_0 such that $\mu(f^n(P)) = g^n(\mu(P)) \notin \mu(\mathrm{Exc}(\mu))$ for $n \ge n_0$. So we have $f^n(P) \notin \mathrm{Exc}(\mu)$ for $n \ge n_0$. Replacing P by $f^{n_0}(P)$, we may assume that $\mathcal{O}_f(P) \subset X \setminus \mathrm{Exc}(\mu)$. Applying Theorem 3.4 (i) to P, it follows that $\alpha_f(P) = \alpha_g(\mu(P))$. We know that $\alpha_g(\mu(P))$ exists since g is a morphism. So $\alpha_f(P)$ also exists. The equality $\alpha_g(\mu(P)) = \delta_g$ holds as a consequence of Conjecture 1.1 for automorphisms on surfaces (see Remark 1.8(3)). Since the dynamical degree is invariant under birational conjugacy, it follows that $\delta_g = \delta_f$. So we obtain the equality $\alpha_f(P) = \delta_f$. $\square$

Theorem 3.7 (Theorem 1.5). *Let X be a smooth projective toric variety and $f\colon X \to X$ a toric surjective endomorphism on X. Then Conjecture 1.1 holds for f.*

Proof. Let $\mathbb{G}_m^d \subset X$ be the torus embedded as an open dense subset in X. Then $f|_{\mathbb{G}_m^d} : \mathbb{G}_m^d \to \mathbb{G}_m^d$ is a homomorphism of algebraic groups by assumption. Let $\mathbb{G}_m^d \subset \mathbb{P}^d$ be the natural embedding of $\mathbb{G}_m^d$ to the projective space $\mathbb{P}^d$ and $g : \mathbb{P}^d \dashrightarrow \mathbb{P}^d$ be the induced rational self-map. Then g is a monomial map.

Take $P \in X(\bar{k})$ such that $\mathcal{O}_f(P)$ is Zariski dense. Note that $\alpha_f(P)$ exists since f is a morphism. Since $\mathcal{O}_f(P)$ is Zariski dense and $f(\mathbb{G}_m^d) \subset \mathbb{G}_m^d$, there is a positive integer n_0 such that $f^n(P) \in \mathbb{G}_m^d$ for $n \geq n_0$. By replacing P by $f^{n_0}(P)$, we may assume that $\mathcal{O}_f(P) \subset \mathbb{G}_m^d$. Applying Theorem 3.4 (i) to P, it follows that $\alpha_f(P) = \alpha_g(P)$.

The equality $\alpha_g(P) = \delta_g$ holds as a consequence of Conjecture 1.1 for monomial maps (see Remark 1.8(4)). Since the dynamical degree is invariant under birational conjugacy, it follows that $\delta_g = \delta_f$. So we obtain the equality $\alpha_f(P) = \delta_f$. □

4. Endomorphisms on surfaces

We start to prove Theorem 1.3. Since Conjecture 1.1 for automorphisms on surfaces is already proved by Kawaguchi (see Remark 1.8(3)), it is sufficient to prove Theorem 1.3 for *noninvertible* endomorphisms, that is, surjective endomorphisms which are not automorphisms.

Let $f : X \to X$ be a noninvertible endomorphism on a surface. We divide the proof of Theorem 1.3 according to the Kodaira dimension of X:

(I) $\kappa(X) = -\infty$; we need the following result due to Nakayama.

Lemma 4.1 [Nakayama 2002, Proposition 10]. *Let $f : X \to X$ be a noninvertible endomorphism on a surface X with $\kappa(X) = -\infty$. Then there is a positive integer m such that $f^m(E) = E$ for any irreducible curve E on X with negative self-intersection.*

Let $\mu : X \to X'$ be the contraction of a (-1)-curve E on X. By Lemma 4.1, there is a positive integer m such that $f^m(E) = E$. Then f^m induces an endomorphism $f' : X' \to X'$ such that $\mu \circ f^m = f' \circ \mu$. Using Lemma 3.1 and Theorem 3.4, the assertion of Theorem 1.3 for f follows from that for f'. Continuing this process, we may assume that X is relatively minimal.

When X is irrational and relatively minimal, X is a $\mathbb{P}^1$-bundle over a curve C with $g(C) \geq 1$.

When X is rational and relatively minimal, X is isomorphic to $\mathbb{P}^2$ or the Hirzebruch surface $\mathbb{F}_n = \mathbb{P}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(-n))$ for some $n \geq 0$ with $n \neq 1$. Note that Conjecture 1.1 holds for surjective endomorphisms on projective spaces (see Remark 1.8(1)).

(II) $\kappa(X) = 0$; for surfaces with nonnegative Kodaira dimension, we use the following result due to Fujimoto.

Lemma 4.2 [Fujimoto 2002, Lemma 2.3 and Proposition 3.1]. *Let $f : X \to X$ be a noninvertible endomorphism on a surface X with $\kappa(X) \geq 0$. Then X is minimal and f is étale.*

So X is either an abelian surface, a hyperelliptic surface, a K3 surface, or an Enriques surface. Since f is étale, we have $\chi(X, \mathcal{O}_X) = \deg(f)\chi(X, \mathcal{O}_X)$. Now $\deg(f) \geq 2$ by assumption, so $\chi(X, \mathcal{O}_X) = 0$ [Fujimoto 2002, Corollary 2.4]. Hence X must be either an abelian surface or a hyperelliptic surface because K3 surfaces and Enriques surfaces have nonzero Euler characteristics. Note that Conjecture 1.1 is valid for endomorphisms on abelian varieties (see Remark 1.8(5)).

(III) $\kappa(X) = 1$; this case will be treated in Section 8.

(IV) $\kappa(X) = 2$; the following fact is well known.

Lemma 4.3. *Let X be a smooth projective variety of general type. Then any surjective endomorphism on X is an automorphism. Furthermore, the group of automorphisms* $\mathrm{Aut}(X)$ *on X has finite order.*

Proof. See [Fujimoto 2002, Proposition 2.6], [Iitaka 1982, Theorem 11.12], or [Matsumura 1963, Corollary 2]. □

So there is no noninvertible endomorphism on X. As a summary, the remaining cases for the proof of Theorem 1.3 are the following:

- Noninvertible endomorphisms on $\mathbb{P}^1$-bundles over a curve.
- Noninvertible endomorphisms on hyperelliptic surfaces.
- Noninvertible endomorphisms on surfaces of Kodaira dimension 1.

These three cases are studied in Sections 5–8 below.

Remark 4.4. Fujimoto and Nakayama gave a complete classification of surfaces which admit noninvertible endomorphisms (see [Fujimoto 2002, Proposition 3.3], [Fujimoto and Nakayama 2008, Theorem 1.1], [Fujimoto and Nakayama 2005, Appendix to Section 4], and [Nakayama 2002, Theorem 3]).

5. Some properties of $\mathbb{P}^1$-bundles over curves

In this section, we recall and prove some properties of $\mathbb{P}^1$-bundles (see [Hartshorne 1977, Chapter V.2] or [Homma 1992; 1999] for details). In this section, let X be a $\mathbb{P}^1$-bundle over a curve C. Let $\pi : X \to C$ be the projection.

Proposition 5.1. *We can represent X as $X \cong \mathbb{P}(\mathcal{E})$, where $\mathcal{E}$ is a locally free sheaf of rank 2 on C such that $H^0(\mathcal{E}) \neq 0$ but $H^0(\mathcal{E} \otimes \mathcal{L}) = 0$ for all invertible sheaves $\mathcal{L}$ on C with $\deg \mathcal{L} < 0$. The integer $e := -\deg \mathcal{E}$ does not depend on the choice of such $\mathcal{E}$. Furthermore, there is a section $\sigma : C \to X$ with image C_0 such that $\mathcal{O}_X(C_0) \cong \mathcal{O}_X(1)$.*

Proof. See [Hartshorne 1977, Proposition 2.8]. □

Lemma 5.2. *The Picard group and the Néron–Severi group of X have the following structure*:

$$\mathrm{Pic}(X) \cong \mathbb{Z} \oplus \pi^* \mathrm{Pic}(C) \quad \textit{and} \quad \mathrm{NS}(X) \cong \mathbb{Z} \oplus \pi^* \mathrm{NS}(C) \cong \mathbb{Z} \oplus \mathbb{Z}.$$

Furthermore, the image C_0 of the section $\sigma : C \to X$ in Proposition 5.1 generates the first direct factor of $\mathrm{Pic}(X)$ *and* $\mathrm{NS}(X)$.

Proof. See [Hartshorne 1977, V, Proposition 2.3]. □

Lemma 5.3. *Let $F \in \mathrm{NS}(X)$ be a fiber $\pi^{-1}(p) = \pi^* p$ over a point $p \in C(\bar{k})$, and e the integer defined in Proposition 5.1. Then the intersection numbers of generators of* $\mathrm{NS}(X)$ *are as follows*:

$$F \cdot F = 0, \quad F \cdot C_0 = 1, \quad C_0 \cdot C_0 = -e.$$

Proof. It is easy to see that the equalities $F \cdot F = 0$ and $F \cdot C_0 = 1$ hold. For the last equality, see [Hartshorne 1977, V, Proposition 2.9]. ☐

We say that f *preserves fibers* if there is an endomorphism f_C on C such that $\pi \circ f = f_C \circ \pi$. In our situation, since there is a section $\sigma : C \to X$, f preserves fibers if and only if, for any point $p \in C$, there is a point $q \in C$ such that $f(\pi^{-1}(p)) \subset \pi^{-1}(q)$.

The following lemma appears in [Amerik 2003, p.18] in a more general form. But we need it only in the case of $\mathbb{P}^1$-bundles on a curve, and the proof in the general case is similar to our case. So we deal only with the case of $\mathbb{P}^1$- bundles on a curve.

Lemma 5.4. *For any surjective endomorphism f on X, the iterate f^2 preserves fibers.*

Proof. By the projection formula, the fibers of $\pi : X \to C$ can be characterized as connected curves having intersection number zero with any fiber $F_p = \pi^* p$, $p \in C$. Hence, to check that the iterate f^2 sends fibers to fibers, it suffices to show that $(f^2)^*(\pi^* \mathrm{NS}(C)_{\mathbb{R}}) = \pi^* \mathrm{NS}(C)_{\mathbb{R}}$. Now $\dim \mathrm{NS}(X)_{\mathbb{R}} = 2$ and the set of the numerical classes in X with self-intersection zero forms two lines, one of which is $\pi^* \mathrm{NS}(C)_{\mathbb{R}}$, and f^* fixes or interchanges them. So $(f^2)^*$ fixes $\pi^* \mathrm{NS}(C)_{\mathbb{R}}$. ☐

The following might be well-known, but we give a proof for the reader's convenience.

Lemma 5.5. *A surjective endomorphism f preserves fibers if and only if there exists a nonzero integer a such that $f^* F \equiv aF$. Here, F is the numerical class of a fiber.*

Proof. Assume $f^* F \equiv aF$. For any point $p \in C$, we set $F_p := \pi^{-1}(p) = \pi^* p$. If f does not preserve fibers, there is a point $p \in C$ such that $f(F_p) \cdot F > 0$. Now we can calculate the intersection number as follows:

$$0 = F \cdot aF = F \cdot (f^* F) = F_p \cdot (f^* F) = (f_* F_p) \cdot F = \deg(f|_{F_p}) \cdot (f(F_p) \cdot F) > 0.$$

This is a contradiction. Hence f preserves fibers.

Next, assume that f preserves fibers. Write $f^* F = aF + bC_0$. Then we can also calculate the intersection number as follows:

$$b = F \cdot (aF + bC_0) = F \cdot f^* F = (f_* F) \cdot F = \deg(f|_F) \cdot (F \cdot F) = 0.$$

Further, by the injectivity of f^*, we have $a \neq 0$. The proof is complete. ☐

Lemma 5.6. *If $\mathcal{E}$ splits, i.e., if there is an invertible sheaf $\mathcal{L}$ on C such that $\mathcal{E} \cong \mathcal{O}_C \oplus \mathcal{L}$, the invariant e of $X = \mathbb{P}(\mathcal{E})$ is nonnegative.*

Proof. See [Hartshorne 1977, V, Example 2.11.3]. ☐

Lemma 5.7. *Assume that $e \geq 0$. Then for a divisor $D = aF + bC_0 \in \mathrm{NS}(X)$, the following properties are equivalent*:

- *D is ample.*
- *$a > be$ and $b > 0$.*

In other words, the nef cone of X is generated by F and $eF + C_0$.

Proof. See [Hartshorne 1977, V, Proposition 2.20]. □

We can prove a result stronger than Lemma 5.4 as follows.

Lemma 5.8. *Assume that $e > 0$. Then any surjective endomorphism $f\colon X \to X$ preserves fibers.*

Proof. By Lemma 5.5, it is enough to prove $f^*F \equiv aF$ for some integer $a > 0$. We can write $f^*F \equiv aF + bC_0$ for some integers $a, b \geq 0$.

Since we have

$$aF + bC_0 = (a - be)F + b(eF + C_0)$$

and f preserves the nef cone and the ample cone, either of the equalities $a - be = 0$ or $b = 0$ holds.

We have

$$0 = \deg(f)(F\cdot F) = (f_*f^*F)\cdot F = (f^*F)\cdot(f^*F) = (aF+bC_0)\cdot(aF+bC_0) = 2ab - b^2e = b(2a - be).$$

So either of the equalities $b = 0$ or $2a - be = 0$ holds.

If we have $b \neq 0$, we have $a - be = 0$ and $2a - be = 0$. So we get $a = 0$. But since $e \neq 0$, we obtain $b = 0$. This is a contradiction. Consequently, we get $b = 0$ and $f^*F \equiv aF$. □

Lemma 5.9. *Fix a fiber $F = F_p$ for a point $p \in C(\bar{k})$. Let f be a surjective endomorphism on X preserving fibers, f_C the endomorphism on C satisfying $\pi \circ f = f_C \circ \pi$, $f_F := f|_F\colon F \to f(F)$ the restriction of f to the fiber F. Set $f^*F \equiv aF$ and $f^*C_0 \equiv cF + dC_0$. Then we have $a = \deg(f_C)$, $d = \deg(f_F)$, $\deg(f) = ad$, and $\delta_f = \max\{a, d\}$.*

Proof. Our assertions follow from the following equalities of divisor classes in $\mathrm{NS}(X)$ and of intersection numbers:

$$aF = f^*F = f^*\pi^*p = \pi^*f_C^*p = \pi^*(\deg(f_C)p) = \deg(f_C)\pi^*p = \deg(f_C)F,$$

$$\begin{aligned}\deg(f)F = f_*f^*F = f_*f^*\pi^*p = f_*\pi^*f_C^*p = f_*\pi^*(\deg(f_C)p)\\ = \deg(f_C)f_*F = \deg(f_C)\deg(f_F)f(F) = \deg(f_C)\deg(f_F)F\end{aligned}$$

$$\deg(f) = \deg(f)C_0\cdot F = (f_*f^*C_0)\cdot F = (f^*C_0)\cdot(f^*F) = (cF + dC_0)\cdot aF = ad.$$

The last assertion $\delta_f = \max\{a, d\}$ follows from the equality $\delta_f = \lim_{n\to\infty}\rho((f^n)^*)^{1/n} = \rho(f^*)$ and from the functoriality of f^* (see Remark 2.1). □

Lemma 5.10. *Using the notation of Lemma 5.9, assume that $e \geq 0$. Then both F and C_0 are eigenvectors of $f^*\colon \mathrm{NS}(X)_{\mathbb{R}} \to \mathrm{NS}(X)_{\mathbb{R}}$. Further, if e is positive, then we have $\deg(f_C) = \deg(f_F)$.*

Proof. Set $f^*F = aF$ and $f^*C_0 = cF + dC_0$ in $\mathrm{NS}(X)$. Then we have

$$-ead = -e \deg f = (f_* f^* C_0) \cdot C_0 = (f^* C_0)^2 = (cF + dC_0)^2 = 2cd - ed^2.$$

Hence, we get $c = e(d-a)/2$. We have the following equalities in $\mathrm{NS}(X)$:

$$f^*(eF + C_0) = aeF + (cF + dC_0) = (ae + c)F + dC_0.$$

By the fact that f^*D is ample if and only if D is ample, it follows that $eF + C_0$ is an eigenvector of f^*. Thus, we have

$$de = ae + c = ae + e(d-a)/2 = e(d+a)/2.$$

Therefore, the equality $e(d-a) = 0$ holds. So $c = e(d-a)/2 = 0$ holds.

Further, we assume that $e > 0$. Then it follows that $d - a = 0$. So we have $\deg(f_C) = a = d = \deg(f_F)$. $\square$

The following lemma is used on page 1650.

Lemma 5.11. *Let $\mathcal{L}$ be a nontrivial invertible sheaf of degree* 0 *on a curve C with $g(C) \geq 1$, $\mathcal{E} = \mathcal{O}_C \oplus \mathcal{L}$, and $X = \mathbb{P}(\mathcal{E})$. Let C_0 and C_1 be sections corresponding to the projections $\mathcal{E} \to \mathcal{L}$ and $\mathcal{E} \to \mathcal{O}_C$. If $\sigma : C \to X$ is a section such that $(\sigma(C))^2 = 0$, then $\sigma(C)$ is equal to C_0 or C_1.*

Proof. Note that $e = 0$ in this case and thus $(C_0^2) = 0$. Moreover, $\mathcal{O}_X(C_0) \cong \mathcal{O}_X(1)$ and $\mathcal{O}_X(C_1) \cong \mathcal{O}_X(1) \otimes \pi^* \mathcal{L}^{-1}$. Set $\sigma(C) \equiv aC_0 + bF$. Then $a = (\sigma(C) \cdot F) = 1$ and $2ab = (\sigma(C)^2) = 0$. Thus $\sigma(C) \equiv C_0$. Therefore, $\mathcal{O}_X(\sigma(C)) \cong \mathcal{O}_X(C_0) \otimes \pi^* \mathcal{N}$ for some invertible sheaf $\mathcal{N}$ of degree 0 on C. Then

$$0 \neq H^0(X, \mathcal{O}_X(\sigma(C))) = H^0(C, \pi_* \mathcal{O}_X(C_0) \otimes \mathcal{N}) = H^0(C, (\mathcal{L} \oplus \mathcal{O}_C) \otimes \mathcal{N})$$

and this implies $\mathcal{N} \cong \mathcal{O}_C$ or $\mathcal{N} \cong \mathcal{L}^{-1}$. Hence $\mathcal{O}_X(\sigma(C))$ is isomorphic to $\mathcal{O}_X(C_0)$ or $\mathcal{O}_X(C_0) \otimes \pi^* \mathcal{L}^{-1} = \mathcal{O}_X(C_1)$. Since $\mathcal{L}$ is nontrivial, we have $H^0(\mathcal{O}_X(C_0)) = H^0(\mathcal{O}_X(C_1)) = \bar{k}$ and we get $\sigma(C) = C_0$ or C_1. $\square$

6. $\mathbb{P}^1$-bundles over curves

In this section, we prove Conjecture 1.1 for noninvertible endomorphisms on $\mathbb{P}^1$-bundles over curves. We divide the proof according to the genus of the base curve.

$\mathbb{P}^1$-bundles over $\mathbb{P}^1$.

Theorem 6.1. *Let $\pi : X \to \mathbb{P}^1$ be a $\mathbb{P}^1$-bundle over $\mathbb{P}^1$ and $f : X \to X$ be a noninvertible endomorphism. Then Conjecture 1.1 holds for f.*

Proof. Take a locally free sheaf $\mathcal{E}$ of rank 2 on $\mathbb{P}^1$ such that $X \cong \mathbb{P}(\mathcal{E})$ and $\deg \mathcal{E} = -e$ (see Proposition 5.1). Then $\mathcal{E}$ splits [Hartshorne 1977, V, Corollary 2.14]. When X is isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$, i.e., the case of $e = 0$, the assertion holds by [Sano 2016, Theorem 1.3]. When X is not isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$, i.e., the case of $e > 0$, the endomorphism f preserves fibers and induces an endomorphism $f_{\mathbb{P}^1}$ on the base curve $\mathbb{P}^1$. By Lemma 5.10, we have $\delta_f = \delta_{f_{\mathbb{P}^1}}$. Fix a point $p \in \mathbb{P}^1$ and set $F = \pi^* p$. Let $P \in X(\bar{k})$ be a point whose forward f-orbit is Zariski dense in X. Then the forward $f_{\mathbb{P}^1}$-orbit of $\pi(P)$ is also Zariski dense in $\mathbb{P}^1$. Now the assertion follows from the following computation.

$$\begin{aligned}\alpha_f(P) &\geq \lim_{n\to\infty} h_F(f^n(P))^{1/n} = \lim_{n\to\infty} h_{\pi^*p}(f^n(P))^{1/n} \\ &= \lim_{n\to\infty} h_p(\pi\circ f^n(P))^{1/n} = \lim_{n\to\infty} h_p(f^n_{\mathbb{P}^1}\circ\pi(P))^{1/n} = \delta_{f_{\mathbb{P}^1}} = \delta_f. \qquad \square\end{aligned}$$

$\mathbb{P}^1$-bundles over genus one curves. In this subsection, we prove Conjecture 1.1 for any endomorphisms on a $\mathbb{P}^1$-bundle on a curve C of genus one.

The following result is due to Amerik. Note that Amerik in fact proved it for $\mathbb{P}^1$-bundles over varieties of arbitrary dimension.

Lemma 6.2. *Let $X = \mathbb{P}(\mathcal{E})$ be a $\mathbb{P}^1$-bundle over a curve C. If X has a fiber-preserving surjective endomorphism whose restriction to a general fiber has degree greater than* 1, *then $\mathcal{E}$ splits into a direct sum of two line bundles after a finite base change. Furthermore*, *if $\mathcal{E}$ is semistable*, *then $\mathcal{E}$ splits into a direct sum of two line bundles after an étale base change.*

Proof. See [Amerik 2003, Theorem 2 and Proposition 2.4]. $\square$

Lemma 6.3. *Let E be a curve of genus one with an endomorphism $f\colon E\to E$. If $g\colon E'\to E$ is a finite étale covering of E*, *there exists a finite étale covering $h\colon E''\to E'$ and an endomorphism $f'\colon E''\to E''$ such that $f\circ g\circ h = g\circ h\circ f'$. Furthermore*, *we can take h as satisfying $E''=E$.*

Proof. At first, since E' is an étale covering of E, a genus one curve, E' is also a genus one curve. By fixing a rational point $p\in E'(\bar{k})$ and $g(p)\in E(\bar{k})$, these curves E and E' can be regarded as elliptic curves, and g can be regarded as an isogeny between elliptic curves. Let $h := \hat{g}\colon E\to E'$ be the dual isogeny of g. The morphism f is decomposed as $f=\tau_c\circ\psi$ for a homomorphism ψ and a translation map τ_c by $c\in E(\bar{k})$. Fix a rational point $c'\in E(\bar{k})$ such that $[\deg(g)](c')=c$ and consider the translation map $\tau_{c'}$, where $[\deg(g)]$ is the multiplication by $\deg(g)$. We set $f'=\tau_{c'}\circ\psi$. Then we have the following equalities.

$$f\circ g\circ h = \tau_c\circ\psi\circ g\circ\hat{g} = \tau_c\circ\psi\circ[\deg(g)] = \tau_c\circ[\deg(g)]\circ\psi = [\deg(g)]\circ\tau_{c'}\circ\psi = g\circ h\circ f'.$$

This is what we want. $\square$

Proposition 6.4. *Let $\mathcal{E}$ be a locally free sheaf of rank* 2 *on a genus one curve C and $X=\mathbb{P}(\mathcal{E})$. Suppose Conjecture 1.1 holds for any noninvertible endomorphism on X with $\mathcal{E}=\mathcal{O}_C\oplus\mathcal{L}$ where $\mathcal{L}$ is a line bundle of degree zero on C. Then Conjecture 1.1 holds for any noninvertible endomorphism on $X=\mathbb{P}(\mathcal{E})$ for any $\mathcal{E}$.*

Proof. By Lemmas 5.4 and 3.1, we may assume that f preserves fibers. We can prove Conjecture 1.1 in the case of $\deg(f|_F)=1$ in the same way as in the case of $g(C)=0$ since $\deg(f|_F)=1\leq\deg(f_C)$. Since we are considering the case of $g(C)=1$, if $\mathcal{E}$ is indecomposable, then $\mathcal{E}$ is semistable (see [Mukai 2003, 10.2(c), 10.49] or [Hartshorne 1977, V, Exercise 2.8(c)]). By Lemma 6.2, if $\deg(f|_F)>1$ and $\mathcal{E}$ is indecomposable, there is a finite étale covering $g\colon E\to C$ satisfying that $E\times_C X\cong\mathbb{P}(\mathcal{O}_E\oplus\mathcal{L})$ for an invertible sheaf $\mathcal{L}$ over E. Furthermore, by Lemma 6.3, we can take E equal to C and there is an endomorphism $f'_C\colon C \to C$ satisfying $f_C \circ g = g \circ f'_C$. Then by the universality of cartesian product $X \times_{C,g} C$, we have an induced endomorphism $f'\colon X \times_{C,g} C \to X \times_{C,g} C$. By Lemma 3.2, it is enough to prove Conjecture 1.1 for the endomorphism f'. Thus, we may assume that $\mathcal{E}$ is decomposable, i.e., $X \cong \mathbb{P}(\mathcal{O}_C \oplus \mathcal{L})$. Then the invariant e is nonnegative by Lemma 5.6. When e is positive, by the same method as the proof of Theorem 1.3 in the case of $g(C) = 0$, the proof is complete. When $e = 0$, we have $\deg \mathcal{L} = 0$ and the assertion holds by the assumption. □

In the rest of this subsection, we keep the following notation. Let C be a genus one curve and $\mathcal{L}$ an invertible sheaf on C with degree 0. Let $X = \mathbb{P}(\mathcal{O}_C \oplus \mathcal{L}) = \operatorname{Proj}(\operatorname{Sym}(\mathcal{O}_C \oplus \mathcal{L}))$ and $\pi : X \to C$ the projection. When $\mathcal{L}$ is trivial, we have $X \cong C \times \mathbb{P}^1$, and by [Sano 2016, Theorem 1.3], Conjecture 1.1 is true for X. Thus we may assume $\mathcal{L}$ is nontrivial. In this case, we have two sections of $\pi : X \to C$ corresponding to the projections $\mathcal{O}_C \oplus \mathcal{L} \to \mathcal{L}$ and $\mathcal{O}_C \oplus \mathcal{L} \to \mathcal{O}_C$. Let C_0 and C_1 denote the images of these sections. Then we have $\mathcal{O}_X(C_0) = \mathcal{O}_X(1)$ and $\mathcal{O}_X(C_1) = \mathcal{O}_X(1) \otimes \pi^*\mathcal{L}^{-1}$. Since $\mathcal{L}$ is nontrivial, we have $C_0 \neq C_1$. But since $\deg \mathcal{L} = 0$, C_0 and C_1 are numerically equivalent. Thus $(C_0 \cdot C_1) = (C_0^2) = 0$ and therefore $C_0 \cap C_1 = \varnothing$.

Let f be a noninvertible endomorphism on X such that there is a surjective endomorphism $f_C\colon C \to C$ with $\pi \circ f = f_C \circ \pi$.

Lemma 6.5. *When $\mathcal{L}$ is a torsion element of* $\operatorname{Pic} C$, *Conjecture 1.1 holds for f.*

Proof. We fix an algebraic group structure on C. Since $\mathcal{L}$ is torsion, there exists a positive integer $n > 0$ such that $[n]^*\mathcal{L} \cong \mathcal{O}_C$. Then the base change of $\pi : X \to C$ by $[n]\colon C \to C$ is the trivial $\mathbb{P}^1$-bundle $\mathbb{P}^1 \times C \to C$. Applying Lemma 6.3 to $g = [n]$, we get a finite morphism $h\colon C \to C$ such that the base change of $\pi : X \to C$ by $h\colon C \to C$ is $\mathbb{P}^1 \times C \to C$ and there exists a finite morphism $f'_C\colon C \to C$ with $f_C \circ h = h \circ f'_C$. Then f induces a noninvertible endomorphism $f'\colon \mathbb{P}^1 \times C \to \mathbb{P}^1 \times C$. By [Sano 2016, Theorem 1.3], Conjecture 1.1 holds for f'. By Lemma 3.2, Conjecture 1.1 holds also for f. □

Now, let F be the numerical class of a fiber of π. By Lemma 5.10, we have

$$f^*F \equiv aF \quad \text{and} \quad f^*C_0 \equiv bC_0$$

for some integers $a, b \geq 1$. Note that $a = \deg f_C$, $b = \deg f|_F$ and $ab = \deg f$ (see Lemma 5.9).

Lemma 6.6. (1) *One of the equalities $f(C_0) = C_0$, $f(C_0) = C_1$ or $f(C_0) \cap C_0 = f(C_0) \cap C_1 = \varnothing$ holds. The same is true for $f(C_1)$.*

(2) *If $f(C_0) \cap C_i = \varnothing$ for $i = 0, 1$, then the base change of $\pi : X \to C$ by $f_C : C \to C$ is isomorphic to $\mathbb{P}^1 \times C$. In particular, $f_C^*\mathcal{L} \cong \mathcal{O}_C$ and $\mathcal{L}$ is a torsion element of* $\operatorname{Pic} C$. *The same conclusion holds under the assumption that $f(C_1) \cap C_i = \varnothing$ for $i = 0, 1$.*

Proof. (1) Since $f^*C_i \equiv bC_i$, $C_0 \equiv C_1$ and $(C_0^2) = 0$, we have $(f_*C_i \cdot C_j) = 0$ for every i and j. Thus the assertion follows.

(2) Assume $f(C_0) \cap C_i = \varnothing$ for $i = 0, 1$. Consider the following Cartesian diagram:

$$\begin{tikzcd} Y \arrow[r, "g"] \arrow[d, "\pi'"'] & X \arrow[d, "\pi"] \\ C \arrow[r, "f_C"] & C \end{tikzcd}$$

Then Y is a $\mathbb{P}^1$-bundle over C associated with the vector bundle $\mathcal{O}_C \oplus f_C^* \mathcal{L}$. The pull-backs $C_i = g^{-1}(C_i)$, $i = 0, 1$ are sections of π'. By the projection formula, we have $(C_i'^2) = 0$. Let $\sigma\colon C \to X$ be the section with $\sigma(C) = C_0$. Since $\pi \circ f \circ \sigma = f_C$, we get a section $s\colon C \to Y$ of π'.

$$\begin{tikzcd} & C \arrow[d, "\sigma"] \arrow[ddl, "s"'] \arrow[dddl, bend right=40, "\mathrm{id}"'] \\ & X \arrow[d, "f"] \\ Y \arrow[r, "g"] \arrow[d, "\pi'"] & X \arrow[d, "\pi"] \\ C \arrow[r, "f_C"'] & C \end{tikzcd}$$

Note that $g(s(C)) = f(C_0) \neq C_0, C_1$. Thus $s(C)$, C_0' and C_1' are distinct sections of π'. Moreover, by the projection formula, we have $(s(C) \cdot C_0') = 0$. Thus we have three sections which are numerically equivalent to each other. Then Lemma 5.11 implies $f_C^* \mathcal{L} \cong \mathcal{O}_C$ and $Y \cong \mathbb{P}^1 \times C$. Since $f_C^*\colon \operatorname{Pic}^0 C \to \operatorname{Pic}^0 C$ is an isogeny, the kernel of f_C^* is finite and thus $\mathcal{L}$ is a torsion element of $\operatorname{Pic} C$. $\square$

Lemma 6.7. (1) *Suppose that*

- $\mathcal{L}$ *is nontorsion in* $\operatorname{Pic} C$,
- $f(C_0) = C_0$ *or* C_1, *and*
- $f(C_1) = C_0$ *or* C_1.

Then $f(C_0) = C_0$ *and* $f(C_1) = C_1$, *or* $f(C_0) = C_1$ *and* $f(C_1) = C_0$.

(2) *If the equalities* $f(C_0) = C_0$ *and* $f(C_1) = C_1$ *hold, then* $f^* C_i \sim_{\mathbb{Q}} bC_i$ *for* $i = 0$ *and* 1.

Proof. (1) Assume that $f(C_0) = C_0$ and $f(C_1) = C_0$. Then $f_* C_0 = aC_0$ and $f_* C_1 = aC_0$ as cycles. Since $f_C^*\colon \operatorname{Pic}^0 C \to \operatorname{Pic}^0 C$ is surjective, there exists a degree zero divisor M on C such that $f_C^* \mathcal{O}_C(M) \cong \mathcal{L}$. Then $C_1 \sim C_0 - \pi^* f_C^* M$. Hence

$$aC_0 = f_* C_1 \sim (f_* C_0 - f_* \pi^* f_C^* M) = (aC_0 - f_* \pi^* f_C^* M)$$

and

$$0 \sim f_* \pi^* f_C^* M \sim f_* f^* \pi^* M \sim (\deg f) \pi^* M.$$

Thus $\pi^* M$ is torsion and so is M. This implies that $\mathcal{L}$ is torsion, which contradicts the assumption.

The same argument shows that the case when $f(C_0) = C_1$ and $f(C_1) = C_1$ does not occur.

(2) In this case, we have $f_*C_0 \sim aC_0$. We can write $f^*C_0 \sim bC_0 + \pi^*D$ for some degree zero divisor D on C. Thus

$$(\deg f)C_0 \sim f_*f^*C_0 \sim abC_0 + f_*\pi^*D = (\deg f)C_0 + f_*\pi^*D$$

and $f_*\pi^*D \sim 0$. Since $f_C^*\colon \operatorname{Pic}^0 C \to \operatorname{Pic}^0 C$ is surjective, there exists a degree zero divisor D' on C such that $f_C^*D' \sim D$. Then

$$0 \sim f_*\pi^*D \sim f_*\pi^*f_C^*D' \sim f_*f^*\pi^*D' \sim (\deg f)\pi^*D'.$$

Hence $\pi^*D' \sim_{\mathbb{Q}} 0$ and $D' \sim_{\mathbb{Q}} 0$. Therefore $D \sim_{\mathbb{Q}} 0$ and $f^*C_0 \sim_{\mathbb{Q}} bC_0$.

Similarly, we have $f^*C_1 \sim_{\mathbb{Q}} bC_1$. ☐

Lemma 6.8. *Suppose $a < b$. If $f^*C_i \sim_{\mathbb{Q}} bC_i$ for $i = 0, 1$, the line bundle $\mathcal{L}$ is a torsion element of* $\operatorname{Pic} C$.

Proof. Let L be a divisor on C such that $\mathcal{O}_C(L) \cong \mathcal{L}$. Note that $C_1 \sim C_0 - \pi^*L$. Thus

$$f^*\pi^*L \sim f^*(C_0 - C_1) \sim_{\mathbb{Q}} bC_0 - bC_1 \sim b\pi^*L$$

and $f_C^*L \sim_{\mathbb{Q}} bL$ hold.

Thus, from the following lemma, $\mathcal{L}$ is a torsion element. ☐

Lemma 6.9. *Let a and b be integers such that $1 \le a < b$. Let C be a curve of genus one defined over an algebraically closed field k. Let $f_C\colon C \to C$ be an endomorphism of* $\deg f_C = a$. *If L is a divisor on C of degree* 0 *satisfying*

$$f_C^*L \sim_{\mathbb{Q}} bL,$$

the divisor L is a torsion element of $\operatorname{Pic}^0(C)$

Proof. By the definition of $\mathbb{Q}$-linear equivalence, we have $f_C^*rL \sim brL$ for some positive integer r. Since the curve C is of genus one, the group $\operatorname{Pic}^0(C)$ is an elliptic curve. Assume the (group) endomorphism

$$f_C^* - [b]\colon \operatorname{Pic}^0(C) \to \operatorname{Pic}^0(C)$$

is the 0 map. Then we have the equalities $a = \deg f_C = \deg f_C^* = \deg[b] = b^2$. But this contradicts to the inequality $1 \le a < b$. Hence the map $f_C^* - [b]$ is an isogeny, and $\operatorname{Ker}(f_C^* - [b]) \subset \operatorname{Pic}^0(C)$ is a finite group scheme. In particular, the order of $rL \in \operatorname{Ker}(f_C^* - [b])(k)$ is finite. Thus, L is a torsion element. ☐

Remark 6.10. We can actually prove the following. Let X be a smooth projective variety over $\overline{\mathbb{Q}}$ and $f\colon X \to X$ be a surjective morphism over $\overline{\mathbb{Q}}$ with first dynamical degree δ. If an $\mathbb{R}$-divisor D on X satisfies

$$f^*D \sim_{\mathbb{R}} \lambda D$$

for some $\lambda > \delta$, then one has $D \sim_{\mathbb{R}} 0$.

Sketch of the proof. Consider the canonical height

$$\hat{h}_D(P) = \lim_{n\to\infty} h_D(f^n(P))/\lambda^n$$

where h_D is a height associated with D [Call and Silverman 1993]. If $\hat{h}_D(P) \neq 0$ for some P, then we can prove $\overline{\alpha}_f(P) \geq \lambda$. This contradicts the fact $\delta \geq \overline{\alpha}_f(P)$ and the assumption $\lambda > \delta$. Thus one has $\hat{h}_D = 0$ and therefore $h_D = \hat{h}_D + O(1) = O(1)$. By a theorem of Serre, we get $D \sim_{\mathbb{R}} 0$ [Serre 1997, 2.9, Theorem]. □

Proposition 6.11. *Let $\mathcal{L}$ be an invertible sheaf of degree zero on a genus one curve C and $X = \mathbb{P}(\mathcal{O}_C \oplus \mathcal{L})$. For any noninvertible endomorphism $f : X \to X$, Conjecture 1.1 holds.*

Proof. By Lemmas 6.5 and 6.9 we may assume $a \geq b$. In this case, $\delta_f = a$ and Conjecture 1.1 can be proved as in the proof of Theorem 6.1. □

Proof of Theorem 1.3 for $\mathbb{P}^1$-bundles over genus one curves. As we argued at the first of Section 4, we may assume that the endomorphism $f : X \to X$ is not an automorphism. Then the assertion follows from Propositions 6.4 and 6.11. □

Remark 6.12. In the above setting, the line bundle $\mathcal{L}$ is actually an eigenvector for f_C^* up to linear equivalence. More precisely, for a $\mathbb{P}^1$-bundle $\pi : X = \mathbb{P}(\mathcal{O}_C \oplus \mathcal{L}) \to C$ over a curve C with $\deg \mathcal{L} = 0$ and an endomorphism $f : X \to X$ that induces an endomorphism $f_C : C \to C$, there exists an integer t such that $f_C^*\mathcal{L} \cong \mathcal{L}^t$. Indeed, let C_0 and C_1 be the sections defined above. Since $(f^*(C_0) \cdot C_0) = 0$, we can write $\mathcal{O}_X(f^{-1}(C_0)) \cong \mathcal{O}_X(mC_0) \otimes \pi^*\mathcal{N}$ for some integer m and degree zero line bundle $\mathcal{N}$ on C. Since

$$0 \neq H^0(\mathcal{O}_X(f^{-1}(C_0))) = H^0(\mathcal{O}_X(mC_0)\otimes\pi^*\mathcal{N}) = H^0(\mathrm{Sym}^m(\mathcal{O}_C \oplus \mathcal{L})\otimes\mathcal{N}) = \bigoplus_{i=0}^{m} H^0(\mathcal{L}^i \otimes \mathcal{N}),$$

we have $\mathcal{N} \cong \mathcal{L}^r$ for some $-m \leq r \leq 0$. Thus $f^*\mathcal{O}_X(C_0) \cong \mathcal{O}_X(mC_0)\otimes\pi^*\mathcal{L}^r$. The key is the calculation of global sections using projection formula. Since $\mathcal{O}_X(C_1) \cong \mathcal{O}_X(C_0)\otimes\pi^*\mathcal{L}^{-1}$, we have $\pi_*\mathcal{O}_X(mC_1) \cong \pi_*\mathcal{O}_X(mC_0)\otimes\mathcal{L}^{-m}$. Moreover, since C_0 and C_1 are numerically equivalent, we can similarly get $f^*\mathcal{O}_X(C_1) \cong \mathcal{O}_X(mC_0)\otimes\pi^*\mathcal{L}^s$ for some integer s. Thus, $f^*\pi^*\mathcal{L} \cong \pi^*\mathcal{L}^{r-s}$. Therefore, $\pi^* f_C^*\mathcal{L} \cong \pi^*\mathcal{L}^{r-s}$. Since $\pi^* : \operatorname{Pic} C \to \operatorname{Pic} X$ is injective, we get $f_C^*\mathcal{L} \cong \mathcal{L}^{r-s}$.

$\mathbb{P}^1$-bundles over curves of genus ≥ 2. By the following proposition, Conjecture 1.1 trivially holds in this case.

Proposition 6.13. *Let C be a curve with $g(C) \geq 2$ and $\pi : X \to C$ be a $\mathbb{P}^1$-bundle over C. Let $f : X \to X$ be a surjective endomorphism. Then there exists an integer $t > 0$ such that f^t is a morphism over C, that is, f^t satisfies $\pi \circ f^t = \pi$. In particular, f admits no Zariski dense orbit.*

Proof. By Lemma 5.4, we may assume that f induces a surjective endomorphism $f_C : C \to C$ with $\pi \circ f = f_C \circ \pi$. Since C is of general type, f_C is an automorphism of finite order and the assertion follows. □

Remark 6.14. One can also show that any surjective endomorphism over a curve of genus at least two admits no dense orbit by using the Mordell conjecture (Faltings's theorem).

7. Hyperelliptic surfaces

Theorem 7.1. *Let X be a hyperelliptic surface and $f\colon X \to X$ a noninvertible endomorphism on X. Then Conjecture 1.1 holds for f.*

Proof. Let $\pi: X \to E$ be the Albanese map of X. By the universality of π, there is a morphism $g\colon E \to E$ satisfying $\pi \circ f = g \circ \pi$. It is well-known that E is a genus one curve, π is a surjective morphism with connected fibers, and there is an étale cover $\phi\colon E' \to E$ such that $X' = X \times_E E' \cong F \times E'$, where F is a genus one curve [Bădescu 2001, Chapter 10]. In particular, X' is an abelian surface. By Lemma 6.3, taking a further étale base change, we may assume that there is an endomorphism $h\colon E' \to E'$ such that $\phi \circ h = g \circ \phi$. Let $\pi'\colon X' \to E'$ and $\psi: X' \to X$ be the induced morphisms. Then, by the universality of fiber products, there is a morphism $f'\colon X' \to X'$ satisfying $\pi' \circ f' = \pi' \circ h$ and $\psi \circ f' = f \circ \psi$. Applying Lemma 3.2, it is enough to prove Conjecture 1.1 for the endomorphism f'. Since X' is an abelian variety, this holds by [Kawaguchi and Silverman 2016a, Corollary 31] and [Silverman 2017, Theorem 2]. □

8. Surfaces with $\kappa(X) = 1$

Let $f\colon X \to X$ be a noninvertible endomorphism on a surface X with $\kappa(X) = 1$. In this section we shall prove that f does not admit any Zariski dense forward f-orbit. Although this result is a special case of [Nakayama and Zhang 2009, Theorem A] (see Remark 1.2), we will give a simpler proof of it.

By Lemma 4.2, X is minimal and f is étale. Since $\deg(f) \geq 2$, we have $\chi(X, \mathcal{O}_X) = 0$.

Let $\phi = \phi_{|mK_X|}\colon X \to \mathbb{P}^N = \mathbb{P}H^0(X, mK_X)$ be the Iitaka fibration of X and set $C_0 = \phi(X)$. Since f is étale, it induces an automorphism $g\colon \mathbb{P}^N \to \mathbb{P}^N$ such that $\phi \circ f = g \circ \phi$ [Fujimoto and Nakayama 2008, Lemma 3.1]. The restriction of g to C_0 gives an automorphism $f_{C_0}\colon C_0 \to C_0$ such that $\phi \circ f = f_{C_0} \circ \phi$. Take the normalization $\nu\colon C \to C_0$ of C_0. Then ϕ factors as $X \xrightarrow{\pi} C \xrightarrow{\nu} C_0$ and π is an elliptic fibration. Moreover, f_{C_0} lifts to an automorphism $f_C\colon C \to C$ such that $\pi \circ f = f_C \circ \pi$.

So we obtain an elliptic fibration $\pi: X \to C$ and an automorphism f_C on C such that $\pi \circ f = f_C \circ \pi$. In this situation, the following holds.

Theorem 8.1. *Let X be a surface with $\kappa(X) = 1$, $\pi\colon X \to C$ an elliptic fibration, $f\colon X \to X$ a noninvertible endomorphism, and $f_C\colon C \to C$ an automorphism such that $\pi \circ f = f_C \circ \pi$. Then $f_C^t = \mathrm{id}_C$ for a positive integer t.*

Proof. Let $\{P_1, \dots, P_r\}$ be the points over which the fibers of π are multiple fibers (possibly $r = 0$, i.e., π does not have any multiple fibers). We denote by m_i denotes the multiplicity of the fiber $\pi^* P_i$ for every i. Then we have the canonical bundle formula:

$$K_X = \pi^*(K_C + L) + \sum_{i=1}^{r} \frac{m_i - 1}{m_i} \pi^* P_i,$$

where L is a divisor on C such that $\deg(L) = \chi(X, \mathcal{O}_X)$. Then $\deg(L) = 0$ because f is étale and $\deg(f) \geq 2$ (see Lemma 4.2). Since $\kappa(X) = 1$, the divisor $K_C + L + \sum_{i=1}^{r}(m_i - 1)/m_i\, P_i$ must have positive degree. So we have

$$2(g(C)-1)+\sum_{i=1}^{r}\frac{m_i-1}{m_i}>0. \tag{$**$}$$

For any i, set $Q_i=f_C^{-1}(P_i)$. Then $\pi^*Q_i=\pi^*f_C^*P_i=f^*\pi^*P_i$ is a multiple fiber. So $(f_C)|_{\{P_1,\dots,P_r\}}$ is a permutation of $\{P_1,\dots,P_r\}$ since f_C is an automorphism.

We divide the proof into three cases according to the genus $g(C)$ of C:

(1) $g(C)\geq 2$; then the automorphism group of C is finite. So $f_C^t=\mathrm{id}_C$ for a positive integer t.

(2) $g(C)=1$; by $(**)$, it follows that $r\geq 1$. For a suitable t, all P_i are fixed points of f_C^t. We put the algebraic group structure on C such that P_1 is the identity element. Then f_C^t is a group automorphism on C. So $f_C^{ts}=\mathrm{id}_C$ for a suitable s since the group of group automorphisms on C is finite.

(3) $g(C)=0$; again by $(**)$, it follows that $r\geq 3$. For a suitable t, all P_i are fixed points of f_C^t. Then f_C^t fixes at least three points, which implies that f_C^t is in fact the identity map. □

Immediately we obtain the following corollary.

Corollary 8.2. *Let $f\colon X\to X$ be a noninvertible endomorphism on a surface X with $\kappa(X)=1$. Then there does not exist any Zariski dense f-orbit.*

Therefore Conjecture 1.1 trivially holds for noninvertible endomorphisms on surfaces of Kodaira dimension 1.

9. Existence of a rational point P satisfying $\alpha_f(P)=\delta_f$

In this section, we prove Theorems 1.6 and 1.7. Theorem 1.6 follows from the following lemma. A subset $\Sigma\subset V(\bar{k})$ is called a *set of bounded height* if for some (or, equivalently, any) ample divisor A on V, the height function h_A associated with A is a bounded function on Σ.

Lemma 9.1. *Let X be a smooth projective variety and $f\colon X\to X$ a surjective endomorphism with $\delta_f>1$. Let $D\not\equiv 0$ be a nef $\mathbb{R}$-divisor such that $f^*D\equiv\delta_f D$. Let $V\subset X$ be a closed subvariety of positive dimension such that $(D^{\dim V}\cdot V)>0$. Then there exists a nonempty open subset $U\subset V$ and a set $\Sigma\subset U(\bar{k})$ of bounded height such that for every $P\in U(\bar{k})\setminus\Sigma$ we have $\alpha_f(P)=\delta_f$.*

Remark 9.2. By a Perron–Frobenius type result of [Birkhoff 1967, Theorem], there is a nef $\mathbb{R}$-divisor $D\not\equiv 0$ satisfying the condition $f^*D\equiv\delta_f D$ since f^* preserves the nef cone.

Proof. Fix a height function h_D associated with D. For every $P\in X(\bar{k})$, the following limit exists [Kawaguchi and Silverman 2016b, Theorem 5]:

$$\hat{h}(P)=\lim_{n\to\infty}\frac{h_D(f^n(P))}{\delta_f^n}.$$

The function $\hat{h}$ has the following properties [Kawaguchi and Silverman 2016b, Theorem 5]:

(i) $\hat{h} = h_D + O(\sqrt{h_H})$ where H is any ample divisor on X and $h_H \geq 1$ is a height function associated with H.

(ii) If $\hat{h}(P) > 0$, then $\alpha_f(P) = \delta_f$.

Since $(D^{\dim V} \cdot V) > 0$, we have $(D|_V{}^{\dim V}) > 0$ and $D|_V$ is big. Thus we can write $D|_V \sim_{\mathbb{R}} A + E$ with an ample $\mathbb{R}$-divisor A and an effective $\mathbb{R}$-divisor E on V. Therefore we have

$$\hat{h}|_{V(\bar{k})} = h_A + h_E + O(\sqrt{h_A})$$

where h_A and h_E are height functions associated with A and E and h_A is taken to be $h_A \geq 1$. In particular, there exists a positive real number $B > 0$ such that $h_A + h_E - \hat{h}|_{V(\bar{k})} \leq B\sqrt{h_A}$. Then we have the following inclusions:

$$\begin{aligned}\{P \in V(\bar{k}) \mid \hat{h}(P) \leq 0\} &\subset \{P \in V(\bar{k}) \mid h_A(P) + h_E(P) \leq B\sqrt{h_A(P)}\}\\ &\subset \operatorname{Supp} E \cup \{P \in V(\bar{k}) \mid h_A(P) \leq B\sqrt{h_A(P)}\}\\ &= \operatorname{Supp} E \cup \{P \in V(\bar{k}) \mid h_A(P) \leq B^2\}.\end{aligned}$$

Hence we can take $U = V \setminus \operatorname{Supp} E$ and $\Sigma = \{P \in U(\bar{k}) \mid \hat{h}(P) \leq 0\}$. □

Corollary 9.3. *Let X be a smooth projective variety of dimension N and $f : X \to X$ a surjective endomorphism. Let C be a irreducible curve which is a complete intersection of ample effective divisors $H_1, \ldots, H_{N-1}$ on X. Then for infinitely many points P on C, we have $\alpha_f(P) = \delta_f$.*

Proof. We may assume $\delta_f > 1$. Let D be as in Lemma 9.1. Then $(D \cdot C) = (D \cdot H_1 \cdots H_{N-1}) > 0$ [Kawaguchi and Silverman 2016b, Lemma 20]. Since $C(\bar{k})$ is not a set of bounded height, the assertion follows from Lemma 9.1. □

To prove Theorem 1.7, we need the following theorem which is a corollary of the dynamical Mordell–Lang conjecture for étale finite morphisms.

Theorem 9.4 (Bell, Ghioca and Tucker [2010, Corollary 1.4]). *Let $f : X \to X$ be an étale finite morphism of smooth projective variety X. Let $P \in X(\bar{k})$. If the orbit $\mathcal{O}_f(P)$ is Zariski dense in X, then any proper closed subvariety of X intersects $\mathcal{O}_f(P)$ in at most finitely many points.*

Proof of Theorem 1.7. We may assume $\dim X \geq 2$. Since we are working over $\bar{k}$, we can write the set of all proper subvarieties of X as

$$\{V_i \subsetneq X \mid i = 0, 1, 2, \ldots\}.$$

By Corollary 9.3, we can take a point $P_0 \in X \setminus V_0$ such that $\alpha_f(P) = \delta_f$. Assume we can construct $P_0, \ldots, P_n$ satisfying the following conditions:

(1) $\alpha_f(P_i) = \delta_f$ for $i = 0, \ldots, n$.

(2) $\mathcal{O}_f(P_i) \cap \mathcal{O}_f(P_j) = \varnothing$ for $i \neq j$.

(3) $P_i \notin V_i$ for $i = 0, \ldots, n$.

Now, take a complete intersection curve $C \subset X$ satisfying the following conditions:

- For $i = 0, \ldots, n$, $C \not\subset \mathcal{O}_f(P_i)$ if $\overline{\mathcal{O}_f(P_i)} \neq X$.
- For $i = 0, \ldots, n$, $C \not\subset \mathcal{O}_{f^{-1}}(P_i)$ if $\overline{\mathcal{O}_{f^{-1}}(P_i)} \neq X$.
- $C \not\subset V_{n+1}$.

By Theorem 9.4, if $\mathcal{O}_{f^{\pm}}(P_i)$ is Zariski dense in X, then $\mathcal{O}_{f^{\pm}}(P_i) \cap C$ is a finite set. By Corollary 9.3, there exists a point

$$P_{n+1} \in C \setminus \left(\bigcup_{0 \le i \le n} \mathcal{O}_f(P_i) \cup \bigcup_{0 \le i \le n} \mathcal{O}_{f^{-1}}(P_i) \cup V_{n+1} \right)$$

such that $\alpha_f(P_{n+1}) = \delta_f$. Then $P_0, \ldots, P_{n+1}$ satisfy the same conditions. Therefore we get a subset $S = \{P_i \mid i = 0, 1, 2, \ldots\}$ of X which satisfies the desired conditions. □

Acknowledgements

The authors would like to thank Professors Tetsushi Ito, Osamu Fujino, and Tomohide Terasoma for helpful advice. They would also like to thank Takeru Fukuoka and Hiroyasu Miyazaki for answering their questions. The first author is supported by the Program for Leading Graduate Schools, MEXT, Japan. The third author is supported by JSPS KAKENHI Grant Number JP17J01912.

References

[Amerik 2003] E. Amerik, "On endomorphisms of projective bundles", *Manuscripta Math.* **111**:1 (2003), 17–28. MR Zbl

[Bell et al. 2010] J. P. Bell, D. Ghioca, and T. J. Tucker, "The dynamical Mordell–Lang problem for étale maps", *Amer. J. Math.* **132**:6 (2010), 1655–1675. MR Zbl

[Birkhoff 1967] G. Birkhoff, "Linear transformations with invariant cones", *Amer. Math. Monthly* **74** (1967), 274–276. MR Zbl

[Bădescu 2001] L. Bădescu, *Algebraic surfaces*, Springer, 2001. MR Zbl

[Call and Silverman 1993] G. S. Call and J. H. Silverman, "Canonical heights on varieties with morphisms", *Compositio Math.* **89**:2 (1993), 163–205. MR Zbl

[Dang 2017] N.-B. Dang, "Degrees of iterates of rational maps on normal projective varieties", preprint, 2017. arXiv

[Diller and Favre 2001] J. Diller and C. Favre, "Dynamics of bimeromorphic maps of surfaces", *Amer. J. Math.* **123**:6 (2001), 1135–1169. MR Zbl

[Dinh and Sibony 2004] T.-C. Dinh and N. Sibony, "Regularization of currents and entropy", *Ann. Sci. École Norm. Sup.* (4) **37**:6 (2004), 959–971. MR Zbl

[Dinh and Sibony 2005] T.-C. Dinh and N. Sibony, "Une borne supérieure pour l'entropie topologique d'une application rationnelle", *Ann. of Math.* (2) **161**:3 (2005), 1637–1644. MR Zbl

[Fujimoto 2002] Y. Fujimoto, "Endomorphisms of smooth projective 3-folds with non-negative Kodaira dimension", *Publ. Res. Inst. Math. Sci.* **38**:1 (2002), 33–92. MR Zbl

[Fujimoto and Nakayama 2005] Y. Fujimoto and N. Nakayama, "Compact complex surfaces admitting non-trivial surjective endomorphisms", *Tohoku Math. J.* (2) **57**:3 (2005), 395–426. MR Zbl

[Fujimoto and Nakayama 2008] Y. Fujimoto and N. Nakayama, "Complex projective manifolds which admit non-isomorphic surjective endomorphisms", pp. 51–79 in *Higher dimensional algebraic varieties and vector bundles*, edited by S. Mukai, RIMS Kôkyûroku Bessatsu **B9**, RIMS, Kyoto, 2008. MR Zbl

[Guedj 2005] V. Guedj, "Ergodic properties of rational mappings with large topological degree", *Ann. of Math.* (2) **161**:3 (2005), 1589–1607. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, 1977. MR Zbl

[Hindry and Silverman 2000] M. Hindry and J. H. Silverman, *Diophantine geometry*, Graduate Texts in Mathematics **201**, Springer, 2000. MR Zbl

[Homma 1992] Y. Homma, "On finite morphisms of rational ruled surfaces", *Math. Nachr.* **158** (1992), 263–281. MR Zbl

[Homma 1999] Y. Homma, "On finite morphisms of ruled surfaces", *Geom. Dedicata* **78**:3 (1999), 259–269. MR Zbl

[Iitaka 1982] S. Iitaka, *Algebraic geometry*, Graduate Texts in Mathematics **76**, Springer, 1982. MR Zbl

[Kawaguchi 2008] S. Kawaguchi, "Projective surface automorphisms of positive topological entropy from an arithmetic viewpoint", *Amer. J. Math.* **130**:1 (2008), 159–186. MR Zbl

[Kawaguchi and Silverman 2014] S. Kawaguchi and J. H. Silverman, "Examples of dynamical degree equals arithmetic degree", *Michigan Math. J.* **63**:1 (2014), 41–63. MR Zbl

[Kawaguchi and Silverman 2016a] S. Kawaguchi and J. H. Silverman, "Dynamical canonical heights for Jordan blocks, arithmetic degrees of orbits, and nef canonical heights on abelian varieties", *Trans. Amer. Math. Soc.* **368**:7 (2016), 5009–5035. MR Zbl

[Kawaguchi and Silverman 2016b] S. Kawaguchi and J. H. Silverman, "On the dynamical and arithmetic degrees of rational self-maps of algebraic varieties", *J. Reine Angew. Math.* **713** (2016), 21–48. MR Zbl

[Lin 2018] J.-L. Lin, "On the arithmetic dynamics of monomial maps", *Ergodic Theory Dynam. Systems* (Published online March 2018).

[Matsumura 1963] H. Matsumura, "On algebraic groups of birational transformations", *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* (8) **34** (1963), 151–155. MR Zbl

[Matsuzawa 2016] Y. Matsuzawa, "On upper bounds of arithmetic degrees", preprint, 2016. arXiv

[Mukai 2003] S. Mukai, *An introduction to invariants and moduli*, Cambridge Studies in Advanced Mathematics **81**, Cambridge University Press, 2003. MR Zbl

[Nakayama 2002] N. Nakayama, "Ruled surfaces with non-trivial surjective endomorphisms", *Kyushu J. Math.* **56**:2 (2002), 433–446. MR Zbl

[Nakayama and Zhang 2009] N. Nakayama and D.-Q. Zhang, "Building blocks of étale endomorphisms of complex projective manifolds", *Proc. Lond. Math. Soc.* (3) **99**:3 (2009), 725–756. MR Zbl

[Sano 2016] K. Sano, "Dynamical degree and arithmetic degree of endomorphisms on product varieties", preprint, 2016. To appear in *Tohoku Math. J.* arXiv

[Serre 1997] J.-P. Serre, *Lectures on the Mordell–Weil theorem*, 3rd ed., Aspects of Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1997. MR Zbl

[Silverman 2014] J. H. Silverman, "Dynamical degree, arithmetic entropy, and canonical heights for dominant rational self-maps of projective space", *Ergodic Theory Dynam. Systems* **34**:2 (2014), 647–678. MR Zbl

[Silverman 2017] J. H. Silverman, "Arithmetic and dynamical degrees on abelian varieties", *J. Théor. Nombres Bordeaux* **29**:1 (2017), 151–167. MR Zbl

[Truong 2015] T. T. Truong, "(Relative) dynamical degrees of rational maps over an algebraic closed field", preprint, 2015. arXiv

Communicated by Hélène Esnault
Received 2017-03-20 Revised 2018-04-05 Accepted 2018-06-20

myohsuke@ms.u-tokyo.ac.jp *Graduate school of Mathematical Sciences, University of Tokyo, Komaba, Tokyo, Japan*

ksano@math.kyoto-u.ac.jp *Department of Mathematics, Faculty of Science, Kyoto University, Japan*

tshibata@math.kyoto-u.ac.jp *Department of Mathematics, Faculty of Science, Kyoto University, Japan*

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1659

Big Cohen–Macaulay algebras and the vanishing conjecture for maps of Tor in mixed characteristic

Raymond Heitmann and Linquan Ma

We prove a version of weakly functorial big Cohen–Macaulay algebras that suffices to establish Hochster and Huneke's vanishing conjecture for maps of Tor in mixed characteristic. As a corollary, we prove an analog of Boutot's theorem that direct summands of regular rings are pseudorational in mixed characteristic. Our proof uses perfectoid spaces and is inspired by the recent breakthroughs on the direct summand conjecture by André and Bhatt.

1. Introduction and preliminaries

In a recent breakthrough, Y. André [2018a] settled Hochster's direct summand conjecture which dates back to 1969.

Theorem 1.1 (André). *Let $A \to R$ be a finite extension of Noetherian rings. If A is regular, then the map is split as a map of A-modules.*

This was previously only known for rings containing a field [Hochster 1975b] and for rings of dimension less than or equal to three [Heitmann 2002]; what is new and striking is the general mixed characteristic case. A simplified and shorter proof of Theorem 1.1 was later found by Bhatt [2018]. But André's argument [2018a] also proved the stronger conjecture that balanced big Cohen–Macaulay algebras exist in mixed characteristic.[1] Recall that B is called a *balanced big Cohen–Macaulay algebra* for the local ring $(R, \mathfrak{m})$ if $\mathfrak{m}B \neq B$ and every system of parameters for R is a regular sequence on B. It is a conjecture of Hochster [1975a; 1975b] that such algebras exist in general and he proved this for rings that contain a field. André's solution in mixed characteristic depends on his deep result in [André 2018b] that gives a generalization of the almost purity theorem: the perfectoid Abhyankar lemma.

The purpose of this paper is to prove that weakly functorial balanced big Cohen–Macaulay algebras exist for *certain* surjective ring homomorphisms in mixed characteristic, a result that has many applications.

MSC2010: primary 13D22; secondary 13H05.
Keywords: big Cohen–Macaulay algebras, direct summand conjecture, perfectoid space, vanishing conjecture for maps of Tor, pseudorational singularities.

[1] André [2018a, Théorème 0.7.1] stated the existence of big Cohen–Macaulay algebras for complete local domains, but the general case follows by killing a minimal prime and taking the completion.

Theorem 3.1. *Let $(R, \mathfrak{m}, k)$ be a complete local domain with k algebraically closed, and let $Q \subseteq R$ be a height one prime ideal. Suppose both R and R/Q have mixed characteristic. Then there exists a commutative diagram*:

$$\begin{array}{ccc} R & \longrightarrow & R/Q \\ \downarrow & & \downarrow \\ B & \longrightarrow & C \end{array}$$

where B, C are balanced big Cohen–Macaulay algebras for R and R/Q respectively.

Our method of proving *avoids* the perfectoid Abhyankar lemma in [André 2018b] and thus is much shorter than André's argument. More importantly, the weakly functorial property we prove is *new*.[2] We should note that, in equal characteristic, the existence of weakly functorial balanced big Cohen–Macaulay algebras was known in general [Hochster and Huneke 1995]. Nonetheless, the version we prove is strong enough to settle Hochster and Huneke's [1995] vanishing conjecture for maps of Tor in mixed characteristic.

Theorem 4.1. *Let $A \to R \to S$ be maps of Noetherian rings such that $A \to S$ is a local homomorphism of mixed characteristic regular local rings and R is a module-finite torsion-free extension of A. Then for all A-modules M, the map $\operatorname{Tor}_i^A(M, R) \to \operatorname{Tor}_i^A(M, S)$ vanishes for all $i \geq 1$.*

As a consequence of Theorem 4.1, we prove the following, which is the mixed characteristic analog of Boutot's theorem [1987].[3]

Corollary 4.3. *If $R \to S$ is a ring extension such that S is regular and the map is split as a map of R-modules, then R is pseudorational* (*in particular Cohen–Macaulay*).

It is well known that Theorem 4.1 implies Theorem 1.1 (for example, see [Ranganathan 2000] or [Ma 2018, Remark 4.6]). Recently, Bhatt [2018] gave an alternative and shorter proof of Theorem 1.1: instead of using the perfectoid Abhyankar lemma, Bhatt established a quantitative form of Scholze's Hebbarkeitssatz (the Riemann extension theorem) for perfectoid spaces, and the same idea leads to a proof of a derived variant, i.e., the derived direct summand conjecture. We point out that Theorem 4.1 formally implies such derived variant by [Ma 2018, Remark 5.12] and hence we recover, and in fact generalize, Bhatt's result (see Remark 4.5). Furthermore, although the idea is inspired by [Bhatt 2018], our argument is independent of that work in exposition. We avoid the use of Scholze's Hebbarkeitssatz and the vanishing theorems of perfectoid spaces; instead we study the colon ideals of $A_\infty\langle p^n/g\rangle$ in Lemma 3.4.

Remark 1.2. We should point out that, to the best of our knowledge, Hochster and Huneke's vanishing conjecture for maps of Tor is still open if A and R have mixed characteristic but S has equal characteristic $p > 0$. This case also implies Theorem 1.1 by [Hochster and Huneke 1995, (4.4)]. However, the discussion above shows that the mixed characteristic case we proved (i.e., Theorem 4.1) is enough for almost all applications.

[2] In fact, our version of the existence of weakly functorial big Cohen–Macaulay algebras does not even seem to follow from the perfectoid Abhyankar lemma [André 2018b].

[3] This corollary can be also proved by combining [André 2018a, Remarque 4.2.1] and [Bhatt 2018, Theorem 1.2] (and an extra small argument), see Remark 4.4. However, to the best of our knowledge, the results of [André 2018a; 2018b; Bhatt 2018] are not enough to establish Theorem 4.1.

This paper is organized as follows. In Section 2 we demonstrate a weakly functorial construction of integral perfectoid algebras in Lemma 2.3. Then, in Section 3, we prove Theorem 3.1, and in Section 4, we prove Theorem 4.1 and Corollary 4.3.

Perfectoid algebras. We will freely use the language of perfectoid spaces [Scholze 2012] and almost mathematics [Gabber and Ramero 2003]. In this paper we will always work in the following situation: for a perfect field k of characteristic $p > 0$, we let $W(k)$ be the ring of Witt vectors with coefficients in k. Let K° be the p-adic completion of $W(k)[p^{1/p^\infty}]$ and $K = K^\circ[1/p]$. Then K is a *perfectoid field* in the sense of [Scholze 2012] with $K^\circ \subseteq K$ its ring of integers.

A *perfectoid K-algebra* is a Banach K-algebra R such that the set of power-bounded elements $R^\circ \subseteq R$ is bounded and the Frobenius is surjective on R°/p. A K°-algebra S is called *integral perfectoid* if it is p-adically complete, p-torsion free, satisfies $S = S_*$[4] and the Frobenius induces an isomorphism $S/p^{1/p} \to S/p$. These two categories are equivalent to each other [Scholze 2012, Theorem 5.2] via the functors $R \to R^\circ$ and $S \to S[1/p]$.

Unless otherwise stated, almost mathematics in this paper will always be measured with respect to the ideal (p^{1/p^∞}) in K°.

Partial algebra modifications. We briefly recall Hochster's partial algebra modifications that play a crucial rule in the construction of balanced big Cohen–Macaulay algebras. Our definition and usage of these modifications is basically the same as that in [Hochster 2002, Sections 3 and 4].

Let $(R, \mathfrak{m})$ be a local ring and let M be an R-module. We define a *partial algebra modification* of M with respect to a system of parameters $x_1, \dots, x_d$ of R to be a map $M \to M'$ obtained as follows: for some integer $s \geq 0$ and relation $x_{s+1}u_{s+1} = \sum_{j=1}^{s} x_j u_j$, where $u_j \in M$, choose indeterminates $X_1, \dots, X_s$ and an integer $N \geq 1$, let $F = u_{s+1} - \sum_{j=1}^{s} x_j X_j$ and let

$$M' = M[X_1, \dots, X_s]_{\leq N}/F \cdot R[X_1, \dots, X_s]_{\leq N-1},$$

where $M[X_1, \dots, X_s] = M \otimes_R R[X_1, \dots, X_s]$ and thus $M[X_1, \dots, X_s]_{\leq N}$ refers to polynomials of degree at most N (with coefficients in M). The definition of M' makes sense since F has degree one in X_j. It is readily seen that in M', the relation $x_{s+1}u_{s+1} = \sum_{j=1}^{s} x_j u_j$ is trivialized in the sense that u_{s+1} is contained in $(x_1, \dots, x_s)M'$ by construction. We shall refer to the integer N as the *degree bound* of the partial algebra modification. We can then recursively define a sequence of partial algebra modifications of an R-module M.

Now given a local map of local rings $(R, \mathfrak{m}) \to (S, \mathfrak{n})$ we can define a *double sequence of partial algebra modifications* of an R-module M with respect to $R \to S$, a system of parameters $x_1, \dots, x_d$ of R and a system of parameters $y_1, \dots, y_{d'}$ of S as follows: we first form a sequence of partial algebra modifications of M over R with respect to $x_1, \dots, x_d$, say $M = M_0, M_1, \dots, M_r$, and then we form a sequence of partial algebra modifications $N_0 = S \otimes_R M_r, N_1, \dots, N_s$ of N_0 over S with respect to $y_1, \dots, y_{d'}$. When M is an R-algebra, we call this double sequence *bad* if the image of $1 \in M$ in N_s is in $\mathfrak{n}N_s$.

[4] $S_* = \{x \in S[1/p] \mid p^{1/p^k} \cdot x \in S \text{ for all } k\}$. Hence S is almost isomorphic to S_* with respect to (p^{1/p^∞}); thus in practice we will often ignore this distinction since one can always pass to S_* without affecting the issue.

The following was essentially taken from [Hochster 2002, Theorem 4.2], and is one of the main ingredients in our construction.

Theorem 1.3. *Let $(R, \mathfrak{m}) \to (S, \mathfrak{n})$ be a local homomorphism of local rings. Then there exists a commutative diagram*

$$\begin{array}{ccc} R & \longrightarrow & S \\ \downarrow & & \downarrow \\ B & \longrightarrow & C \end{array}$$

such that B is a balanced big Cohen–Macaulay algebra for R and C is a balanced big Cohen–Macaulay algebra for S if and only if there is no bad double sequence of partial algebra modifications of R over $R \to S$ with respect to $x_1, \dots, x_d$ of R and $y_1, \dots, y_{d'}$ of S.

This theorem is actually a bit stronger than [Hochster 2002, Theorem 4.2]. Whereas Hochster allows the system of parameters to vary throughout the double sequence, we fix a system of parameters of R and S. But the idea of the proof is the same: one first constructs B' as a direct limit of finite sequences of modifications of R and then constructs C' as a direct limit of finite sequences of modifications of $S \otimes_R B$ over S. It is readily seen that $x_1, \dots, x_d$ and $y_1, \dots, y_{d'}$ are improper-regular sequences on B' and C' respectively. To guarantee that $\mathfrak{m}B' \neq B'$ and $\mathfrak{n}C' \neq C'$ one needs precisely that there is no bad double sequence of partial algebra modifications over $R \to S$. Now B' and C' are not balanced, but that problem is easily remedied. We invoke [Bruns and Herzog 1993, Corollary 8.5.3] to note that $B' \to C'$ induces $B = \widehat{B'}^{\mathfrak{m}} \to C = \widehat{C'}^{\mathfrak{n}}$, a map of balanced big Cohen–Macaulay algebras of $R \to S$.

2. Weakly functorial construction of integral perfectoid algebras

Notation. Throughout this section, $(A, \mathfrak{m}, k)$ will always be a complete and unramified regular local ring of mixed characteristic with k perfect, i.e., $A \cong W(k)[[x_1, \dots, x_{d-1}]]$, where $W(k)$ is the ring of Witt vectors with coefficients in k. Let K° be the p-adic completion of $W(k)[p^{1/p^\infty}]$ and $K = K^\circ[1/p]$. Let $A_{\infty,0}$ be the p-adic completion of $A[p^{1/p^\infty}, x_1^{1/p^\infty}, \dots, x_{d-1}^{1/p^\infty}]$, which is an integral perfectoid K°-algebra.

For any nonzero element $g \in A$, we let $A_{\infty,0} \to A_\infty$ be André's construction of integral perfectoid K°-algebras (for example see [Bhatt 2018, Theorem 2.3]): A_∞ is almost faithfully flat over $A_{\infty,0}$ modulo p such that g admits a compatible system of p^k-th roots in A_∞. We will denote by $A_\infty\langle p^n/g\rangle$ the integral perfectoid K°-algebra which is the ring of bounded functions on the rational subset $\{x \in X \mid |p^n| \leq |g(x)|\}$, where $X = \mathrm{Spa}(A_\infty[1/p], A_\infty)$ is the perfectoid space associated to A_∞. Since g admits a compatible system of p^k-th roots in A_∞, $A_\infty\langle p^n/g\rangle$ can be described almost explicitly as the p-adic completion of $A_\infty[(p^n/g)^{1/p^\infty}]$ [Scholze 2012, Lemma 6.4].

We begin by observing the following:

Lemma 2.1. *Suppose $g \neq 0$ in A/x_1A. Then we have a natural map $A_\infty \to (A/x_1A)_\infty$ sending g^{1/p^k} to $\bar{g}^{1/p^k}$.*

Proof. We first note that there are natural maps

$$A_{\infty,0} \to (A/x_1A)_{\infty,0} \to (A/x_1A)_\infty,$$

where the first map is simply obtained by killing x_1^{1/p^∞}. Thus we have a map

$$A_{\infty,0}\langle T^{1/p^\infty}\rangle \to (A/x_1A)_\infty$$

of integral perfectoid K°-algebras sending T^{1/p^k} to $\bar{g}^{1/p^k}$. Since A_∞ is the ring of functions on the Zariski closed subset of $Y = \mathrm{Spa}(A_{\infty,0}\langle T^{1/p^\infty}\rangle[1/p], A_{\infty,0}\langle T^{1/p^\infty}\rangle)$ defined by $T - g$, the map $A_{\infty,0}\langle T^{1/p^\infty}\rangle \to (A/x_1A)_\infty$ induces a map $A_\infty \to (A/x_1A)_\infty$ sending g^{1/p^k} to $\bar{g}^{1/p^k}$. □

Lemma 2.2. *Let $(R, \mathfrak{m}, k)$ be a complete normal local domain with k perfect, and let $Q \subseteq R$ be a height one prime ideal. Suppose both R and R/Q have mixed characteristic. Then we can find a complete and unramified regular local ring $A \cong W(k)[[x_1, \ldots, x_{d-1}]]$ with $A \to R$ a module-finite extension such that*

(1) $Q \cap A = (x_1)$;

(2) $A_{(x_1)} \to R_Q$ *is essentially étale.*

Proof. Let $\{P_i\}$ be all the minimal primes of (p); they all have height one. Since R/Q has mixed characteristic, $p \notin Q$. Thus Q is not contained in any of the P_i. By prime avoidance we can choose $x \in Q$ that is not in $\left(\bigcup_i P_i\right) \cup Q^{(2)}$. Thus the image of x in R_Q generates QR_Q since R is normal, and p, x is part of a system of parameters of R.

Cohen's structure theorem implies the existence of a complete and unramified regular local ring $A \cong W(k)[[x_1, \ldots, x_{d-1}]]$ and a module-finite extension $A \to R$ such that the image of x_1 in R is x. It is clear that $Q \cap A = (x_1)$ because $Q \cap A$ is a height one prime of A that contains (x_1), so it must be (x_1). To see $A_{(x_1)} \to R_Q$ is essentially étale, note that the image of x_1, x, generates the maximal ideal QR_Q of R_Q and the extension of residue fields $A_{(x_1)}/(x_1)A_{(x_1)} \to R_Q/QR_Q$ is finite separable since both fields have characteristic 0 (p is inverted when we localize). Thus $A_{(x_1)} \to R_Q$ is unramified. But it is clearly flat because R_Q is x_1-torsion free. Therefore $A_{(x_1)} \to R_Q$ is essentially étale. □

The following is the main result of this section. It is crucial in proving the version of weakly functorial balanced big Cohen–Macaulay algebras that we need.

Lemma 2.3. *Let $(R, \mathfrak{m}, k)$ be a complete normal local domain with k perfect, and let $Q \subseteq R$ be a height one prime ideal. Suppose both R and R/Q have mixed characteristic. We pick $A \cong W(k)[[x_1, \ldots, x_{d-1}]]$ such that $A \to R$ is a module-finite extension satisfying the conclusion of Lemma 2.2. Then there exists an element $g \in A$, whose image is nonzero in A/x_1A, such that $A_g \to R_g$ and $(A/x_1A)_g \to (R/Q)_g$ are both finite étale. Furthermore, for every $n > 0$, we have a commutative diagram*:

$$\begin{tikzcd} R \arrow[r] \arrow[d] & R/Q \arrow[d] \\ R_{\infty,n} \arrow[r] & (R/Q)_{\infty,n} \end{tikzcd}$$

where $R_{\infty,n}$ (resp. $(R/Q)_{\infty,n}$) is an integral perfectoid K°-algebra that is almost finite étale over $A_\infty\langle p^n/g\rangle$ (resp. $(A/x_1A)_\infty\langle p^n/g\rangle$).

Proof. Let $g \in A$ be the discriminant of the map $A \to R$; i.e., it defines the locus of $\operatorname{Spec} A$ such that the map $A \to R$ is not essentially étale when localizing. Since $A_{(x_1)} \to R_Q$ is essentially étale, g is nonzero in A/x_1A. Since x_1 generates Q when localizing at Q and we know that $A_g \to R_g$ and hence $(A/x_1A)_g \to (R/x_1R)_g$ are finite étale, replacing g by a multiple we have $A_g \to R_g$ and $(A/x_1A)_g \to (R/Q)_g$ are both finite étale. By Lemma 2.2 we have a commutative diagram:

$$\begin{tikzcd} A \arrow[r] \arrow[d] & R \arrow[d] \\ A/x_1A \arrow[r] & R/Q. \end{tikzcd}$$

By Lemma 2.1 we also have a commutative diagram:

$$\begin{tikzcd} A \arrow[r] \arrow[d] & A_\infty \arrow[r] \arrow[d] & A_\infty\langle \frac{p^n}{g}\rangle \arrow[d] \\ A/x_1A \arrow[r] & (A/x_1A)_\infty \arrow[r] & (A/x_1A)_\infty\langle \frac{p^n}{g}\rangle. \end{tikzcd}$$

Tensoring over A we get a natural commutative diagram:

$$\begin{tikzcd} R \arrow[r] \arrow[d] & R\otimes A_\infty\langle \frac{p^n}{g}\rangle \arrow[d] \\ R/Q \arrow[r] & (R/Q)\otimes (A/x_1A)_\infty\langle \frac{p^n}{g}\rangle. \end{tikzcd}$$

Since $A_g \to R_g$ and $(A/x_1A)_g \to (R/Q)_g$ are both finite étale and g divides p^n in $A_\infty\langle p^n/g\rangle$ and $(A/x_1A)_\infty\langle p^n/g\rangle$, we know that $(R\otimes A_\infty\langle p^n/g\rangle)[1/p]$ and $((R/Q)\otimes (A/x_1A)_\infty\langle p^n/g\rangle)[1/p]$ are finite étale over $(A_\infty\langle p^n/g\rangle)[1/p]$ and $((A/x_1A)_\infty\langle p^n/g\rangle)[1/p]$ respectively. Therefore

$$\Big(R\otimes A_\infty\Big\langle\frac{p^n}{g}\Big\rangle\Big)\Big[\frac{1}{p}\Big] \to \Big((R/Q)\otimes (A/x_1A)_\infty\Big\langle\frac{p^n}{g}\Big\rangle\Big)\Big[\frac{1}{p}\Big]$$

is a morphism of perfectoid K-algebras; thus it induces a map on the ring of power-bounded elements

$$R_{\infty,n} := \Big(R\otimes A_\infty\Big\langle\frac{p^n}{g}\Big\rangle\Big)\Big[\frac{1}{p}\Big]^\circ \to \Big(R/Q)_{\infty,n} := ((R/Q)\otimes (A/x_1A)_\infty\Big\langle\frac{p^n}{g}\Big\rangle\Big)\Big[\frac{1}{p}\Big]^\circ.$$

The almost purity theorem [Scholze 2012, Theorem 7.9] implies that $R_{\infty,n}$ and $(R/Q)_{\infty,n}$ are integral perfectoid K°-algebras that are almost finite étale over $A_\infty\langle p^n/g\rangle$ and $(A/x_1A)_\infty\langle p^n/g\rangle$ respectively. Therefore we have the desired commutative diagram:

$$\begin{tikzcd} R \arrow[r] \arrow[d] & R/Q \arrow[d] \\ R_{\infty,n} \arrow[r] & (R/Q)_{\infty,n} \end{tikzcd}$$

□

3. The main result

In this section we continue to use the notation from the beginning of Section 2. The main theorem we want to prove is the following:

Theorem 3.1. *Let $(R, \mathfrak{m}, k)$ be a complete local domain with k algebraically closed, and let $Q \subseteq R$ be a height one prime ideal. Suppose both R and R/Q have mixed characteristic. Then there exists a commutative diagram*:

$$\begin{array}{ccc} R & \longrightarrow & R/Q \\ \downarrow & & \downarrow \\ B & \longrightarrow & C \end{array}$$

where B, C are balanced big Cohen–Macaulay algebras for R and R/Q respectively.

To prove this we need several lemmas.

Lemma 3.2. *Let $A \cong W(k)[[x_1, \dots, x_{d-1}]]$ be a complete and unramified regular local ring with k perfect, and let $I = (p, y_1, \dots, y_s)$ be an ideal of A that contains p. Fix a nonzero element $g = p^m g_0 \in A$, where $p \nmid g_0$, and consider the extension $A \to A_\infty \to A_\infty\langle p^n/g\rangle$. Suppose $z \in IA_\infty\langle p^n/g\rangle \cap A_\infty$ for some $n > p^a + m$ (one should think that $n \gg p^a \gg 0$ here). Then we have $(pg)^{1/p^a} z \in IA_\infty$.*

Proof. Using the almost explicit description of $A_\infty\langle p^n/g\rangle$ [Scholze 2012, Lemma 6.4], we have

$$p^{1/p^t} z \in IA_\infty\Big[\widehat{\Big(\frac{p^n}{g}\Big)^{1/p^\infty}}\Big]$$

for some $t > a$. This implies that the image of $p^{1/p^t} z$ in $A_\infty[(p^n/g)^{1/p^\infty}]/p = A_\infty[\widehat{(p^n/g)^{1/p^\infty}}]/p$ is contained in the ideal $(y_1, \dots, y_s)$. Therefore we can write

$$p^{1/p^t} z = pf_0 + y_1 f_1 + \dots + y_s f_s,$$

where $f_0, f_1, \dots, f_s \in A_\infty[(p^n/g)^{1/p^\infty}]$. Then there exists integers k and h such that $f_0, f_1, \dots, f_s$ are elements in $A_\infty[(p^n/g)^{1/p^k}]$ of degree bounded by $p^k h$. Multiplying by g_0^h to clear all the denominators in f_i, one gets:

$$p^{1/p^t} g_0^h z \in (g_0^{h-(1/p^a)}, p^{(n-m)/p^a}) \cdot (p, y_1, \dots, y_s)A_\infty.$$

From this we know:

$$p^{1/p^t} g_0^h z = g_0^{h-(1/p^a)}(ph_0 + y_1 h_1 + \dots + y_s h_s) \quad \text{in } A_\infty/p^{(n-m)/p^a},$$

where $h_0, h_1, \dots, h_s \in A_\infty$. Rewriting this we have

$$g_0^{h-(1/p^a)}(p^{1/p^t} g_0^{1/p^a} z - ph_0 - y_1 h_1 - \dots - y_s h_s) = 0 \quad \text{in } A_\infty/p^{(n-m)/p^a}.$$

Since $p \nmid g_0$, g_0 is a nonzero divisor on A/p. This implies $g_0^{h-(1/p^a)}$ is an almost nonzero divisor on $A_\infty/p^{(n-m)/p^a}$ since $A \to A_{\infty,0}$ is faithfully flat and $A_{\infty,0} \to A_\infty$ is almost faithfully flat modulo p.

Hence $p^{1/p^t} g_0^{1/p^a} z - ph_0 - y_1h_1 - \cdots - y_sh_s$ is killed by (p^{1/p^∞}). In particular, since $t > a$, we know

$$(pg_0)^{1/p^a} z \in (p, y_1, \ldots, y_s) \quad \text{in } A_\infty/p^{(n-m)/p^a}.$$

Finally, since $n > p^a + m$ and g is a multiple of g_0, we have

$$(pg)^{1/p^a} z \in (p, y_1, \ldots, y_s)A_\infty.$$

This finishes the proof. □

Lemma 3.3. *Let $A \cong W(k)[\![x_1, \ldots, x_{d-1}]\!]$ be a complete and unramified regular local ring with k perfect. Fix a nonzero element $g = p^m g_0 \in A$ where $p \nmid g_0$ and $n > m$. Suppose $z \in A_\infty[(p^n/g)^{1/p^\infty}]$ and $p^D z \in A_\infty$ for some $D > 0$. Then $p^D z \in p^{D-(1/p^t)} A_\infty$ for all t.*

Proof. There exist $k \gg 0$ such that $z \in A_\infty[(p^n/g)^{1/p^k}]$. Choosing a high enough power of g_0 to clear denominators, we get $g_0^h z \in A_\infty$. So $g_0^h(p^D z) \in p^D A_\infty$. Since g_0 is a nonzerodivisor on A/p^D and A_∞/p^D is almost faithfully flat over A/p^D, $p^{1/p^t} p^D z \in p^D A_\infty$ for all t. Since A_∞ is p-torsion free, $p^D z \in p^{D-(1/p^t)} A_\infty$ for all t. □

Lemma 3.4. *Let $A \cong W(k)[\![x_1, \ldots, x_{d-1}]\!]$ be a complete and unramified regular local ring with k perfect. Fix a nonzero element $g = p^m g_0 \in A$ where $p \nmid g_0$, and consider the extension $A \to A_\infty \to A_\infty\langle p^n/g\rangle$ for every n. Suppose S is an almost finite projective $A_\infty\langle p^n/g\rangle$-algebra. If $p^a + m < n$, then we have $(p^{1/p^\infty})(pg)^{1/p^a}$ annihilates $(p, x_1, \ldots, x_s)S : x_{s+1}/(p, x_1, \ldots, x_s)S$ for all $s < d-1$.*

Proof. Suppose $y \in (p, x_1, \ldots, x_s)A_\infty\langle p^n/g\rangle : x_{s+1}$. Since y is an element of $A_\infty\langle p^n/g\rangle$, for every $t > 0$, $p^{1/p^t} y \in A_\infty[\widehat{(p^n/g)^{1/p^\infty}}]$ and $x_{s+1} p^{1/p^t} y \in (p, x_1, \ldots, x_s)A_\infty[\widehat{(p^n/g)^{1/p^\infty}}]$ by [Scholze 2012, Lemma 6.4]. Thus modulo p, $p^{1/p^t} y$ gives an element in $A_\infty[\widehat{(p^n/g)^{1/p^\infty}}]/p = A_\infty[(p^n/g)^{1/p^\infty}]/p$. We pick $z \in A_\infty[(p^n/g)^{1/p^\infty}]$ such that $z \equiv p^{1/p^t} y$ modulo $pA_\infty[\widehat{(p^n/g)^{1/p^\infty}}]$.

Now the image of $x_{s+1}z \in A_\infty[(p^n/g)^{1/p^\infty}]$ in $A_\infty[(p^n/g)^{1/p^\infty}]/p = A_\infty[\widehat{(p^n/g)^{1/p^\infty}}]/p$ is contained in the ideal $(x_1, \ldots, x_s)(A_\infty[\widehat{(p^n/g)^{1/p^\infty}}]/p)$. Therefore, we know

$$x_{s+1}z \in (p, x_1, \ldots, x_s)A_\infty\Big[\Big(\frac{p^n}{g}\Big)^{1/p^\infty}\Big]$$

and thus

$$z \in (p, x_1, \ldots, x_s)A_\infty\Big[\Big(\frac{p^n}{g}\Big)^{1/p^\infty}\Big] : x_{s+1}.$$

Next we write $z = u + (p^n/g)^{1/p^a} u'$ such that $g_0^{1/p^a} u \in A_\infty$, $u' \in A_\infty[(p^n/g)^{1/p^\infty}]$, and we also write $x_{s+1}z = v + (p^n/g)^{1/p^a} v'$ such that $g_0^{1/p^a} v \in (p, x_1, \ldots, x_s)A_\infty$, $v' \in (p, x_1, \ldots, x_s)A_\infty[(p^n/g)^{1/p^\infty}]$. We consider two expressions of $x_{s+1} g_0^{1/p^a} z$:

$$x_{s+1} g_0^{1/p^a} u + p^{(n-m)/p^a} x_{s+1} u' = x_{s+1} g_0^{1/p^a} z = g_0^{1/p^a} v + p^{(n-m)/p^a} v'.$$

From this we know that

$$x_{s+1}(g_0^{1/p^a} u) = g_0^{1/p^a} v + p^{(n-m)/p^a}(v' - x_{s+1}u'). \tag{3.4.1}$$

It follows from (3.4.1) that $p^{(n-m)/p^a}(v'-x_{s+1}u')\in A_\infty$ (since the other two terms are in A_∞). Thus by Lemma 3.3, $p^{(n-m)/p^a}(v'-x_{s+1}u')\in pA_\infty$ since $n>p^a+m$. But now (3.4.1) tells us that

$$x_{s+1}(g_0^{1/p^a}u)\in(p,x_1,\dots,x_s)A_\infty+pA_\infty=(p,x_1,\dots,x_s)A_\infty.$$

Since $g_0^{1/p^a}u\in A_\infty$ and $p,x_1,\dots,x_{s+1}$ is an almost regular sequence on A_∞,

$$(pg_0)^{1/p^a}u\in(p,x_1,\dots,x_s)A_\infty.$$

But now

$$(pg_0)^{1/p^a}z=(pg_0)^{1/p^a}u+p^{1/p^a}p^{(n-m)/p^a}u'.$$

Therefore

$$(pg_0)^{1/p^a}z\in(p,x_1,\dots,x_s)A_\infty+pA_\infty\Big[\Big(\frac{p^n}{g}\Big)^{1/p^\infty}\Big]\subseteq(p,x_1,\dots,x_s)A_\infty\Big[\Big(\frac{p^n}{g}\Big)^{1/p^\infty}\Big].$$

Because $z\equiv p^{1/p^t}y$ modulo $pA_\infty[\widehat{(p^n/g)}^{1/p^\infty}]$ and g is a multiple of g_0, we have

$$p^{1/p^t}(pg)^{1/p^a}y\in(p,x_1,\dots,x_s)A_\infty\Big[\widehat{\Big(\frac{p^n}{g}\Big)^{1/p^\infty}}\Big]\subseteq(p,x_1,\dots,x_s)A_\infty\Big\langle\frac{p^n}{g}\Big\rangle.$$

Since this is true for all $t>0$, we have $(p^{1/p^\infty})(pg)^{1/p^a}$ annihilates

$$\frac{(p,x_1,\dots,x_s)A_\infty\langle\frac{p^n}{g}\rangle:x_{s+1}}{(p,x_1,\dots,x_s)A_\infty\langle\frac{p^n}{g}\rangle}.$$

Finally, since S is an almost finite projective $A_\infty\langle p^n/g\rangle$-algebra, by [Gabber and Ramero 2003, Lemma 2.4.31],

$$\frac{(p,x_1,\dots,x_s)S:x_{s+1}}{(p,x_1,\dots,x_s)S}=\operatorname{Hom}_S(S/x_{s+1},S/(p,x_1,\dots,x_s))$$

is almost isomorphic to

$$S\otimes\operatorname{Hom}_{A_\infty\langle p^n/g\rangle}\Big(A_\infty\Big\langle\frac{p^n}{g}\Big\rangle/x_{s+1},A_\infty\Big\langle\frac{p^n}{g}\Big\rangle/(p,x_1,\dots,x_s)\Big)=S\otimes\frac{(p,x_1,\dots,x_s)A_\infty\langle\frac{p^n}{g}\rangle:x_{s+1}}{(p,x_1,\dots,x_s)A_\infty\langle\frac{p^n}{g}\rangle}.$$

Therefore $(p^{1/p^\infty})(pg)^{1/p^a}$ annihilates $(p,x_1,\dots,x_s)S:x_{s+1}/(p,x_1,\dots,x_s)S$ as well. □

We need the following lemma:

Lemma 3.5 [Hochster 2002, Lemma 5.1]. *Let M be an R-module and let T be an R-algebra with a map $\alpha:M\to T[1/c]$. Let $M\to M'$ be a partial algebra modification of M with respect to part of a system of parameters $p,x_1,\dots,x_s,x_{s+1}$ with degree bound D. Suppose $x_{s+1}t_{s+1}=pt_0+x_1t_1+\dots+x_st_s$ with $t_j\in T$ implies $ct_{s+1}\in(p,x_1,\dots,x_s)T$ and $\alpha(M)\subseteq c^{-N}T$. Then there is an R-linear map $\beta:M'\to T[1/c]$ extending α with image contained in $c^{-N'}T$ where $N'=ND+N+D$ depends only on N and D.*

Proof of Theorem 3.1. Let R' be the normalization of R and let Q' be a height one prime of R' that lies over Q. Note that the residue field of R' is still k since we assumed k is algebraically closed. If we can construct weakly functorial big Cohen–Macaulay algebras for $R' \to R'/Q'$ then the same follows for $R \to R/Q$. Thus without loss of generality we can assume $(R, \mathfrak{m}, k)$ is normal. Let

$$\begin{tikzcd} R \arrow[r] \arrow[d] & R/Q \arrow[d] \\ R_{\infty,n} \arrow[r] & (R/Q)_{\infty,n} \end{tikzcd}$$

be the commutative diagram constructed in Lemma 2.3. Moreover, abusing notation slightly, suppose $g = p^{m_1} g_0$ in A and $\bar{g} = p^{m_2} \bar{g}_0$ in A/x_1A such that $p \nmid g_0$ and $p \nmid \bar{g}_0$.

Now $R_{\infty,n}$ and $(R/Q)_{\infty,n}$ are almost finite étale over $A_\infty\langle p^n/g\rangle$ and $(A/x_1A)_\infty\langle p^n/g\rangle$ respectively, in particular they are almost finite projective over $A_\infty\langle p^n/g\rangle$ and $(A/x_1A)_\infty\langle p^n/g\rangle$ respectively (see [Scholze 2012, Definition 4.3 and Proposition 4.10]). Lemma 3.4 shows that, for every n and p^a such that $n > p^a + m_1 + m_2$, with $c = (pg)^{2/p^a}$, if $x_{s+1}t_{s+1} = pt_0 + x_1t_1 + \cdots + x_st_s$ with $t_j \in R_{\infty,n}$ (resp. $(R/Q)_{\infty,n}$), we have that $ct_{s+1} \in (p, x_1, \ldots, x_s)R_{\infty,n}$ (resp. $(R/Q)_{\infty,n}$).

By Theorem 1.3, it suffices to show that there is no bad double sequence of partial algebra modifications of R. Suppose there is one:

$$R \to M_1 \to \cdots \to M_r \to (R/Q) \otimes M_r \to N_1 \to \cdots \to N_s.$$

We claim that there exists a commutative diagram:

$$\begin{tikzcd} R \arrow[r] \arrow[d] & M_1 \arrow[r] \arrow[d] & \cdots \arrow[r] & M_r \arrow[r] \arrow[d, "\alpha"] & (R/Q) \otimes M_r \arrow[r] \arrow[d] & N_1 \arrow[r] \arrow[d] & \cdots \arrow[r] & N_s \arrow[d, "\beta"] \\ R_{\infty,n}\left[\frac{1}{c}\right] \arrow[r, "="] & R_{\infty,n}\left[\frac{1}{c}\right] \arrow[r, "="] & \cdots \arrow[r, "="] & R_{\infty,n}\left[\frac{1}{c}\right] \arrow[r] & (R/Q)_{\infty,n}\left[\frac{1}{c}\right] \arrow[r, "="] & (R/Q)_{\infty,n}\left[\frac{1}{c}\right] \arrow[r, "="] & \cdots \arrow[r, "="] & (R/Q)_{\infty,n}\left[\frac{1}{c}\right] \end{tikzcd}$$

The leftmost vertical map is the natural one; the first half of the diagram exists by Lemma 3.5; the middle commutative diagram exists because the composite map $M_r \to R_{\infty,n}[1/c] \to (R/Q)_{\infty,n}[1/c]$ induces a map $(R/Q) \otimes M_r \to (R/Q)_{\infty,n}[1/c]$ since $(R/Q)_{\infty,n}[1/c]$ is an R/Q-algebra; the second half of the diagram exists by Lemma 3.5 again.

Let $D > 0$ be an integer larger than the degree bounds for all the partial algebra modifications in this sequence. Applying Lemma 3.5 repeatedly to the first half of the diagram, we know there is an integer M depending only on D, but not on n and p^a, such that the image of α is contained in $c^{-M}R_{\infty,n}$. The image of the map $(R/Q) \otimes M_r \to (R/Q)_{\infty,n}[1/c]$ is contained in $c^{-M}(R/Q)_{\infty,n}$ because $R_{\infty,n}[1/c] \to (R/Q)_{\infty,n}[1/c]$ is induced by $R_{\infty,n} \to (R/Q)_{\infty,n}$. But then applying Lemma 3.5 repeatedly to the second half of the diagram, we know that there exists an integer N depending on M and D (and hence only on D), but not on n and p^a, such that the image of β is contained in $c^{-N}(R/Q)_{\infty,n}$.

Now we chase the above diagram and we see that on the one hand, the element $1 \in R$ maps to $1 \in (R/Q)_{\infty,n}[1/c]$. But on the other hand, since the sequence is bad, the image of $1 \in R$ in N_s is in $\mathfrak{m}N_s$ and hence the image of $1 \in R$ is contained in $\mathfrak{m}c^{-N}(R/Q)_{\infty,n}$ in $(R/Q)_{\infty,n}[1/c]$. Therefore we have $1 \in \mathfrak{m}((pg)^{2/p^a})^{-N}(R/Q)_{\infty,n}$, that is,

$$(pg)^{2N/p^a} \in \mathfrak{m}(R/Q)_{\infty,n}.$$

Because $\mathfrak{m}$ is the maximal ideal of R and $A = W(k)[[x_1, \dots, x_{d-1}]] \to R$ is module-finite, $\mathfrak{m}^{N'} \subseteq (p, x_1, \dots, x_{d-1})R$ for some fixed N'. We thus have:

$$(pg)^{2NN'/p^a} \in (p, x_2, \dots, x_{d-1})(R/Q)_{\infty,n}.$$

Since $(R/Q)_{\infty,n}$ is almost finite étale over $(A/x_1A)_\infty\langle p^n/g\rangle$, we know that

$$(pg)^{(2NN'+1)/p^a} \in (p, x_2, \dots, x_{d-1})(A/x_1A)_\infty\left\langle \frac{p^n}{g}\right\rangle \cap (A/x_1A)_\infty.$$

But now Lemma 3.2 implies $(pg)^{(2NN'+2)/p^a} \in (p, x_2, \dots, x_{d-1})(A/x_1A)_\infty$ for all p^a. Because N, N' do not depend on p^a, we know that $pg \in (p, x_2, \dots, x_{d-1})^m(A/x_1A)_\infty$ for all $m > 0$. Since $(A/x_1A)_\infty$ is almost faithfully flat over $(A/x_1A)_{\infty,0}$ mod p^m, we know that

$$p^2g \in (p, x_2, \dots, x_{d-1})^m(A/x_1A)_{\infty,0} \cap (A/x_1A) = (p, x_2, \dots, x_{d-1})^m(A/x_1A)$$

for all $m > 0$ by faithful flatness of $(A/x_1A)_{\infty,0}$ over A/x_1A. But then

$$p^2g \in \cap_m(p, x_2, \dots, x_{d-1})^m(A/x_1A) = 0,$$

which is a contradiction. □

Remark 3.6. We point out that the quantitative form of Scholze's Hebbarkeitssatz [Bhatt 2018, Theorem 4.2] implies Lemma 3.2 and the following weaker form of Lemma 3.4: if $\{S_n\}_n$ is a pro-system such that S_n is an almost finite projective $A_\infty\langle p^n/g\rangle$-algebra, then for every $k \geq 1$ and $n \geq p^a + m$, $(p^{1/p^\infty})(pg)^{1/p^a}$ annihilates the image of $(p,x_1,\dots,x_s)S_{k+n} : x_{s+1}/(p,x_1,\dots,x_s)S_{k+n}$ in $(p,x_1,\dots,x_s)S_k : x_{s+1}/(p,x_1,\dots,x_s)S_k$. This weaker form is enough to establish Theorem 3.1, but one needs to modify the proof of Lemma 3.5 and Theorem 3.1: to extend each partial algebra modification to $R_{\infty,n}[1/c]$ one needs to decrease n roughly by p^a in order to trivialize bad relations (and keep control on the denominators). We leave it to the interested reader to carry out the details.

4. Applications

The results obtained in the preceding section are strong enough to establish the mixed-characteristic case of Hochster and Huneke's vanishing conjecture for maps of Tor [1995].

Theorem 4.1. *Let $A \to R \to S$ be maps of Noetherian rings such that $A \to S$ is a local homomorphism of mixed characteristic regular local rings and R is a module-finite torsion-free extension of A. Then for all A-modules M, the map* $\mathrm{Tor}_i^A(M, R) \to \mathrm{Tor}_i^A(M, S)$ *vanishes for all $i \geq 1$.*

We need the following important reduction. This reduction is known to experts and is proved implicitly in [Ranganathan 2000, Chapter 5.2] and [Hochster 2017, Section 13]. We will give a sketch of the proof.

Lemma 4.2. *To prove Theorem 4.1, we can assume* $(A, \mathfrak{m})$ *is complete, R is a complete local domain, and $S = A/xA$ where $x \in \mathfrak{m} - \mathfrak{m}^2$.*

Sketch of proof. We can assume M is finitely generated. Replacing M by its first module of syzygies over A repeatedly, we only need to prove the case $i = 1$. We may further assume $M = A/I$ by [Ranganathan 2000, Lemma 5.2.1] or [Hochster 2017, Page 15].[5] Next by [Hochster and Huneke 1995, (4.5)(a)], we can assume A and S are both complete, R is a complete local domain, and $A \to S$ is surjective; i.e., $S = A/P$ where P is generated by part of a regular system of parameters of A (note that $p \notin P$ since S has mixed characteristic). It follows that $S = R/Q$ for some prime ideal Q of R lying over P. After all these reductions, we note that by [Hochster 2017, Lemma 13.6], $\mathrm{Tor}_1^A(A/I, R) \to \mathrm{Tor}_1^A(A/I, S)$ vanishes if and only if $IQ \cap P = IP$.

We next want to reduce to the case that P is generated by one element. The trick is to replace A by its extended Rees ring $\widetilde{A} = A[Pt, t^{-1}]$, R by $\widetilde{R} = R[Pt, t^{-1}]$ and S by $\widetilde{S} = \widetilde{A}/t^{-1}\widetilde{A}$. Since P is generated by part of a regular system of parameters, $\widetilde{A}$ and $\widetilde{S}$ are still regular. The point is that there is a homogeneous prime ideal $\widetilde{Q} \subseteq \widetilde{R}$ that contains Q and contracts to $t^{-1}\widetilde{A} \subseteq \widetilde{A}$ (see [Ranganathan 2000, Proof of Theorem 5.2.6] or [Hochster 2017, Page 16]), thus we have $\widetilde{A} \to \widetilde{R} \to \widetilde{S}$. Therefore if we can prove Theorem 4.1 for $\widetilde{A} \to \widetilde{R} \to \widetilde{S}$ and $M = \widetilde{A}/I\widetilde{A}$, then [Hochster 2017, Lemma 13.6] implies that $I\widetilde{Q} \cap t^{-1}\widetilde{A} = It^{-1}\widetilde{A}$. Comparing the degree 0 part, we see that $IQ \cap P = IP$.

Finally, we can localize $\widetilde{A}$ and $\widetilde{S}$ and complete, and reduce to the case $\widetilde{R}$ is a complete local domain as in [Hochster and Huneke 1995, (4.5)(a)]. Note that $\widetilde{S}$ is obtained from $\widetilde{A}$ by killing one element (and we may assume $\widetilde{S}$ still has mixed characteristic after localization). We thus obtain all the desired reductions. ☐

Proof of Theorem 4.1. By Lemma 4.2, we may assume R is a complete local domain and $S = A/xA$. It follows that $S = R/Q$ for a height one prime Q of R. Since $A \to S$ and $R \to S$ are both surjective, A, R, S have the same residue field k. We fix a coefficient ring $W(k)$ of A, then the images of $W(k)$ in R and S are also coefficient rings of R and S. Replacing A, R, S by their faithfully flat extensions $A\widehat{\otimes}_{W(k)}W(\bar{k})$, $R\widehat{\otimes}_{W(k)}W(\bar{k})$, $S\widehat{\otimes}_{W(k)}W(\bar{k})$ does not affect whether the map on Tor vanishes or not. Thus without loss of generality we may assume k is algebraically closed.

By Theorem 3.1, we have a commutative diagram:

$$\begin{tikzcd} R \arrow[r] \arrow[d] & S = R/Q \arrow[d] \\ B \arrow[r] & C \end{tikzcd}$$

[5] In this process we may lose A and S being local, but we can always localize A and S again to assume they are local (and have mixed characteristic, since otherwise Theorem 4.1 is known).

where B and C are balanced big Cohen–Macaulay algebras for R and S respectively. This induces a commutative diagram:

$$\begin{array}{ccc} \mathrm{Tor}_i^A(M, R) & \longrightarrow & \mathrm{Tor}_i^A(M, S) \\ \downarrow & & \downarrow \\ \mathrm{Tor}_i^A(M, B) & \longrightarrow & \mathrm{Tor}_i^A(M, C) \end{array}$$

Since B is a balanced big Cohen–Macaulay algebra over R (and hence also over A), it is faithfully flat over A so $\mathrm{Tor}_i^A(M, B) = 0$ for all $i \geq 1$. Moreover, C is faithfully flat over S since it is a balanced big Cohen–Macaulay algebra over S and S is regular, thus $\mathrm{Tor}_i^A(M, S) \to \mathrm{Tor}_i^A(M, C)$ is injective. Chasing the diagram above we know that the map $\mathrm{Tor}_i^A(M, R) \to \mathrm{Tor}_i^A(M, S)$ vanishes for all $i \geq 1$. $\square$

A local ring $(R, \mathfrak{m})$ of dimension d is called *pseudorational* if it is normal, Cohen–Macaulay, analytically unramified (i.e., the completion $\widehat{R}$ is reduced), and if for every projective and birational map $\pi\colon W \to \operatorname{Spec} R$, the canonical map $H^d_{\mathfrak{m}}(R) \to H^d_E(W, O_W)$ is injective where $E = \pi^{-1}(\mathfrak{m})$ denotes the closed fiber. In characteristic 0, pseudorational singularities are the same as rational singularities. Very recently, Kovács [2017] has proved a remarkable result that, in *all* characteristics, if $\pi\colon X \to \operatorname{Spec} R$ is projective and birational, where X is Cohen–Macaulay and R is pseudorational, then $\boldsymbol{R}\pi_* O_X = R$.

In equal characteristic, direct summands of regular rings are pseudorational [Boutot 1987; Hochster and Huneke 1990]. This is usually called Boutot's theorem. It is well known that the vanishing conjecture for maps of Tor in a given characteristic implies that direct summands of regular rings are Cohen–Macaulay [Hochster and Huneke 1995, (4.3)]. What we want to prove next is the analog of Boutot's theorem that direct summands of regular rings are pseudorational in mixed characteristic. This in fact also follows formally from the vanishing conjecture for maps of Tor [Ma 2018]. Since the full details were not written down explicitly there, we give a complete argument here. We first recall the following Sancho de Salas exact sequence [1987].

Let $T = R[Jt] = R \oplus Jt \oplus J^2t^2 \oplus \cdots$ and let $W = \operatorname{Proj} T \to \operatorname{Spec} R$ be the blow up with $E = \pi^{-1}(\mathfrak{m})$. Pick $f_1, \dots, f_n \in Jt = [T]_1$ such that $U = \{U_i = \operatorname{Spec}[T_{f_i}]_0\}$ is an affine open cover of W. We have an exact sequence of chain complexes:

$$0 \to \check{C}^\bullet(U, O_W)[-1] \to [C^\bullet(f_1, \dots, f_n, T)]_0 \to R \to 0.$$

Since $\check{C}^\bullet(U, O_W) \cong \boldsymbol{R}\pi_* O_W$ and $C^\bullet(f_1, \dots, f_n, T) = [\boldsymbol{R}\Gamma_{T_{>0}} T]_0$, the above sequence gives us (after rotating) an exact triangle:

$$[\boldsymbol{R}\Gamma_{T_{>0}} T]_0 \to R \to \boldsymbol{R}\pi_* O_W \xrightarrow{+1}$$

Applying $\boldsymbol{R}\Gamma_{\mathfrak{m}}$, we get:

$$[\boldsymbol{R}\Gamma_{\mathfrak{m}+T_{>0}} T]_0 \to \boldsymbol{R}\Gamma_{\mathfrak{m}} R \to \boldsymbol{R}\Gamma_{\mathfrak{m}} \boldsymbol{R}\pi_* O_W \xrightarrow{+1} .$$

Taking cohomology we get the Sancho de Salas exact sequence:

$$\begin{tikzcd} {} \arrow[r] & h^d([\boldsymbol{R}\Gamma_{\mathfrak{m}+T_{>0}}T]_0) \arrow[r] \arrow[d, "="] & h^d(\boldsymbol{R}\Gamma_{\mathfrak{m}}R) \arrow[r] \arrow[d, "="] & h^d(\boldsymbol{R}\Gamma_{\mathfrak{m}}\boldsymbol{R}\pi_*O_W) \arrow[r] \arrow[d, "="] & {} \\ {} \arrow[r] & {[H^d_{\mathfrak{m}+T_{>0}}(T)]_0} \arrow[r] & H^d_{\mathfrak{m}}(R) \arrow[r] & H^d_E(W, O_W) \arrow[r] & {} \end{tikzcd} \tag{4.2.1}$$

We also recall that $R \to S$ is *pure* if $R \otimes M \to S \otimes M$ is injective for every R-module M. This is slightly weaker than saying that $R \to S$ splits as a map of R-modules. If R is an A-algebra and $R \to S$ is pure, then $\operatorname{Tor}_i^A(M, R) \to \operatorname{Tor}_i^A(M, S)$ is injective for every i [Hochster and Huneke 1995, (2.1)(h)], in particular, $H^i_{\mathfrak{m}}(R) \to H^i_{\mathfrak{m}}(S)$ is injective for every i.

We are ready to prove the following corollary. We state the result in the local setting, but the general case reduces immediately to the local case.

Corollary 4.3. *Let $(R, \mathfrak{m}) \to (S, \mathfrak{n})$ be a pure map of local rings such that $(S, \mathfrak{n})$ is regular of mixed characteristic. Then R is pseudorational. In particular, direct summands of regular rings are pseudorational.*

Proof. We can complete R and S at $\mathfrak{m}$ and $\mathfrak{n}$ respectively to assume both R and S are complete; R is normal since pure subrings of normal domains are normal. By Cohen's structure theorem, we have a module-finite extension $A \to R$ such that A is a complete regular local ring. Let $x_1, \dots, x_d$ be a regular system of parameters of A. We apply Theorem 4.1 to $M = A/(x_1, \dots, x_d)$. We have

$$\operatorname{Tor}_i^A(A/(x_1, \dots, x_d), R) \to \operatorname{Tor}_i^A(A/(x_1, \dots, x_d), S)$$

vanishes for all $i \geq 1$. However, we also know that this map is injective because $R \to S$ is pure. Thus we have $\operatorname{Tor}_i^A(A/(x_1, \dots, x_d), R) = H_i(x_1, \dots, x_d, R) = 0$ for all $i \geq 1$. This implies $x_1, \dots, x_d$ is a regular sequence on R and hence R is Cohen–Macaulay. Obviously, the complete local domain R is analytically unramified.

We now check the last condition of pseudorationality. Let $W \to \operatorname{Spec} R$ be a projective birational map, thus $W \cong \operatorname{Proj} T = \operatorname{Proj} R \oplus Jt \oplus J^2t^2 \oplus \cdots$ for some ideal $J \subseteq R$. We now apply the Sancho de Salas exact sequence (4.2.1) to get:

$$\begin{tikzcd} {[H^d_{\mathfrak{m}+T_{>0}}(T)]_0} \arrow[r] \arrow[d, hook] & H^d_{\mathfrak{m}}(R) \arrow[r] \arrow[d, "="] & H^d_E(W, O_W) \\ H^d_{\mathfrak{m}+T_{>0}}(T) \arrow[r] & H^d_{\mathfrak{m}}(R). & \end{tikzcd}$$

Thus in order to show $H^d_{\mathfrak{m}}(R) \to H^d_E(W, O_W)$ is injective, it suffices to show $H^d_{\mathfrak{m}+T_{>0}}(T) \to H^d_{\mathfrak{m}}(R)$ vanishes. We can localize T at the maximal ideal $\mathfrak{m} + T_{>0}$, complete, and kill a minimal prime without affecting whether the map vanishes or not. Hence it is enough to show that if $(T, \mathfrak{m}) \twoheadrightarrow (R, \mathfrak{m})$ is a surjection such that T is a complete local domain of dimension $d+1$, then $H^d_{\mathfrak{m}}(T) \to H^d_{\mathfrak{m}}(R)$ vanishes. By Cohen's structure theorem there exists $(A, \mathfrak{m}_0) \to (T, \mathfrak{m})$ a module-finite extension such that A is a complete regular local ring. We consider the chain of maps

$$A \to T \to R \to S$$

Applying Theorem 4.1 to $A \to T \to S$ and $M = H^{d+1}_{\mathfrak{m}_0}(A)$, we know that the composite map

$$\mathrm{Tor}_1^A(H^{d+1}_{\mathfrak{m}_0}(A), T) \to \mathrm{Tor}_1^A(H^{d+1}_{\mathfrak{m}_0}(A), R) \to \mathrm{Tor}_1^A(H^{d+1}_{\mathfrak{m}_0}(A), S)$$

vanishes. Since the Čech complex on a regular system of parameters gives a flat resolution of $H^{d+1}_{\mathfrak{m}_0}(A)$ over A, we know that $\mathrm{Tor}_1^A(H^{d+1}_{\mathfrak{m}_0}(A), N) \cong H^d_{\mathfrak{m}_0}(N)$ for every A-module N. Thus the composite map

$$H^d_{\mathfrak{m}}(T) \to H^d_{\mathfrak{m}}(R) \to H^d_{\mathfrak{m}}(S)$$

vanishes. But then $H^d_{\mathfrak{m}}(T) \to H^d_{\mathfrak{m}}(R)$ vanishes because $H^d_{\mathfrak{m}}(R) \to H^d_{\mathfrak{m}}(S)$ is injective since $R \to S$ is pure. □

Remark 4.4. Corollary 4.3 can be also obtained by combining the main results of [André 2018a; Bhatt 2018] and using the following argument: the existence of weakly functorial big Cohen–Macaulay algebras for injective ring homomorphisms [André 2018a, Remarque 4.2.1] implies that direct summands of regular rings are Cohen–Macaulay, but we also know they are derived splinters (because this is true for regular rings by [Bhatt 2018, Theorem 1.2] and it is easy to see that direct summand of derived splinters are still derived splinters). Now the argument of [Kovács 2017, Lemma 7.5] implies that Cohen–Macaulay derived splinters are pseudorational.

Remark 4.5. Last we point out that by [Ma 2018, Remark 5.12], Theorem 4.1 gives a new proof of the derived direct summand conjecture [Bhatt 2018, Theorem 6.1], that is, if R is a complete regular local ring of mixed characteristic and $\pi\colon X \to \operatorname{Spec} R$ is a proper surjective map, then $R \to \boldsymbol{R}\pi_* O_X$ splits in the derived category of R-modules. Our proof is different from Bhatt's in that it does not use Scholze's vanishing theorem [2012, Proposition 6.14]. In fact, tracing the arguments of [Ma 2018, Theorem 5.11 and Remark 5.13], one can show that our Theorem 3.1 leads to a stronger result that complete local rings that are pure inside all their big Cohen–Macaulay algebras (e.g., complete regular local rings) are derived splinters.

Acknowledgement

It is a pleasure to thank Mel Hochster for many enjoyable discussions on the vanishing conjecture for maps of Tor and other homological conjectures. We would like to thank Bhargav Bhatt for explaining the basic theory of perfectoid spaces to us and for pointing out Remark 3.6. We would also like to thank Kiran Kedlaya, Karl Schwede, Kazuma Shimomoto and Chris Skalit for very helpful discussions. Ma is partially supported by NSF Grant DMS #1836867/1600198, and NSF CAREER Grant DMS #1252860/1501102.

References

[André 2018a] Y. André, "La conjecture du facteur direct", *Publ. Math. Inst. Hautes Études Sci.* **127** (2018), 71–93. MR

[André 2018b] Y. André, "Le lemme d'Abhyankar perfectoïde", *Publ. Math. Inst. Hautes Études Sci.* **127** (2018), 1–70. MR

[Bhatt 2018] B. Bhatt, "On the direct summand conjecture and its derived variant", *Invent. Math.* **212**:2 (2018), 297–317. MR

[Boutot 1987] J.-F. Boutot, "Singularités rationnelles et quotients par les groupes réductifs", *Invent. Math.* **88**:1 (1987), 65–68. MR Zbl

[Bruns and Herzog 1993] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Adv. Math. **39**, Cambridge Univ. Press, 1993. MR Zbl

[Gabber and Ramero 2003] O. Gabber and L. Ramero, *Almost ring theory*, Lect. Notes in Math. **1800**, Springer, 2003. MR Zbl

[Heitmann 2002] R. C. Heitmann, "The direct summand conjecture in dimension three", *Ann. of Math.* (2) **156**:2 (2002), 695–712. MR Zbl

[Hochster 1975a] M. Hochster, "Big Cohen–Macaulay modules and algebras and embeddability in rings of Witt vectors", pp. 106–195 in *Conference on commutative algebra*, edited by A. V. Geramita, Queen's Papers on Pure and Appl. Math. **42**, Queen's Univ., Kingston, ON, 1975. MR Zbl

[Hochster 1975b] M. Hochster, *Topics in the homological theory of modules over commutative rings* (Lincoln, NE, 1974), CBMS Regional Conference Series in Math. **24**, Amer. Math. Soc., Providence, RI, 1975. MR Zbl

[Hochster 2002] M. Hochster, "Big Cohen–Macaulay algebras in dimension three via Heitmann's theorem", *J. Algebra* **254**:2 (2002), 395–408. MR Zbl

[Hochster 2017] M. Hochster, "Homological conjectures and lim Cohen–Macaulay sequences", pp. 173–197 in *Homological and computational methods in commutative algebra* (Cortona, Italy, 2016), edited by A. Conca et al., Springer INdAM Ser. **20**, Springer, 2017. MR

[Hochster and Huneke 1990] M. Hochster and C. Huneke, "Tight closure, invariant theory, and the Briançon–Skoda theorem", *J. Amer. Math. Soc.* **3**:1 (1990), 31–116. MR Zbl

[Hochster and Huneke 1995] M. Hochster and C. Huneke, "Applications of the existence of big Cohen–Macaulay algebras", *Adv. Math.* **113**:1 (1995), 45–117. MR Zbl

[Kovács 2017] S. J. Kovács, "Rational singularities", preprint, 2017. arXiv

[Ma 2018] L. Ma, "The vanishing conjecture for maps of Tor and derived splinters", *J. Eur. Math. Soc.* **20**:2 (2018), 315–338. MR Zbl

[Ranganathan 2000] N. Ranganathan, *Splitting in module-finite extension rings and the vanishing conjecture for maps of Tor*, Ph.D. thesis, University of Michigan, 2000, Available at https://search.proquest.com/docview/304609163.

[Sancho de Salas 1987] J. B. Sancho de Salas, "Blowing-up morphisms with Cohen–Macaulay associated graded rings", pp. 201–209 in *Géométrie algébrique et applications, I* (La Rábida, Spain, 1984), edited by J.-M. Aroca et al., Travaux en Cours **22**, Hermann, Paris, 1987. MR Zbl

[Scholze 2012] P. Scholze, "Perfectoid spaces", *Publ. Math. Inst. Hautes Études Sci.* **116** (2012), 245–313. MR Zbl

Communicated by Craig Huneke
Received 2017-05-18 Revised 2018-02-07 Accepted 2018-03-13

heitmann@math.utexas.edu *University of Texas, Austin, Austin, TX, United States*

ma326@purdue.edu *Purdue University, West Lafayette, IN, United States*

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1675

Blocks of the category of smooth ℓ-modular representations of $\mathrm{GL}(n, F)$ and its inner forms: reduction to level 0

Gianmarco Chinello

Let G be an inner form of a general linear group over a nonarchimedean locally compact field of residue characteristic p, let R be an algebraically closed field of characteristic different from p and let $\mathscr{R}_R(G)$ be the category of smooth representations of G over R. In this paper, we prove that a block (indecomposable summand) of $\mathscr{R}_R(G)$ is equivalent to a level-0 block (a block in which every simple object has nonzero invariant vectors for the pro-p-radical of a maximal compact open subgroup) of $\mathscr{R}_R(G')$, where G' is a direct product of groups of the same type of G.

Introduction

Let F be a nonarchimedean locally compact field of residue characteristic p and let D be a central division algebra of finite dimension over F whose reduced degree is denoted by d. Given $m \in \mathbb{N}^*$, we consider the group $G = \mathrm{GL}_m(D)$ which is an inner form of $\mathrm{GL}_{md}(F)$. Let R be an algebraically closed field of characteristic $\ell \neq p$ and let $\mathscr{R}_R(G)$ be the category of smooth representations of G over R, that are called ℓ-modular when ℓ is positive. In this paper, we are interested in the Bernstein decomposition of $\mathscr{R}_R(G)$ (see [Sécherre and Stevens 2016] or [Vignéras 1998] for $d = 1$) that is its decomposition as a direct sum of full indecomposable subcategories, called *blocks*. Actually a full understanding of blocks of $\mathscr{R}_R(G)$ is equivalent to a full understanding of the whole category.

The main purpose of this paper is to find an equivalence of categories between any block of $\mathscr{R}_R(G)$ and a level-0 block of $\mathscr{R}_R(G')$ where G' is a suitable direct product of inner forms of general linear groups over finite extensions of F. We recall that a level-0 block of $\mathscr{R}_R(G')$ is a block in which every object has nonzero invariant vectors for the pro-p-radical of a maximal compact open subgroup of G'. This result is an important step in the attempt to describe blocks of $\mathscr{R}_R(G)$ because it reduces the problem to the description of level-0 blocks.

In the case of complex representations, Bernstein [1984] found a block decomposition of $\mathscr{R}_{\mathbb{C}}(G)$ indexed by pairs (M, σ) where M is a Levi subgroup of G and σ is an irreducible cuspidal representation

This work is partially part of the PhD thesis of the author and he wants to thank his supervisor, Vincent Sécherre, for his support and his comments on this paper.
MSC2010: primary 20C20; secondary 22E50.
Keywords: equivalence of categories, blocks, modular representations of p-adic reductive groups, type theory, semisimple types, Hecke algebras, level-0 representations.

of M, up to a certain equivalence relation called *inertial equivalence*. In particular an irreducible representation π of G is in the block associated to the inertial class of (M, σ) if its cuspidal support is in this class. Bushnell and Kutzko [1998] introduced a method to describe the blocks of $\mathscr{R}_{\mathbb{C}}(G)$: the *theory of types*. This method consists in associating to every block of $\mathscr{R}_{\mathbb{C}}(G)$ a pair (J, λ), called a type, where J is a compact open subgroup of G and λ is an irreducible representation of J, such that the simple objects of the block are the irreducible subquotients of the compactly induced representation $\mathrm{ind}_J^G(\lambda)$. In this case the block is equivalent to the category of modules over the $\mathbb{C}$-algebra $\mathscr{H}_{\mathbb{C}}(G, \lambda)$ of G-endomorphisms of $\mathrm{ind}_J^G(\lambda)$. Sécherre and Stevens [2012] (see [Bushnell and Kutzko 1999] for $d = 1$) described explicitly this algebra as a tensor product of algebras of type A.

In the case of ℓ-modular representations, Sécherre and Stevens [2016] (see [Vignéras 1998] for $d = 1$) found a block decomposition of $\mathscr{R}_R(G)$ indexed by inertial classes of pairs (M, σ) where M is a Levi subgroup of G and σ is an irreducible supercuspidal representation of M. In particular an irreducible representation π of G is in the block associated to the inertial class of (M, σ) if its supercuspidal support is in this class. We recall that the notions of cuspidal and supercuspidal representations are not equivalent as in complex case; however, Mínguez and Sécherre [2014a] proved the uniqueness of supercuspidal support, up to conjugation, for every irreducible representation of G. We remark that to obtain the block decomposition of $\mathscr{R}_R(G)$, Sécherre and Stevens do not use the same method as Bernstein, but they rely, like us in this paper, on the theory of semisimple types developed in [Sécherre and Stevens 2012] (see [Bushnell and Kutzko 1999] for $d = 1$). Actually, they associate to every block of $\mathscr{R}_R(G)$ a pair $(\boldsymbol{J}, \boldsymbol{\lambda})$, called a *semisimple supertype*. Unfortunately the construction of the equivalence, as in the complex case, between the block and the category of modules over $\mathscr{H}_R(G, \boldsymbol{\lambda})$ does not hold and one of the problems that occurs is that the pro-order of $\boldsymbol{J}$ can be divisible by ℓ. Some partial results on descriptions of algebras which are Morita equivalent to blocks of $\mathscr{R}_R(\mathrm{GL}_n(F))$ are given in [Dat 2012; Helm 2016; Guiraud 2013].

The idea of this paper is the following. We fix a block $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ of $\mathscr{R}_R(G)$ associated to the semisimple supertype $(\boldsymbol{J}, \boldsymbol{\lambda})$ and, as in [Sécherre and Stevens 2016], we can associate to it a compact open subgroup $\boldsymbol{J}_{\max}$ of G, its pro-p-radical $\boldsymbol{J}^1_{\max}$ and an irreducible representation $\boldsymbol{\eta}_{\max}$ of $\boldsymbol{J}^1_{\max}$. We remark that we can extend, not uniquely, $\boldsymbol{\eta}_{\max}$ to an irreducible representation $\boldsymbol{\kappa}_{\max}$ of $\boldsymbol{J}_{\max}$. Thus, we denote $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ the direct sum of blocks of $\mathscr{R}_R(G)$ associated to $(\boldsymbol{J}^1_{\max}, \boldsymbol{\eta}_{\max})$ and we consider the functor

$$\boldsymbol{M}_{\boldsymbol{\eta}_{\max}} = \mathrm{Hom}_G(\mathrm{ind}_{\boldsymbol{J}^1_{\max}}^G \boldsymbol{\eta}_{\max}, -) : \mathscr{R}(G, \boldsymbol{\eta}_{\max}) \longrightarrow \text{Mod-}\,\mathscr{H}_R(G, \boldsymbol{\eta}_{\max}),$$

where $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max}) \cong \mathrm{End}_G(\mathrm{ind}_{\boldsymbol{J}^1_{\max}}^G(\boldsymbol{\eta}_{\max}))$. Using the fact that $\boldsymbol{\eta}_{\max}$ is a projective representation, since $\boldsymbol{J}^1_{\max}$ is a pro-p-group, we prove that $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}$ is an equivalence of categories (Theorem 5.10). This result generalizes Corollary 3.3 of [Chinello 2017] where $\boldsymbol{\eta}_{\max}$ is a trivial character. We can also associate to $(\boldsymbol{J}, \boldsymbol{\lambda})$ a Levi subgroup L of G and a group $B_L^{\times}$, which is a direct product of inner forms of general linear groups over finite extensions of F and which we have denoted G' above. If K_L is a maximal compact open subgroup of $B_L^{\times}$ and K_L^1 is its pro-p-radical then $K_L/K_L^1 \cong \boldsymbol{J}_{\max}/\boldsymbol{J}^1_{\max} = \mathscr{G}$ is a direct product of finite general linear groups. Actually, in [Chinello 2017] it is proved that the K_L^1-invariants functor $\mathrm{inv}_{K_L^1}$ is an equivalence of categories between the level-0 subcategory $\mathscr{R}(B_L^{\times}, K_L^1)$ of $\mathscr{R}_R(B_L^{\times})$, which is the direct sum of its level-0 blocks, and the category of modules over the algebra $\mathscr{H}_R(B_L^\times, K_L^1) \cong \mathrm{End}_{B_L^\times}(\mathrm{ind}_{K_L^1}^{B_L^\times} 1_{K_L^1})$. Now, thanks to the explicit presentation by generators and relations of $\mathscr{H}_R(B_L^\times, K_L^1)$ presented in [Chinello 2017], in this paper we construct a homomorphism $\Theta_{\gamma,\boldsymbol{\kappa}_{\max}} : \mathscr{H}_R(B_L^\times, K_L^1) \longrightarrow \mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ finding elements in $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ satisfying all relations defining $\mathscr{H}_R(B_L^\times, K_L^1)$. This homomorphism depends on the choice of the extension $\boldsymbol{\kappa}_{\max}$ of $\boldsymbol{\eta}_{\max}$ to $\boldsymbol{J}_{\max}$ and on the choice of an intertwining element γ of $\boldsymbol{\eta}_{\max}$. Moreover, using some properties of $\boldsymbol{\eta}_{\max}$, we prove that this homomorphism is actually an isomorphism. We remark that finding this isomorphism is one of the most difficult results obtained in this article and the proof in the case $L = G$ takes about half of the paper (Section 3). In this way we obtain an equivalence of categories $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}} : \mathscr{R}(G, \boldsymbol{\eta}_{\max}) \longrightarrow \mathscr{R}(B_L^\times, K_L^1)$ such that the following diagram commutes:

$$\begin{tikzcd}
\mathscr{R}(G, \boldsymbol{\eta}_{\max}) \arrow[r, "\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}"] \arrow[d, "\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}"', "\wr"] & \mathscr{R}(B_L^\times, K_L^1) \arrow[d, "\mathrm{inv}_{K_L^1}", "\wr"'] \\
\text{Mod-}\mathscr{H}_R(G, \boldsymbol{\eta}_{\max}) \arrow[r, "\Theta^*_{\gamma,\boldsymbol{\kappa}_{\max}}"] & \text{Mod-}\mathscr{H}_R(B_L^\times, K_L^1).
\end{tikzcd}$$

Then we obtain

$$\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(\pi, V) = \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\pi, V) \otimes_{\mathscr{H}_R(B_L^\times, K_L^1)} \mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1})$$

for every (π, V) in $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$, where the action of $\mathscr{H}_R(B_L^\times, K_L^1)$ on $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\pi, V)$ depends on $\Theta_{\gamma,\boldsymbol{\kappa}_{\max}}$. Hence, $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ induces an equivalence of categories between the block $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ and a level-0 block of $\mathscr{R}_R(B_L^\times)$. To understand this correspondence we need to use the functor

$$\mathsf{K}_{\boldsymbol{\kappa}_{\max}} : \mathscr{R}(G, \boldsymbol{\eta}_{\max}) \longrightarrow \mathscr{R}_R(\boldsymbol{J}_{\max}/\boldsymbol{J}^1_{\max}) = \mathscr{R}_R(\mathscr{G}),$$

where $\boldsymbol{J}_{\max}$ acts on $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(\pi) = \mathrm{Hom}_{\boldsymbol{J}^1_{\max}}(\boldsymbol{\eta}_{\max}, \pi)$ by $x.\varphi = \pi(x) \circ \varphi \circ \boldsymbol{\kappa}_{\max}(x)^{-1}$ for every representation π of G, $\varphi \in \mathrm{Hom}_{\boldsymbol{J}^1_{\max}}(\boldsymbol{\eta}_{\max}, \pi)$ and $x \in \boldsymbol{J}_{\max}$. This functor is strongly used in [Sécherre and Stevens 2016] to define $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ and to prove the Bernstein decomposition of $\mathscr{R}_R(G)$. We also consider the functor $\mathsf{K}_{K_L} : \mathscr{R}(B_L^\times, K_L^1) \to \mathscr{R}_R(K_L/K_L^1) = \mathscr{R}_R(\mathscr{G})$ given by $\mathsf{K}_{K_L}(Z) = Z^{K_L^1}$ for every representation (ϱ, Z) of $B_L^\times$ where $x \in K_L$ acts on $z \in Z^{K_L^1}$ by $x.z = \varrho(x)z$. Then the functors $\mathsf{K}_{K_L} \circ \boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ and $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}$ are naturally isomorphic (Proposition 5.14) and so $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ is equivalent to the level-0 block $\mathscr{B}$ of $\mathscr{R}_R(B_L^\times)$ such that $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})) = \mathsf{K}_{K_L}(\mathscr{B})$. More precisely, if $\boldsymbol{J}^1$ is the pro-p-radical of $\boldsymbol{J}$, then $\boldsymbol{J}/\boldsymbol{J}^1 = \mathscr{M}$ is a Levi subgroup of $\mathscr{G}$ and the choice of $\boldsymbol{\kappa}_{\max}$ defines a decomposition $\boldsymbol{\lambda} = \boldsymbol{\kappa} \otimes \boldsymbol{\sigma}$ where $\boldsymbol{\kappa}$ is an irreducible representation of $\boldsymbol{J}$ and $\boldsymbol{\sigma}$ is a cuspidal representation of $\mathscr{M}$ viewed as an irreducible representation of $\boldsymbol{J}$ trivial on $\boldsymbol{J}^1$. If we can consider the pair $(\mathscr{M}, \boldsymbol{\sigma})$ up to the equivalence relation given in Definition 1.14 of [Sécherre and Stevens 2016], then a representation (ϱ, Z) of $B_L^\times$ is in $\mathscr{B}$ if it is generated by the maximal subspace of $Z^{K_L^1}$ such that every irreducible subquotient has supercuspidal support in the class of $(\mathscr{M}, \boldsymbol{\sigma})$.

One question we do not address in this paper is the structure of level-0 blocks of $\mathscr{R}_R(B_L^\times)$ when the characteristic of R is positive. Thanks to results of [Chinello 2017] we know that there is a correspondence between these blocks and the set $\mathscr{E}$ of primitive central idempotents of $\mathscr{H}_R(B_L^\times, K_L^1)$, which are described in Sections 2.5 and 2.6 of [Chinello 2015]. Hence, one possibility for understanding level-0 blocks of

$\mathscr{R}_R(B_L^\times)$ is to describe the algebras $e\mathscr{H}_R(B_L^\times, K_L^1)$ with $e \in \mathscr{E}$. On the other hand, we recall that Dat [2018] proved that every level-0 block of $\mathscr{R}_R(\mathrm{GL}_n(F))$ is equivalent to the unipotent block of $\mathscr{R}_R(G'')$, where G'' is a suitable product of general linear groups over nonarchimedean locally compact fields. Hence, putting together the result of Dat and results of this article, we obtain a method to reduce the description of any block of $\mathscr{R}_R(\mathrm{GL}_n(F))$ to that of a unipotent block. Unfortunately the description of the unipotent block of $\mathscr{R}_R(\mathrm{GL}_n(F))$, or of $\mathscr{R}_R(G)$, is nowadays a hard question and it has no answer yet.

We now summarize the contents of each section of this paper. In Section 1 we present general results on the convolution Hecke algebras $\mathscr{H}_R(\mathtt{G}, \sigma)$ where $\mathtt{G}$ is an arbitrary locally profinite group and σ a representation of an open subgroup $\mathtt{H}$ of $\mathtt{G}$. We see that if σ is finitely generated then $\mathscr{H}_R(\mathtt{G}, \sigma)$ is isomorphic to the endomorphism algebra of $\mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}\,\sigma$. We define two subcategories of $\mathscr{R}_R(\mathtt{G})$ and prove that, when they coincide, they are equivalent to the category of modules over $\mathscr{H}_R(\mathtt{G}, \sigma)$. In Section 2 we introduce the theory of maximal simple types; we consider the Heisenberg representation η associated to a simple character (see Section 2A) and define the groups $B^\times = B_G^\times$ and $K^1 = K_G^1$. In Section 3 we prove that the algebras $\mathscr{H}_R(G, \eta)$ and $\mathscr{H}_R(B^\times, K^1)$ are isomorphic. In Section 4 we introduce the theory of semisimple types, define the representation $\boldsymbol{\eta}_{\max}$ and the group $B_L^\times$, and prove that the algebras $\mathscr{H}_R(B_L^\times, K_L^1)$ and $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ are isomorphic. In Section 5 we prove that $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}$ and $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ are equivalences of categories; we describe the correspondence between blocks of $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and of $\mathscr{R}(B_L^\times, K_L^1)$ and investigate the dependence of these results on the choice of the extension of $\boldsymbol{\eta}_{\max}$ to $\boldsymbol{J}_{\max}$.

1. Preliminaries

This section is written in much more generality than the remainder of the paper. We present general results for an arbitrary locally profinite group.

Let $\mathtt{G}$ be a locally profinite group (i.e., a locally compact and totally disconnected topological group) and let R be a unitary commutative ring. We recall that a representation (π, V) of $\mathtt{G}$ over R is smooth if for every $v \in V$ the stabilizer $\{g \in \mathtt{G} \mid \pi(g)v = v\}$ is an open subgroup of $\mathtt{G}$. We denote by $\mathscr{R}_R(\mathtt{G})$ the (abelian) category of smooth representations of $\mathtt{G}$ over R. From now on all representations considered are smooth.

1A. ***Hecke algebras for a locally profinite group.*** In this section we introduce an algebra associated to a representation σ of a subgroup of $\mathtt{G}$ and we prove that it is isomorphic to the endomorphism algebra of the compact induction of σ. This definition generalizes those in Section 1 of [Chinello 2017] that corresponds to the case in which σ is trivial.

Let $\mathtt{H}$ be an open subgroup of $\mathtt{G}$ such that every $\mathtt{H}$-double coset is a finite union of left $\mathtt{H}$-cosets (or equivalently $\mathtt{H} \cap g\mathtt{H}g^{-1}$ is of finite index in $\mathtt{H}$ for every $g \in \mathtt{G}$) and let (σ, V_σ) be a smooth representation of $\mathtt{H}$ over R.

Definition 1.1. Let $\mathscr{H}_R(\mathtt{G}, \sigma)$ be the R-algebra of functions $\Phi : \mathtt{G} \to \mathrm{End}_R(V_\sigma)$ such that $\Phi(hgh') = \sigma(h) \circ \Phi(g) \circ \sigma(h')$ for every $h, h' \in \mathtt{H}$ and $g \in \mathtt{G}$ and whose supports are a finite union of $\mathtt{H}$-double cosets, endowed with convolution product

$$(\Phi_1 * \Phi_2)(g) = \sum_x \Phi_1(x)\Phi_2(x^{-1}g), \tag{1}$$

where x runs over a system of representatives of $\mathtt{G}/\mathtt{H}$ in $\mathtt{G}$. This algebra is unitary and the identity element is σ seen as a function on $\mathtt{G}$ with support equal to $\mathtt{H}$. To simplify the notation, from now on we denote $\Phi_1\Phi_1 = \Phi_1 * \Phi_2$ for all $\Phi_1, \Phi_2 \in \mathscr{H}_R(\mathtt{G},\sigma)$.

We observe that the sum in (1) is finite since the support of Φ_1 is a finite union of $\mathtt{H}$-double cosets and by hypothesis, every $\mathtt{H}$-double coset is a finite union of left $\mathtt{H}$-cosets. Furthermore, the formula (1) is well defined because for every $h \in \mathtt{H}$ and $x, g \in \mathtt{G}$ we have

$$\Phi_1(xh)\Phi_2((xh)^{-1}g) = \Phi_1(x)\circ\sigma(h)\circ\sigma(h^{-1})\circ\Phi_2(x^{-1}g) = \Phi_1(x)\circ\Phi_2(x^{-1}g).$$

For every $g \in \mathtt{G}$ we denote by $\mathscr{H}_R(\mathtt{G},\sigma)_{\mathtt{H}g\mathtt{H}}$ the submodule of $\mathscr{H}_R(\mathtt{G},\sigma)$ of functions with support in $\mathtt{H}g\mathtt{H}$. If $g_1, g_2 \in \mathtt{G}$, $\Phi_1 \in \mathscr{H}_R(\mathtt{G},\sigma)_{\mathtt{H}g_1\mathtt{H}}$ and $\Phi_2 \in \mathscr{H}_R(\mathtt{G},\sigma)_{\mathtt{H}g_2\mathtt{H}}$ then the support of $\Phi_1\Phi_2$ is in $\mathtt{H}g_1\mathtt{H}g_2\mathtt{H}$ and the support of $x \mapsto \Phi_1(x)\Phi_2(x^{-1}g)$ is in $\mathtt{H}g_1\mathtt{H} \cap g\mathtt{H}g_2^{-1}\mathtt{H}$.

Remark 1.2. If g_1 or g_2 normalizes $\mathtt{H}$ then the support of $\Phi_1\Phi_2$ is in $\mathtt{H}g_1g_2\mathtt{H}$ and the support of $x \mapsto \Phi_1(x)\Phi_2(x^{-1}g_1g_2)$ is in $g_1\mathtt{H}$. Hence, we obtain $(\Phi_1\Phi_2)(g_1g_2) = \Phi_1(g_1)\circ\Phi_2(g_2)$.

For every $g \in \mathtt{G}$ we denote by $\mathtt{H}^g = g^{-1}\mathtt{H}g$ and (σ^g, V_σ) the representation of $\mathtt{H}^g$ given by $\sigma^g(x) = \sigma(gxg^{-1})$ for every $x \in \mathtt{H}^g$. We denote by $I_g(\sigma)$ the R-module $\mathrm{Hom}_{\mathtt{H}\cap\mathtt{H}^g}(\sigma,\sigma^g)$ and $I_{\mathtt{G}}(\sigma)$ the set, called the *intertwining* of σ in $\mathtt{G}$, of $g \in \mathtt{G}$ such that $I_g(\sigma) \neq 0$. For every $g \in I_{\mathtt{G}}(\sigma)$ the map $\Phi \mapsto \Phi(g)$ is an isomorphism of R-modules between $\mathscr{H}_R(\mathtt{G},\sigma)_{\mathtt{H}g\mathtt{H}}$ and $I_g(\sigma)$ and so $g \in \mathtt{G}$ intertwines σ if and only if there exists an element $\Phi \in \mathscr{H}_R(\mathtt{G},\sigma)$ such that $\Phi(g) \neq 0$.

Let $\mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma)$ be the compactly induced representation of σ to $\mathtt{G}$. It is the R-module of functions $f : \mathtt{G} \to V_\sigma$, compactly supported modulo $\mathtt{H}$, such that $f(hg) = \sigma(h)f(g)$ for every $h \in \mathtt{H}$ and $g \in \mathtt{G}$ endowed with the action of $\mathtt{G}$ defined by $x.f : g \mapsto f(gx)$ for every $x, g \in \mathtt{G}$ and $f \in \mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma)$. We remark that, since $\mathtt{H}$ is open, by I.5.2(b) of [Vignéras 1996] it is a smooth representation of $\mathtt{G}$. For every $v \in V_\sigma$ let $i_v \in \mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma)$ be the function with support in $\mathtt{H}$ defined by $i_v(h) = \sigma(h)v$ for every $h \in \mathtt{H}$. Then for every $x \in \mathtt{G}$ the function $x^{-1}.i_v$ has support $\mathtt{H}x$ and takes the value v on x. Hence, for every $f \in \mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma)$ we have

$$f = \sum_{x\in\mathtt{H}\backslash\mathtt{G}} x^{-1}.i_{f(x)} \tag{2}$$

with the sum finite since the support of f is compact modulo H, and so the image i_{V_σ} of $v \mapsto i_v$ generates $\mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma)$ as representation of $\mathtt{G}$.

Frobenius reciprocity (I.5.7 of [Vignéras 1996]) states that the map $\mathrm{Hom}_{\mathtt{H}}(\sigma, V) \to \mathrm{Hom}_{\mathtt{G}}(\mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma), V)$ given by $\phi \mapsto \psi$ where $\phi(v) = \psi(i_v)$ for every $v \in V_\sigma$ is an isomorphism of R-modules.

Lemma 1.3. *If V_σ is a finitely generated R-module, the map $\xi : \mathscr{H}_R(\mathtt{G},\sigma) \to \mathrm{End}_{\mathtt{G}}(\mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma))$ given by*

$$\xi(\Phi)(f)(g) = (\Phi * f)(g) = \sum_{x\in\mathtt{G}/\mathtt{H}} \Phi(x)f(x^{-1}g)$$

for every $\Phi \in \mathscr{H}_R(\mathtt{G},\sigma)$, $f \in \mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma)$ and $g \in \mathtt{G}$ is an R-algebra isomorphism whose inverse is given by $\xi^{-1}(\vartheta)(g)(v) = \vartheta(i_v)(g)$ for every $\vartheta \in \mathrm{End}_{\mathtt{G}}(\mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\sigma))$, $g \in \mathtt{G}$ and $v \in V_\sigma$.

Proof. See I.8.5–6 of [Vignéras 1996]. □

1B. ***The categories $\mathscr{R}_\sigma(\mathtt{G})$ and $\mathscr{R}(\mathtt{G},\sigma)$.*** In this section we associate to an irreducible projective representation of a compact open subgroup of $\mathtt{G}$ two subcategories of $\mathscr{R}_R(\mathtt{G})$.

Let $\mathtt{K}$ be a compact open subgroup of $\mathtt{G}$ and (σ, V_σ) be an irreducible projective representation of $\mathtt{K}$ such that V_σ is a finitely generated R-module. Then $\rho = \mathrm{ind}_{\mathtt{K}}^{\mathtt{G}}(\sigma)$ is a projective representation of $\mathtt{G}$ by I.5.9(d) of [Vignéras 1996] and so the functor

$$\boldsymbol{M}_\sigma = \mathrm{Hom}_{\mathtt{G}}(\rho, -) : \mathscr{R}_R(\mathtt{G}) \to \text{Mod-}\mathscr{H}_R(\mathtt{G},\sigma)$$

is exact. We remark that for every representation (π, V) of $\mathtt{G}$ the right-action of $\Phi \in \mathscr{H}_R(\mathtt{G},\sigma)$ on $\varphi \in \mathrm{Hom}_{\mathtt{G}}(\rho, V)$ is given by $\varphi.\Phi = \varphi \circ \xi(\Phi)$ where ξ is the isomorphism of Lemma 1.3. Moreover, if V_1 and V_2 are representations of $\mathtt{G}$ and $\epsilon \in \mathrm{Hom}_{\mathtt{G}}(V_1, V_2)$ then $\boldsymbol{M}_\sigma(\phi)$ maps φ to $\phi \circ \varphi$ for every $\varphi \in \mathrm{Hom}_{\mathtt{G}}(\rho, V_1)$.

Definition 1.4. Let $\mathscr{R}_\sigma(\mathtt{G})$ be the full subcategory of $\mathscr{R}_R(\mathtt{G})$ whose objects are representations V such that $\boldsymbol{M}_\sigma(V') \neq 0$ for every irreducible subquotient V' of V.

For every representation V of $\mathtt{G}$ we denote by $V^\sigma = \sum_{\phi \in \mathrm{Hom}_{\mathtt{K}}(\sigma, V)} \phi(\sigma)$ which is a subrepresentation of the restriction of V to $\mathtt{K}$. We denote by $V[\sigma]$ the representation of $\mathtt{G}$ generated by V^σ. If σ is the trivial character of $\mathtt{K}$ then $V^\sigma = V^{\mathtt{K}} = \{v \in V \mid \pi(k)v = v \text{ for all } k \in \mathtt{K}\}$ is the set of $\mathtt{K}$-invariant vectors of V.

Proposition 1.5. *For every representation V of $\mathtt{G}$ we have $V[\sigma] = \sum_{\psi \in \boldsymbol{M}_\sigma(V)} \psi(\rho)$ and so $\boldsymbol{M}_\sigma(V) = \boldsymbol{M}_\sigma(V[\sigma])$. Moreover, if W is a subrepresentation of V then $\boldsymbol{M}_\sigma(W) = \boldsymbol{M}_\sigma(V)$ if and only if $W[\sigma] = V[\sigma]$.*

Proof. By Frobenius reciprocity we have $\mathrm{Hom}_{\mathtt{K}}(\sigma, V) \cong \boldsymbol{M}_\sigma(V)$ and so using (2) we obtain

$$V[\sigma] = \sum_{g \in \mathtt{G}} \pi(g) \sum_{\psi \in \boldsymbol{M}_\sigma(V)} \psi(i_{V_\sigma}) = \sum_{\psi \in \boldsymbol{M}_\sigma(V)} \psi\Big(\sum_{g \in \mathtt{G}} g.i_{V_\sigma}\Big) = \sum_{\psi \in \boldsymbol{M}_\sigma(V)} \psi(\rho),$$

which implies $\boldsymbol{M}_\sigma(V) = \boldsymbol{M}_\sigma(V[\sigma])$. Furthermore, if $W[\sigma] = V[\sigma]$ then $\boldsymbol{M}_\sigma(W) = \boldsymbol{M}_\sigma(V)$ and if $\boldsymbol{M}_\sigma(W) = \boldsymbol{M}_\sigma(V)$ then

$$W[\sigma] = \sum_{\psi \in \boldsymbol{M}_\sigma(W)} \psi(\rho) = \sum_{\psi \in \boldsymbol{M}_\sigma(V)} \psi(\rho) = V[\sigma]. \qquad \square$$

Definition 1.6. Let $\mathscr{R}(\mathtt{G},\sigma)$ be the full subcategory of $\mathscr{R}_R(\mathtt{G})$ whose objects are representations V such that $V = V[\sigma]$. If σ is the trivial character of $\mathtt{K}$ we denote by $\mathscr{R}(\mathtt{G},\mathtt{K})$ the subcategory of representations V generated by $V^{\mathtt{K}}$.

Proposition 1.7. *Let V be a representation of $\mathtt{G}$. The following conditions are equivalent*:

(i) *For every irreducible subquotient U of V we have $\boldsymbol{M}_\sigma(U) \neq 0$.*

(ii) *For every nonzero subquotient W of V we have $\boldsymbol{M}_\sigma(W) \neq 0$.*

(iii) *For every subquotient Z of V we have $Z = Z[\sigma]$.*

(iv) *For every subrepresentation Z of V we have $Z = Z[\sigma]$.*

Proof. (i)$\to$(ii): Let W be a nonzero subquotient of V and $W_1 \subset W_2$ two subrepresentations of W such that $U = W_2/W_1$ is irreducible. By (i) we have $\boldsymbol{M}_\sigma(U) \neq 0$ which implies $\boldsymbol{M}_\sigma(W_2) \neq 0$ and so $\boldsymbol{M}_\sigma(W) \neq 0$.

(ii)$\to$(iii): Let Z be a subquotient of V. By Proposition 1.5 we have $\boldsymbol{M}_\sigma(Z) = \boldsymbol{M}_\sigma(Z[\sigma])$ and so $\boldsymbol{M}_\sigma(Z/Z[\sigma]) = 0$. Hence, by (ii) we obtain $Z = Z[\sigma]$.

(iv)$\to$(i): Let U be an irreducible subquotient of V and $Z_1 \subsetneq Z_2$ be two subrepresentations of V such that $U = Z_2/Z_1$. By (iv) we have $Z_1[\sigma] = Z_1 \neq Z_2 = Z_2[\sigma]$ and by Proposition 1.5 we have $\boldsymbol{M}_\sigma(Z_1) \neq \boldsymbol{M}_\sigma(Z_2)$. Hence, we obtain $\boldsymbol{M}_\sigma(U) \neq 0$. $\square$

Remark 1.8. Proposition 1.7 implies that $\mathscr{R}_\sigma(\mathtt{G})$ is a subcategory of $\mathscr{R}(\mathtt{G}, \sigma)$.

1C. ***Equivalence of categories.*** In this section we suppose that there exists a compact open subgroup $\mathtt{K}_0$ of $\mathtt{G}$ whose pro-order is invertible in R and we consider the Haar measure $\mathtt{dg}$ on $\mathtt{G}$ with values in R such that $\int_{\mathtt{K}_0} \mathtt{dg} = 1$ (see I.2 of [Vignéras 1996]). We prove that if the two categories introduced in Section 1B are equal then they are equivalent to the category of modules over the algebra introduced in Section 1A.

The *global Hecke algebra* $\mathscr{H}_R(\mathtt{G})$ of $\mathtt{G}$ is the R-algebra of locally constant and compactly supported functions $f : \mathtt{G} \to R$ endowed with convolution product given by $(f_1 * f_2)(x) = \int_{\mathtt{G}} f_1(g) f_2(g^{-1}x)\, \mathtt{dg}$ for every $f_1, f_2 \in \mathscr{H}_R(\mathtt{G})$ and $x \in \mathtt{G}$ (see I.3.1 of [Vignéras 1996]). In general $\mathscr{H}_R(\mathtt{G})$ is not unitary but it has enough idempotents by I.3.2 of [loc. cit.]. The categories $\mathscr{R}_R(\mathtt{G})$ and $\mathscr{H}_R(\mathtt{G})$ -Mod are equivalent by I.4.4 of [loc. cit.] and we have $\mathrm{ind}_{\mathtt{H}}^{\mathtt{G}}(\tau) = \mathscr{H}_R(\mathtt{G}) \otimes_{\mathscr{H}_R(\mathtt{H})} V_\tau$ for every representation (τ, V_τ) of an open subgroup $\mathtt{H}$ of $\mathtt{G}$ by I.5.2 of [loc. cit.].

Let $\mathtt{K}$ be a compact open subgroup of $\mathtt{G}$, let (σ, V_σ) be an irreducible projective representation of $\mathtt{K}$ as in Section 1B and let $\rho = \mathrm{ind}_{\mathtt{K}}^{\mathtt{G}}(\sigma)$. Since V_σ is a simple projective module over the unitary algebra $\mathscr{H}_R(\mathtt{K})$, it is isomorphic to a direct summand of $\mathscr{H}_R(\mathtt{K})$ itself because any nonzero map $\mathscr{H}_R(\mathtt{K}) \to V_\sigma$ is surjective and splits. Then it is isomorphic to a minimal ideal of $\mathscr{H}_R(\mathtt{K})$ and so there exists an idempotent e of $\mathscr{H}_R(\mathtt{K})$ such that $V_\sigma = \mathscr{H}_R(\mathtt{K})e$. Hence, we obtain $\rho = \mathscr{H}_R(\mathtt{G})e$ because the map $\sum_i (f_i \otimes h_i e) \mapsto \left(\sum_i f_i h_i\right)e$ is an isomorphism of $\mathscr{H}_R(\mathtt{G})$-modules between $\mathscr{H}_R(\mathtt{G}) \otimes_{\mathscr{H}_R(\mathtt{K})} \mathscr{H}_R(\mathtt{K})e$ and $\mathscr{H}_R(\mathtt{G})e$ whose inverse is $fe \mapsto fe \otimes e$.

The algebra $\mathscr{H}_R(\mathtt{G}, \sigma)$ is isomorphic to $\mathrm{End}_{\mathtt{G}}(\rho) \cong \mathrm{End}_{\mathscr{H}_R(\mathtt{G})}(\mathscr{H}_R(\mathtt{G})e)$ by Lemma 1.3 and the map $e\mathscr{H}_R(\mathtt{G})e \to (\mathrm{End}_{\mathscr{H}_R(\mathtt{G})}(\mathscr{H}_R(\mathtt{G})e))^{\mathrm{op}}$ which maps $efe \in e\mathscr{H}_R(\mathtt{G})e$ to the endomorphism $f'e \mapsto f'efe$ of $\mathscr{H}_R(\mathtt{G})e$ is an algebra isomorphism whose inverse is $\varphi \mapsto \varphi(e)$. Then we have $\mathscr{H}_R(\mathtt{G}, \sigma)^{\mathrm{op}} \cong e\mathscr{H}_R(\mathtt{G})e$ and so the categories $e\mathscr{H}_R(\mathtt{G})e$ -Mod and Mod- $\mathscr{H}_R(\mathtt{G}, \sigma)$ are equivalent.

Theorem 1.9. *If $\mathscr{R}_\sigma(\mathtt{G}) = \mathscr{R}(\mathtt{G}, \sigma)$ then $V \mapsto \boldsymbol{M}_\sigma(V)$ is an equivalence of categories between $\mathscr{R}(\mathtt{G}, \sigma)$ and* Mod- $\mathscr{H}_R(\mathtt{G}, \sigma)$ *whose quasiinverse is $W \mapsto W \otimes_{\mathscr{H}_R(\mathtt{G},\sigma)} \rho$.*

Proof. We take $A = \mathscr{H}_R(\mathtt{G})$ and $\mathscr{H}_R(\mathtt{G})e = \rho$ as in I.6.6 of [Vignéras 1996]. Since $\mathscr{H}_R(\mathtt{G}, \sigma)^{\mathrm{op}} \cong e\mathscr{H}_R(\mathtt{G})e$, left-actions of $e\mathscr{H}_R(\mathtt{G})e$ become right-actions of $\mathscr{H}_R(\mathtt{G}, \sigma)$. The functor $V \mapsto eV$ of [loc. cit.] from $\mathscr{H}_R(\mathtt{G})$ -Mod to $e\mathscr{H}_R(\mathtt{G})e$ -Mod becomes the functor $V \mapsto \mathrm{Hom}_{\mathscr{H}_R(\mathtt{G})}(\mathscr{H}_R(\mathtt{G})e, V)$ and so the functor $\boldsymbol{M}_\sigma$. The hypotheses of the theorem "*équivalence de catégories*" in I.6.6 of [Vignéras 1996] are satisfied by the condition $\mathscr{R}_\sigma(\mathtt{G}) = \mathscr{R}(\mathtt{G}, \sigma)$ and so we obtain the result. $\square$

2. Maximal simple types

In this section we introduce the theory of simple types of an inner form of a general linear group over a nonarchimedean locally compact field in the case of modular representations. We refer to Sections 2.1–5 of [Mínguez and Sécherre 2014b] for more details.

Let p be a prime number and let F be a nonarchimedean locally compact field of residue characteristic p. For F' a finite extension of F, or more generally a division algebra over a finite extension of F, we denote by $\mathcal{O}_{F'}$ its ring of integers, by $\varpi_{F'}$ a uniformizer of $\mathcal{O}_{F'}$, by $\wp_{F'}$ the maximal ideal of $\mathcal{O}_{F'}$ and by $\mathfrak{k}_{F'}$ its residue field. Let D be a central division algebra of finite dimension over F whose reduced degree is denoted by d. Given a positive integer m, we consider the ring $A = M_m(D)$ and the group $G = \mathrm{GL}_m(D)$ which is an inner form of $\mathrm{GL}_{md}(F)$. Let R be an algebraically closed field of characteristic different from p.

Let Λ be an $\mathcal{O}_D$-lattice sequence of $V = D^m$. It defines a hereditary $\mathcal{O}_F$-order $\mathfrak{A} = \mathfrak{A}(\Lambda)$ of A whose radical is denoted by $\mathfrak{P}$, a compact open subgroup $U(\Lambda) = U_0(\Lambda) = \mathfrak{A}(\Lambda)^\times$ of G and a filtration $U_k(\Lambda) = 1 + \mathfrak{P}^k$ with $k \geq 1$ of $U(\Lambda)$ (see Section 1 of [Sécherre 2004]). Let $[\Lambda, n, 0, \beta]$ be a simple stratum of A (see for instance Section 1.6 of [Sécherre and Stevens 2008]). Then $\beta \in A$ and the F-subalgebra $F[\beta]$ of A generated by β is a field denoted by E. The centralizer B of E in A is a simple central E-algebra and $\mathfrak{B} = \mathfrak{A} \cap B$ is a hereditary $\mathcal{O}_E$-order of B whose radical is $\mathfrak{Q} = \mathfrak{P} \cap B$.

As in Sections 1.2 and 1.3 of [Sécherre 2005b] we can choose a simple right $E \otimes_F D$-module N such that the functor $V \mapsto \mathrm{Hom}_{E\otimes_F D}(N, V)$ defines a Morita equivalence between the category of modules over $E \otimes_F D$ and the category of vector spaces over $D' = \mathrm{End}_{E\otimes_F D}(N)^{\mathrm{op}}$ which is a central division algebra over E. We set $A(E) = \mathrm{End}_D(N)$ which is a central simple F-algebra. If d' is the reduced degree of D' over E and m' is the dimension of $V' = \mathrm{Hom}_{E\otimes_F D}(N, V)$ over D', then we have $m'd' = md/[E:F]$. Fixing a basis of V' over D' we obtain, via the Morita equivalence above, an isomorphism $N^{m'} \cong V$ of $E \otimes_F D$-modules. If for every $i \in \{1, \dots, m'\}$ we denote by V^i the image of the i-th copy of N by this isomorphism, we obtain a decomposition $V = V^1 \oplus \cdots \oplus V^{m'}$ into simple $E \otimes_F D$-submodules. By Section 1.5 of [Sécherre 2005b] we can choose a basis $\mathscr{B}$ of V' over D' so that Λ decomposes as the direct sum of the $\Lambda^i = \Lambda \cap V^i$ for $i \in \{1, \dots, m'\}$. For every $i \in \{1, \dots, m'\}$, let $\mathrm{e}_i : V \to V^i$ be the projection on V^i with kernel $\bigoplus_{j \neq i} V^j$. In accordance with [Sécherre 2004, 2.3.1] (see also [Bushnell and Henniart 1996]) the family of idempotents $\mathrm{e} = (\mathrm{e}_1, \dots, \mathrm{e}_{m'})$ is a decomposition which conforms to Λ over E.

By 1.4.8 and 1.5.2 of [Sécherre 2005b] there exists a unique hereditary order $\mathfrak{A}(E)$ normalized by $E^\times$ in $A(E)$ whose radical is denoted by $\mathfrak{P}(E)$. For every $i \in \{1, \dots, m'\}$ we have an isomorphism $\mathrm{End}_D(V^i) \cong A(E)$ of F-algebras which induces an isomorphism of $\mathcal{O}_F$-algebras between the hereditary orders $\mathfrak{A}(\Lambda^i)$ and $\mathfrak{A}(E)$. Moreover, to the choice of the basis $\mathscr{B}$ corresponds the isomorphisms $M_{m'}(D') \cong B$ of E-algebras and $M_{m'}(A(E)) \cong A$ of F-algebras.

Remark 2.1. If $U(\Lambda) \cap B^\times$ is a maximal compact open subgroup of $B^\times$, these isomorphisms induce an isomorphism $\mathfrak{B} \cong M_{m'}(\mathcal{O}_{D'})$ of $\mathcal{O}_E$-algebras and, by Lemma 1.6 of [Sécherre 2005a], two isomorphisms $\mathfrak{A} \cong M_{m'}(\mathfrak{A}(E))$ and $\mathfrak{P} \cong M_{m'}(\mathfrak{P}(E))$ of $\mathcal{O}_F$-algebras.

We can associate to $[\Lambda, n, 0, \beta]$ two compact open subgroups $J = J(\beta, \Lambda)$, $H = H(\beta, \Lambda)$ of $U(\Lambda)$ (see 2.4 of [Sécherre and Stevens 2008]). For every integer $k \geq 1$ we set $J^k = J^k(\beta, \Lambda) = J(\beta, \Lambda) \cap U_k(\Lambda)$ and $H^k = H^k(\beta, \Lambda) = H(\beta, \Lambda) \cap U_k(\Lambda)$ which are pro-p-groups. In particular J^1 and H^1 are normal pro-p-subgroups of J and the quotient J^1/H^1 is a finite abelian p-group.

Remark 2.2. We have $J = (U(\Lambda) \cap B^\times)J^1$ and this induce a canonical group isomorphism

$$J/J^1 \cong (U(\Lambda) \cap B^\times)/(U_1(\Lambda) \cap B^\times)$$

(see Section 2.3 of [Mínguez and Sécherre 2014b]). It allows us to associate canonically and bijectively a representation of J trivial on J^1 to a representation of $U(\Lambda) \cap B^\times$ trivial on $U_1(\Lambda) \cap B^\times$.

2A. ***Simple characters, Heisenberg representation and β-extensions.*** Let $[\Lambda, n, 0, \beta]$ be a simple stratum of A. We denote by $\mathscr{C}_R(\Lambda, 0, \beta)$ the set of *simple R-characters* (see Section 2.2 of [Mínguez and Sécherre 2014b] and [Sécherre 2004]) that is a finite set of R-characters of H^1 which depends on the choice of an additive R-character of F which has been fixed once and for all. If $\tilde{m} \in \mathbb{N}^*$ and $[\tilde{\Lambda}, \tilde{n}, 0, \tilde{\beta}]$ is a simple stratum of $M_{\tilde{m}}(D)$ such that there exists an isomorphism of F-algebras $\nu : F[\beta] \to F[\tilde{\beta}]$ with $\nu(\beta) = \tilde{\beta}$, then there exists a bijection $\mathscr{C}_R(\Lambda, 0, \beta) \to \mathscr{C}_R(\tilde{\Lambda}, 0, \tilde{\beta})$ canonically associated to ν, called the *transfer map*. There also exists an equivalence relation, called *endoequivalence*, among simple characters in $\mathscr{C}_R(\Lambda, 0, \beta)$ (see [Broussous et al. 2012]) such that two of them are endoequivalent if they have transfers which intertwine. The equivalence classes of this relation are called *endoclasses*. Let $\theta \in \mathscr{C}_R(\Lambda, 0, \beta)$. By Proposition 2.1 of [Mínguez and Sécherre 2014b] there exists a finite dimensional irreducible representation η of J^1, unique up to isomorphism, whose restriction to H^1 contains θ. It is called the *Heisenberg representation* associated to θ. The intertwining of η is $I_G(\eta) = J^1 B^\times J^1 = J B^\times J$ and for every $y \in B^\times$ the R-vector space $I_y(\eta) = \mathrm{Hom}_{J^1 \cap (J^1)^y}(\eta, \eta^y)$ has dimension 1.

A *β-extension* of η (or of θ) is an irreducible representation κ of J extending η such that $I_G(\kappa) = J B^\times J$. By Proposition 2.4 of [Mínguez and Sécherre 2014b], every simple character $\theta \in \mathscr{C}_R(\Lambda, 0, \beta)$ admits a β-extension κ and by formula (2.2) of [Mínguez and Sécherre 2014b] the set of β-extensions of θ is equal to

$$\mathcal{B}(\theta) = \{\kappa \otimes (\chi \circ N_{B/E}) \mid \chi \text{ is a character of } \mathcal{O}_E^\times, \text{ trivial on } 1 + \wp_E\},$$

where $N_{B/E}$ is the reduced norm of B over E and $\chi \circ N_{B/E}$ is seen as a character of J trivial on J^1 thanks to Remark 2.2. We observe that for every $\kappa \in \mathcal{B}(\theta)$ and every $y \in B^\times$, the R-vector space $I_y(\kappa)$ has dimension 1 because it is nonzero and it is contained in $I_y(\eta)$.

2B. ***Maximal simple types.*** Let $[\Lambda, n, 0, \beta]$ be a simple stratum of A such that $U(\Lambda) \cap B^\times$ is a maximal compact open subgroup of $B^\times$. By Remarks 2.1 and 2.2, there exists a group isomorphism $J/J^1 \cong \mathrm{GL}_{m'}(\mathfrak{k}_{D'})$, which depends on the choice of $\mathscr{B}$.

A *maximal simple type* of G associated to $[\Lambda, n, 0, \beta]$ is a pair (J, λ) where λ is an irreducible representation of J of the form $\lambda = \kappa \otimes \sigma$ where $\kappa \in \mathcal{B}(\theta)$ with $\theta \in \mathscr{C}_R(\Lambda, 0, \beta)$ and σ is a cuspidal representation of $\mathrm{GL}_{m'}(\mathfrak{k}_{D'})$ identified with an irreducible representation of J trivial on J^1. If σ is a supercuspidal representation of $\mathrm{GL}_{m'}(\mathfrak{k}_{D'})$ then (J, λ) is called maximal simple *supertype*.

Remark 2.3. The choice of a β-extension $\kappa \in \mathcal{B}(\theta)$ determines the decomposition $\lambda = \kappa \otimes \sigma$. If we choose another β-extension $\kappa' = \kappa \otimes (\chi \circ N_{B/E}) \in \mathcal{B}(\theta)$ we obtain the decomposition $\lambda = \kappa' \otimes \sigma'$ where $\sigma' = \sigma \otimes (\chi^{-1} \circ N_{B/E})$.

2C. ***Covers.*** Let $\mathcal{M}$ be a Levi subgroup of G, let $\mathcal{P}$ be a parabolic subgroup of G with Levi component $\mathcal{M}$ and unipotent radical $\mathcal{U}$ and let $\mathcal{U}^-$ be the unipotent subgroup opposite to $\mathcal{U}$. We say that a compact open subgroup K of G is *decomposed with respect to* $(\mathcal{M}, \mathcal{P})$ if every element $k \in K$ decomposes uniquely as $k = k_1k_2k_3$ with $k_1 \in K \cap \mathcal{U}^-$, $k_2 \in K \cap \mathcal{M}$ and $k_3 \in K \cap \mathcal{U}$. Furthermore, if π is a representation of K we say that the pair (K, π) is *decomposed with respect to* $(\mathcal{M}, \mathcal{P})$ if K is decomposed with respect to $(\mathcal{M}, \mathcal{P})$ and if $K \cap \mathcal{U}$ and $K \cap \mathcal{U}^-$ are in the kernel of π.

Let $\mathcal{M}$ be a Levi subgroup of G. Let K and $K_{\mathcal{M}}$ be two compact open subgroups of G and $\mathcal{M}$ respectively and let ϱ and $\varrho_{\mathcal{M}}$ be two irreducible representations of K and $K_{\mathcal{M}}$ respectively. We say that the pair (K, ϱ) is *decomposed above* $(K_{\mathcal{M}}, \varrho_{\mathcal{M}})$ if (K, ϱ) is decomposed with respect to $(\mathcal{M}, \mathcal{P})$ for every parabolic subgroup $\mathcal{P}$ with Levi component $\mathcal{M}$, if $K \cap \mathcal{M} = K_{\mathcal{M}}$ and if the restriction of ϱ to $K_{\mathcal{M}}$ is equal to $\varrho_{\mathcal{M}}$. For a parabolic subgroup $\mathcal{P}$ of G with Levi component $\mathcal{M}$ and unipotent radical $\mathcal{U}$, let $\varrho_{\mathcal{U}}$ be the Jacquet module of ϱ and $r_{\mathcal{U}}$ be the canonical quotient map $\varrho \to \varrho_{\mathcal{U}}$. A pair (K, ϱ) is a *cover* of $(K_{\mathcal{M}}, \varrho_{\mathcal{M}})$ if it is decomposed above $(K_{\mathcal{M}}, \varrho_{\mathcal{M}})$ and if for every irreducible representations π of G the map $\mathrm{Hom}_K(\varrho, \pi) \to \mathrm{Hom}_{K_{\mathcal{M}}}(\varrho_{\mathcal{M}}, \pi_{\mathcal{U}})$, given by $\varphi \mapsto r_{\mathcal{U}} \circ \varphi$ for every $\varphi \in \mathrm{Hom}_K(\varrho, \pi)$, is injective (see Condition (0.5) of [Blondel 2005]). For more details see [Blondel 2005; Vignéras 1998].

3. The isomorphisms $\mathscr{H}_R(G, \eta) \cong \mathscr{H}_R(B^\times, U_1(\Lambda) \cap B^\times)$

Using the notation of Section 2, let $[\Lambda, n, 0, \beta]$ be a simple stratum of A such that $U(\Lambda) \cap B^\times$ is a maximal compact open subgroup of $B^\times$. Let $\theta \in \mathscr{C}_R(\Lambda, 0, \beta)$ and let η be the Heisenberg representation associated to θ. In this section we want to prove that the algebras $\mathscr{H}_R(G, \eta)$ and $\mathscr{H}_R(B^\times, U_1(\Lambda) \cap B^\times)$ are isomorphic (Theorem 3.43).

Henceforth, for a given $m \in \mathbb{N}$, we denote by $\mathbb{I}_m$ the identity matrix of size m. Thanks to Section 2, from now on we identify A with $M_{m'}(A(E))$, G with $\mathrm{GL}_{m'}(A(E))$, $U(\Lambda)$ with $\mathrm{GL}_{m'}(\mathfrak{A}(E))$, $U_1(\Lambda)$ with $\mathbb{I}_{m'} + M_{m'}(\mathfrak{P}(E))$, $B^\times$ with $\mathrm{GL}_{m'}(D')$, $K_B = U(\Lambda) \cap B^\times$ with $\mathrm{GL}_{m'}(\mathcal{O}_{D'})$ and $K_B^1 = U_1(\Lambda) \cap B^\times$ with $\mathbb{I}_{m'} + M_{m'}(\wp_{D'})$. By Section 2.4 of [Chinello 2017] we know a presentation by generators and relations of the algebra $\mathscr{H}_R(B^\times, K_B^1) \cong \mathscr{H}_{\mathbb{Z}}(B^\times, K_B^1) \otimes_{\mathbb{Z}} R$. Using this presentation we want to find an isomorphism between $\mathscr{H}_R(B^\times, K_B^1)$ and $\mathscr{H}_R(G, \eta)$.

3A. ***Root system of*** $\mathbf{GL}_{m'}$**.** In this section we recall some notation and results on the root system of $\mathrm{GL}_{m'}$ contained in Section 2.1 of [Chinello 2017].

We denote by $\mathbf{\Phi} = \{\alpha_{ij} \mid 1 \le i \ne j \le m'\}$ the set of roots of $\mathrm{GL}_{m'}$ relative to the torus of diagonal matrices. Let $\mathbf{\Phi}^+ = \{\alpha_{ij} \mid 1 \le i < j \le m'\}$, $\mathbf{\Phi}^- = -\mathbf{\Phi}^+ = \{\alpha_{ij} \mid 1 \le j < i \le m'\}$ and $\Sigma = \{\alpha_{i,i+1} \mid 1 \le i \le m' - 1\}$

be, respectively, the sets of positive, negative and simple roots relative to the Borel subgroup of upper triangular matrices. For every $\alpha = \alpha_{i,i+1} \in \Sigma$ we write s_α or s_i for the transposition $(i, i+1)$. Let W be the group generated by the s_i which is the group of permutations of m' elements and so the Weyl group of $\mathrm{GL}_{m'}$. Let $\ell : W \to \mathbb{N}$ be the length function of W relative to $s_1, \dots, s_{m'-1}$. The group W acts on $\mathbf{\Phi}$ by $w\alpha_{ij} = \alpha_{w(i)w(j)}$ and for every $w \in W$ and $\alpha \in \Sigma$ we have (see (2.2) of [loc. cit.])

$$\ell(ws_\alpha) = \begin{cases} \ell(w) + 1 & \text{if } w\alpha \in \mathbf{\Phi}^+, \\ \ell(w) - 1 & \text{if } w\alpha \in \mathbf{\Phi}^-. \end{cases} \tag{3}$$

Remark 3.1. By Proposition 2.2 of [loc. cit.] we have $\ell(w) = |\mathbf{\Phi}^+ \cap w\mathbf{\Phi}^-| = |\mathbf{\Phi}^- \cap w\mathbf{\Phi}^+|$.

For every $P \subset \Sigma$ we denote by $\mathbf{\Phi}_P^+$ the set of positive roots generated by P, $\mathbf{\Phi}_P^- = -\mathbf{\Phi}_P^+$, $\Psi_P^+ = \mathbf{\Phi}^+ \setminus \mathbf{\Phi}_P^+$ and $\Psi_P^- = -\Psi_P^+$. We denote by W_P the subgroup of W generated by the s_α with $\alpha \in P$ and by $\hat{P}$ the complement of P in Σ. We abbreviate $\hat{\alpha} = \widehat{\{\alpha\}}$.

Example. If $\alpha = \alpha_{i,i+1}$ then $\hat{\alpha} = \{\alpha_{j,j+1} \in \Sigma \mid j \neq i\}$, $\mathbf{\Psi}_{\hat{\alpha}}^+ = \{\alpha_{hk} \in \mathbf{\Phi}^+ \mid 1 \le h \le i < k \le m'\}$ and $\mathbf{\Phi}_{\hat{\alpha}}^+ = \{\alpha_{hk} \in \mathbf{\Phi}^+ \mid 1 \le h < k \le i \text{ or } i+1 \le h < k \le m'\}$.

Proposition 3.2. *Let $P \subset \Sigma$ and let w be an element of minimal length in $wW_P \in W/W_P$. Then $w\alpha \in \mathbf{\Phi}^+$ for every $\alpha \in \mathbf{\Phi}_P^+$ and for every $w' \in W_P$ we have $\ell(ww') = \ell(w) + \ell(w')$.*

Proof. Proposition 2.4 and Lemma 2.5 of [Chinello 2017]. □

Proposition 3.2 implies that in each class of W/W_P with $P \subset \Sigma$, there exists a unique element of minimal length and the same holds in each class of $W_P \backslash W$.

If ϖ is a uniformizer of $\mathcal{O}_{D'}$ we identify $\tau_i = \left(\begin{smallmatrix} \mathbb{I}_i & 0 \\ 0 & \varpi \mathbb{I}_{m'-i} \end{smallmatrix}\right)$ with $i \in \{0, \dots, m'\}$, defined in Section 2.2 of [loc. cit.], with elements of $B^\times$ and then of G. For $\alpha = \alpha_{i,i+1} \in \Sigma$ we write $\tau_\alpha = \tau_i$. Let $\mathbf{\Delta}$ and $\hat{\Delta}$ be the commutative monoid and group, respectively, generated by τ_α with $\alpha \in \Sigma$. Then we can write every element τ of $\mathbf{\Delta}$ uniquely as $\tau = \prod_{\alpha \in \Sigma} \tau_\alpha^{i_\alpha}$ with i_α in $\mathbb{N}$ and uniquely as $\tau = \mathrm{diag}(1, \varpi^{a_1}, \dots, \varpi^{a_{m-1}})$ with $0 \le a_1 \le \cdots \le a_{m-1}$. In this case we set $P(\tau) = \{\alpha \in \Sigma \mid i_\alpha = 0\}$ and if $P \subset \{0, \dots, m'\}$ or if $P \subset \Sigma$ we write τ_P in place of $\prod_{x \in P} \tau_x$. We remark that if $P \subset \Sigma$ then $P(\tau_P) = \hat{P}$.

3B. *The representation $\eta_{\mathcal{P}}$.* Let $\mathcal{M} = A(E)^\times \times \cdots \times A(E)^\times$ (m' copies) which is a Levi subgroup of G and let $\mathcal{P}$ be the parabolic subgroup of G of upper triangular matrices with Levi component $\mathcal{M}$ and unipotent radical $\mathcal{U}$. Let $\mathcal{P}^-$ be the opposite parabolic subgroup of $\mathcal{P}$ and $\mathcal{U}^-$ its unipotent radical.

We write $U = K_B \cap \mathcal{U}$, $M = K_B \cap \mathcal{M}$ and $I_B = K_B^1 MU$. Then U is the group of unipotent upper triangular matrices with coefficients in $\mathcal{O}_{D'}$, M is the group of diagonal matrices with coefficients in $\mathcal{O}_{D'}^\times$ and I_B is the standard Iwahori subgroup of K_B.

We denote by $\tilde{W}$ the group $W \ltimes \hat{\Delta}$ of monomial matrices with coefficients in $\varpi^{\mathbb{Z}}$ which is called the *extended affine Weyl group of $B^\times$*. We recall that $B^\times = I_B \tilde{W} I_B$ and actually it is the disjoint union of $I_B \tilde{w} I_B$ with $\tilde{w} \in \tilde{W}$.

Remark 3.3. By Proposition 2.16 of [Sécherre 2005a], which works for every decomposition that conforms to Λ over E and not necessarily subordinate to $\mathfrak{B}$, the groups J^1 and H^1 are decomposed with respect to $(\mathcal{M}, \mathcal{P})$. Moreover, if $\mathcal{M}' = \prod_{i=1}^{r} \mathrm{GL}_{m_i'}(A(E))$ with $\sum_{i=1}^{r} m_i' = m'$ is a standard Levi subgroup of G containing $\mathcal{M}$ and $\mathcal{P}'$ is the upper standard parabolic subgroup of G with Levi component $\mathcal{M}'$, then J^1 and H^1 are decomposed with respect to $(\mathcal{M}', \mathcal{P}')$.

Let $\mathfrak{J}^1 = \mathfrak{J}^1(\beta, \Lambda)$ and $\mathfrak{H}^1 = \mathfrak{H}^1(\beta, \Lambda)$ be the $\mathcal{O}_F$-lattices of A such that $J^1 = 1 + \mathfrak{J}^1$ and $H^1 = 1 + \mathfrak{H}^1$ (see Section 3.3 of [Sécherre 2004] or Chapter 3 of [Bushnell and Kutzko 1993]). Then they are $(\mathfrak{B}, \mathfrak{B})$-bimodules and we have $\varpi \mathfrak{J}^1 \subset \mathfrak{H}^1 \subset \mathfrak{J}^1 \subset M_{m'}(\mathfrak{P}(E))$.

Since $V^i \cong N$ for every $i \in \{1, \dots, m'\}$, we can identify every Λ^i to a lattice sequence Λ_0 of N with the same period as Λ, every $e^i\beta$ to an element $\beta_0 \in A(E)$ and $\mathfrak{A}(\Lambda_0)$ to $\mathfrak{A}(E)$. By Proposition 2.28 of [Sécherre 2004] the stratum $[\Lambda_0, n, 0, \beta_0]$ of $A(E)$ is simple and the critical exponents $k_0(\beta, \Lambda)$ and $k_0(\beta_0, \Lambda_0)$ are equal (for a definition of the critical exponent see Section 2.1 of [Sécherre 2004]). This implies that β is minimal (i.e., $-k_0(\beta, \Lambda) = n$) if and only if β_0 is minimal. We write $\mathfrak{J}_0^1 = \mathfrak{J}^1(\beta_0, \Lambda_0)$, $\mathfrak{H}_0^1 = \mathfrak{H}^1(\beta_0, \Lambda_0)$, $J_0^1 = J^1(\beta_0, \Lambda_0) = 1 + \mathfrak{J}_0^1$ and $H_0^1 = H^1(\beta_0, \Lambda_0) = 1 + \mathfrak{H}_0^1$.

Proposition 3.4. *We have* $\mathfrak{J}^1 = M_{m'}(\mathfrak{J}_0^1)$ *and* $\mathfrak{H}^1 = M_{m'}(\mathfrak{H}_0^1)$.

Proof. We prove the result only for $\mathfrak{J}^1$ since the case of $\mathfrak{H}^1$ is similar. We have to prove that for every $i, j \in \{1, \dots, m'\}$ we have $e^i \mathfrak{J}^1 e^j = \mathfrak{J}_0^1$. We need to recall the definition of $\mathfrak{J}(\beta, \Lambda) = \mathfrak{J}^0(\beta, \Lambda)$ and of $\mathfrak{J}^k(\beta, \Lambda)$ with $k \geq 1$. By Proposition 3.42 of [Sécherre 2004] if we set $q = -k_0(\beta, \Lambda)$ and $s = [(q+1)/2]$ (where $[x]$ denotes the integer part of $x \in \mathbb{Q}$) we have $\mathfrak{J}(\beta, \Lambda) = \mathfrak{B} + \mathfrak{P}^s$ if β is minimal and $\mathfrak{J}(\beta, \Lambda) = \mathfrak{B} + \mathfrak{J}^s(\gamma, \Lambda)$ if $[\Lambda, n, q, \gamma]$ is a simple stratum equivalent to $[\Lambda, n, q, \beta]$. Then, if β is minimal, $\mathfrak{J}^k(\beta, \Lambda) = \mathfrak{J}(\beta, \Lambda) \cap \mathfrak{P}^k$ is equal to $\mathfrak{Q}^k + \mathfrak{P}^s$ if $0 \leq k \leq s-1$ and to $\mathfrak{P}^k$ if $k \geq s$. Otherwise, if $[\Lambda, n, q, \gamma]$ is a simple stratum equivalent to $[\Lambda, n, q, \beta]$, $\mathfrak{J}^k(\beta, \Lambda)$ is equal to $\mathfrak{Q}^k + \mathfrak{J}^s(\gamma, \Lambda)$ if $0 \leq k \leq s-1$ and to $\mathfrak{J}^k(\gamma, \Lambda)$ if $k \geq s$. Similarly we obtain that if β_0 is minimal then $\mathfrak{J}^k(\beta_0, \Lambda_0)$ is equal to $\wp_{D'}^k + \mathfrak{P}(E)^s$ if $0 \leq k \leq s-1$ and to $\mathfrak{P}(E)^k$ if $k \geq s$. Otherwise, if $[\Lambda_0, n, q, \gamma_0]$ is a simple stratum equivalent to $[\Lambda_0, n, q, \beta_0]$, $\mathfrak{J}^k(\beta_0, \Lambda_0)$ is equal to $\wp_{D'}^k + \mathfrak{J}^s(\gamma_0, \Lambda_0)$ if $k \leq s-1$ and to $\mathfrak{J}^k(\gamma_0, \Lambda_0)$ if $k \geq s$. We prove that $e^i \mathfrak{J}^k(\beta, \Lambda) e^j = \mathfrak{J}^k(\beta_0, \Lambda_0)$ for every $k \geq 0$ by induction on q. If $q = n$ and so if β and β_0 are minimal, since $\mathfrak{Q} = M_{m'}(\wp_{D'})$ and $\mathfrak{P} = M_{m'}(\mathfrak{P}(E))$ we have $e^i \mathfrak{Q}^k e^j = \wp_{D'}^k$ and $e^i \mathfrak{P}^k e^j = \mathfrak{P}(E)^k$ for every k and so $e^i \mathfrak{J}^k(\beta, \Lambda) e^j = \mathfrak{J}^k(\beta_0, \Lambda_0)$ for every $k \geq 0$. Now if $q < n$ and so if β and β_0 are not minimal, by Proposition 1.20 of [Sécherre and Stevens 2008] (see also the proof of Theorem 2.2 of [Sécherre 2005b]) we can choose a simple stratum $[\Lambda_0, n, q, \gamma_0]$ equivalent to $[\Lambda_0, n, q, \beta_0]$ such that if γ is the image of γ_0 by the diagonal embedding $A(E) \to A$ then $[\Lambda, n, q, \gamma]$ is a simple stratum equivalent to $[\Lambda, n, q, \beta]$. By the inductive hypothesis we have $e^i \mathfrak{J}^k(\gamma, \Lambda) e^j = \mathfrak{J}^k(\gamma_0, \Lambda_0)$ for every $k \geq 0$ and then we obtain $e^i \mathfrak{J}^k(\beta, \Lambda) e^j = \mathfrak{J}^k(\beta_0, \Lambda_0)$. $\square$

Let θ_0 be the transfer of θ to $\mathscr{C}_R(\Lambda_0, 0, \beta)$. Since H^1 is a pro-p-group, proceeding as in Proposition 2.16 of [Sécherre 2005a], the pair (H^1, θ) is decomposed with respect to $(\mathcal{M}, \mathcal{P})$ and the restriction of θ to $H^1 \cap \mathcal{M} = H_0^1 \times \cdots \times H_0^1$ is $\theta_0^{\otimes m'}$. We remark that in general (J^1, η) is not decomposed with respect to $(\mathcal{M}, \mathcal{P})$. We denote by η_0 the Heisenberg representation of θ_0 and we can consider the irreducible representation $\eta_{\mathcal{M}} = \eta_0^{\otimes m'}$ of $J_{\mathcal{M}}^1 = J^1 \cap \mathcal{M} = J_0^1 \times \cdots \times J_0^1$.

We put $J^1_{\mathcal{P}} = (J^1 \cap \mathcal{P})H^1$ and $H^1_{\mathcal{P}} = (J^1 \cap \mathcal{U})H^1$ which are subgroups of J^1. They are normal in J^1 because H^1 contains the derived group of J^1. Moreover, $J \cap \mathcal{P}$ normalizes $J^1_{\mathcal{P}}$ because H^1 is normal in J and $J^1 \cap \mathcal{P}$ is normal in $J \cap \mathcal{P}$. Then $J^1_{\mathcal{P}}$ is normal in $J^1(J \cap \mathcal{P})$.

Remark 3.5. Taking into account Remark 5.7 of [Sécherre and Stevens 2008], Proposition 5.3 of [Sécherre and Stevens 2008] states that $J^1_{\mathcal{P}}$ and $H^1_{\mathcal{P}}$ are decomposed with respect to $(\mathcal{M}, \mathcal{P})$ and so we have $J^1_{\mathcal{P}} = (H^1 \cap \mathcal{U}^-)J^1_{\mathcal{M}}(J^1 \cap \mathcal{U})$ and $H^1_{\mathcal{P}} = (H^1 \cap \mathcal{U}^-)(H^1 \cap \mathcal{M})(J^1 \cap \mathcal{U})$. Moreover, if $\mathcal{M}' = \prod_{i=1}^{r} \mathrm{GL}_{m'_i}(A(E))$ with $\sum_{i=1}^{r} m'_i = m'$ is a standard Levi subgroup of G containing $\mathcal{M}$ and $\mathcal{P}'$ is the upper standard parabolic subgroup of G with Levi component $\mathcal{M}'$, then $J^1_{\mathcal{P}}$ and $H^1_{\mathcal{P}}$ are decomposed with respect to $(\mathcal{M}', \mathcal{P}')$.

Let $\theta_{\mathcal{P}}$ be the character of $H^1_{\mathcal{P}}$ defined by $\theta_{\mathcal{P}}(uh) = \theta(h)$ for every $u \in J^1 \cap \mathcal{U}$ and every $h \in H^1$. Since J^1 is a pro-p-group, proceeding as in Proposition 5.5 of [Sécherre and Stevens 2008] we can construct an irreducible representation $\eta_{\mathcal{P}}$ of $J^1_{\mathcal{P}}$, unique up to isomorphism, whose restriction to $H^1_{\mathcal{P}}$ contains $\theta_{\mathcal{P}}$. Actually it is the natural representation of $J^1_{\mathcal{P}}$ on the $J^1 \cap \mathcal{U}$-invariants of η. Furthermore, $\mathrm{ind}_{J^1_{\mathcal{P}}}^{J^1}(\eta_{\mathcal{P}})$ is isomorphic to η, $I_G(\eta_{\mathcal{P}}) = J^1_{\mathcal{P}} B^\times J^1_{\mathcal{P}}$ and for every $y \in B^\times$ we have $\dim_R(I_y(\eta_{\mathcal{P}})) = 1$. We remark that $(J^1_{\mathcal{P}}, \eta_{\mathcal{P}})$ is decomposed with respect to $(\mathcal{M}, \mathcal{P})$ and the restriction of η_P to $J^1_{\mathcal{M}}$ is $\eta_{\mathcal{M}}$. We denote by $V_{\mathcal{M}}$ the R-vector space of $\eta_{\mathcal{M}}$ and $\eta_{\mathcal{P}}$.

Since $\mathrm{ind}_{J^1_{\mathcal{P}}}^{J^1}(\eta_{\mathcal{P}})$ is isomorphic to η, we can identify the R-vector space V_η of η with the vector space of functions $\varphi : J^1 \to V_{\mathcal{M}}$ such that $\varphi(xj) = \eta_{\mathcal{P}}(x)\varphi(j)$ for every $x \in J^1_{\mathcal{P}}$ and $j \in J^1$. In this case $\eta(j)\varphi : x \mapsto \varphi(xj)$. By the Mackey formula, $V_{\mathcal{M}}$ is a direct summand of V_η and we can identify it with the subspace of functions $\varphi \in V_\eta$ with support in $J^1_{\mathcal{P}}$. This identification is given by $\varphi \mapsto \varphi(1)$ whose inverse is $v \mapsto \varphi_v$ where the support of φ_v is $J^1_{\mathcal{P}}$ and $\varphi_v(1) = v$. Let $\boldsymbol{p} : V_\eta \to V_{\mathcal{M}}$ be the canonical projection, i.e., the restriction of a function in V_η to $J^1_{\mathcal{P}}$, and let $\iota : V_{\mathcal{M}} \to V_\eta$ be the inclusion.

Remark 3.6. In general we cannot define a representation κ_P of $J_P = (J \cap P)H^1$ as in Section 2.3 of [Sécherre 2005a] or in Section 5.5 of [Sécherre and Stevens 2008], because the decomposition e conforms to Λ over E but it is not subordinate to $\mathfrak{B}$. In our case ($\mathfrak{B}$ maximal) the only decomposition which conforms to Λ over E and is subordinate to $\mathfrak{B}$ is the trivial one.

Lemma 3.7. (1) *For every $j \in J^1_{\mathcal{P}}$ we have $\eta(j) \circ \iota = \iota \circ \eta_{\mathcal{P}}(j)$ and $\boldsymbol{p} \circ \eta(j) = \eta_{\mathcal{P}}(j) \circ \boldsymbol{p}$.*

(2) *For every $j \in J^1$ we have*

$$\boldsymbol{p} \circ \eta(j) \circ \iota = \begin{cases} \eta_{\mathcal{P}}(j) & \text{if } j \in J^1_{\mathcal{P}}, \\ 0 & \text{otherwise.} \end{cases}$$

(3) $\sum_{j \in J^1/J^1_{\mathcal{P}}} \eta(j) \circ \iota \circ \boldsymbol{p} \circ \eta(j^{-1})$ *is the identity of* $\mathrm{End}_R(V_{\mathcal{M}})$.

Proof. To prove the first point, let $\varphi_v \in V_{\mathcal{M}}$ and $\varphi \in V_\eta$. Then $\eta(j)(\iota(\varphi_v))(1) = \varphi_v(j) = \eta_P(j)v$ and $\boldsymbol{p}(\eta(j)(\varphi))(1) = \varphi(j) = \eta_{\mathcal{P}}(j)\varphi(1)$. To prove the second point we observe that if $j \in J^1_P$ then $\boldsymbol{p} \circ \eta(j) \circ \iota = \boldsymbol{p} \circ \iota \circ \eta_{\mathcal{P}}(j) = \eta_{\mathcal{P}}(j)$ while if $j \notin J^1_{\mathcal{P}}$ the support of $\eta(j)(\iota(\varphi_v))$ is in $J^1_{\mathcal{P}} j^{-1}$ for every $\varphi_v \in V_{\mathcal{M}}$ and so $\boldsymbol{p} \circ \eta(j) \circ \iota = 0$. Finally, to prove the third point we observe that for every $\varphi \in V_\eta$ the function $\varphi_j = (\eta(j) \circ \iota \circ \boldsymbol{p} \circ \eta(j^{-1}))\varphi$ has support in $J^1_{\mathcal{P}} j^{-1}$ and $\varphi_j(j^{-1}) = \varphi(j^{-1})$. □

We consider the surjective linear map $\mu : \mathrm{End}_R(V_\eta) \to \mathrm{End}_R(V_{\mathcal{M}})$ given by $f \mapsto \boldsymbol{p} \circ f \circ \iota$.

Lemma 3.8. *The map $\zeta : \mathscr{H}_R(G,\eta) \to \mathscr{H}_R(G,\eta_{\mathcal{P}})$ defined by $\Phi \mapsto \mu \circ \Phi$ for every $\Phi \in \mathscr{H}_R(G,\eta)$ is an isomorphism of R-algebras. Moreover, if the support of $\Phi \in \mathscr{H}_R(G,\eta)$ is in $J^1 x J^1$ with $x \in B^\times$ then the support of $\zeta(\Phi)$ is in $J^1_{\mathcal{P}} x J^1_{\mathcal{P}}$.*

Proof. Let $\Phi \in \mathscr{H}_R(G,\eta)$. Then the support of $\mu \circ \Phi$ is contained in the support of Φ which is compact. Furthermore, for every $x_1, x_2 \in J^1_{\mathcal{P}}$ and every $j \in J^1$ we have $\mu(\Phi(x_1 j x_2)) = \boldsymbol{p} \circ \eta(x_1) \circ \Phi(j) \circ \eta(x_2) \circ \iota$ which, by Lemma 3.7, is $\eta_{\mathcal{P}}(x_1) \circ \mu(\Phi(j)) \circ \eta_{\mathcal{P}}(x_2)$. Hence, ζ is well defined and it is R-linear. Let $\Phi_1, \Phi_2 \in \mathscr{H}_R(G,\eta)$. For every $g \in G$ we have

$$\begin{aligned}
((\mu\circ\Phi_1)*(\mu\circ\Phi_2))(g) &= \sum_{x\in G/J^1_{\mathcal{P}}} \boldsymbol{p}\circ\Phi_1(x)\circ\iota\circ\boldsymbol{p}\circ\Phi_2(x^{-1}g)\circ\iota \\
&= \sum_{y\in G/J^1}\sum_{z\in J^1/J^1_{\mathcal{P}}} \boldsymbol{p}\circ\Phi_1(yz)\circ\iota\circ\boldsymbol{p}\circ\Phi_2(z^{-1}y^{-1}g)\circ\iota \\
&= \sum_{y\in G/J^1} \boldsymbol{p}\circ\Phi_1(y)\circ\Bigg(\sum_{z\in J^1/J^1_{\mathcal{P}}} \eta(z)\circ\iota\circ\boldsymbol{p}\circ\eta(z^{-1})\Bigg)\circ\Phi_2(y^{-1}g)\circ\iota \\
&= \sum_{y\in G/J^1} \boldsymbol{p}\circ\Phi_1(y)\circ\Phi_2(y^{-1}g)\circ\iota \\
\text{(Lemma 3.7)}\ &= (\mu\circ(\Phi_1*\Phi_2))(g)
\end{aligned}$$

and so ζ is a homomorphism of R-algebras. Let $\Phi \in \mathscr{H}_R(G,\eta)$ such that $\boldsymbol{p}\circ\Phi(g)\circ\iota = 0$ for every $g \in G$. Then by Lemma 3.7, for every $g' \in G$ we have

$$\begin{aligned}
\Phi(g') &= \sum_{j_1\in J^1/J^1_{\mathcal{P}}} \eta(j_1)\circ\iota\circ\boldsymbol{p}\circ\eta(j_1^{-1})\circ\Phi(g')\circ\sum_{j_2\in J^1/J^1_{\mathcal{P}}} \eta(j_2)\circ\iota\circ\boldsymbol{p}\circ\eta(j_2^{-1}) \\
&= \sum_{j_1,j_2\in J^1/J^1_{\mathcal{P}}} \eta(j_1)\circ\iota\circ(\boldsymbol{p}\circ\Phi(j_1^{-1}g'j_2)\circ\iota)\circ\boldsymbol{p}\circ\eta(j_2^{-1}) \\
&= 0
\end{aligned}$$

and then ζ is injective. Now, we know that $\mathscr{H}_R(G,\eta) \cong \mathrm{End}_G(\mathrm{ind}_{J^1}^G(\eta))$, $\mathscr{H}_R(G,\eta_{\mathcal{P}}) \cong \mathrm{End}_G(\mathrm{ind}_{J^1_{\mathcal{P}}}^G(\eta_{\mathcal{P}}))$ and $\mathrm{ind}_{J^1_{\mathcal{P}}}^{J^1}(\eta_{\mathcal{P}}) \cong \eta$. Then by transitivity of the induction we have $\mathscr{H}_R(G,\eta) \cong \mathscr{H}_R(G,\eta_{\mathcal{P}})$ and then ζ must be bijective. Furthermore, if $\Phi \in \mathscr{H}_R(G,\eta)$ has support in J^1xJ^1 with $x \in B^\times$ then the support of $\zeta(\Phi)$ is in $J^1xJ^1 \cap I_G(\eta_{\mathcal{P}}) = J^1xJ^1 \cap J^1_{\mathcal{P}}B^\times J^1_{\mathcal{P}} = J^1_{\mathcal{P}}xJ^1_{\mathcal{P}}$. □

Lemma 3.9. *Let $x_1, x_2 \in B^\times$ and let $\tilde{f}_i \in \mathscr{H}_R(G,\eta)_{J^1x_iJ^1}$ and $\hat{f}_i = \zeta(\tilde{f}_i)$ for $i \in \{1,2\}$.*

(1) *If x_1 or x_2 normalizes $J^1_{\mathcal{P}}$ then the support of $\hat{f}_1 * \hat{f}_2$ is in $J^1_{\mathcal{P}}x_1x_2J^1_{\mathcal{P}}$ and*

$$(\hat{f}_1 * \hat{f}_2)(x_1x_2) = \hat{f}_1(x_1)\circ\hat{f}_2(x_2).$$

(2) *If x_1 or x_2 normalizes J^1 then the support of $\hat{f}_1 * \hat{f}_2$ is in $J^1_{\mathcal{P}}x_1x_2J^1_{\mathcal{P}}$ and*

$$(\hat{f}_1 * \hat{f}_2)(x_1x_2) = \boldsymbol{p}\circ\tilde{f}_1(x_1)\circ\tilde{f}_2(x_2)\circ\iota.$$

Proof. The first point follows from Remark 1.2. If x_1 or x_2 normalizes J^1, by Remark 1.2 the support of $\tilde{f}_1 * \tilde{f}_2$ is in $J^1 x_1 x_2 J^1$ and so the support of $\hat{f}_1 * \hat{f}_2 = \zeta(\tilde{f}_1 * \tilde{f}_2)$ is in $J^1 x_1 x_2 J^1 \cap I_G(\eta_{\mathcal{P}}) = J^1_{\mathcal{P}} x_1 x_2 J^1_{\mathcal{P}}$ and moreover

$$(\hat{f}_1 * \hat{f}_2)(x_1 x_2) = \zeta(\tilde{f}_1 * \tilde{f}_2)(x_1 x_2) = \boldsymbol{p} \circ \tilde{f}_1(x_1) \circ \tilde{f}_2(x_2) \circ \iota. \qquad \square$$

Lemma 3.10. *For every $x \in B^\times \cap \mathcal{M}$ and every $y \in I_G(\eta_{\mathcal{P}})$ which normalizes $J^1_{\mathcal{M}}$ we have $I_x(\eta_{\mathcal{P}}) = I_x(\eta_{\mathcal{M}})$ and $I_y(\eta_{\mathcal{P}}) = I_y(\eta_{\mathcal{M}})$. Moreover, every nonzero element in $I_z(\eta_{\mathcal{P}})$, with $z \in I_G(\eta_{\mathcal{P}})$, is invertible.*

Proof. For the first assertion, in both cases the R-vector spaces are 1-dimensional and so it suffices to prove an inclusion. Since $\eta_{\mathcal{M}}$ is the restriction of $\eta_{\mathcal{P}}$ to $J^1_{\mathcal{M}}$, for every $x' \in I_G(\eta_{\mathcal{P}})$ we have $I_{x'}(\eta_{\mathcal{P}}) \subseteq I_{x'}(\eta_{\mathcal{M}})$. For the second assertion, we observe that $I_G(\eta_{\mathcal{P}}) = J^1_{\mathcal{P}} B^\times J^1_{\mathcal{P}} = J^1_{\mathcal{P}} I_B \tilde{W} I_B J^1_{\mathcal{P}}$. Now I_B normalizes $J^1_{\mathcal{P}}$ since it is contained in $J^1(J \cap \mathcal{P})$ while $\tilde{W}$ normalizes $J^1_{\mathcal{M}}$. Take $z = z_1 z_2 z_3 \in I_G(\eta_{\mathcal{P}})$ with $z_1 \in J^1_{\mathcal{P}} I_B$, $z_2 \in \tilde{W}$ and $z_3 \in I_B J^1_{\mathcal{P}}$ and take a nonzero element γ in $I_z(\eta_P)$. Let γ_1 and γ_3 be invertible elements in $I_{z_1^{-1}}(\eta_{\mathcal{P}})$ and in $I_{z_3^{-1}}(\eta_{\mathcal{P}})$ respectively. Then $\gamma_1 \circ \gamma \circ \gamma_3$ is a nonzero element in $I_{z_2}(\eta_{\mathcal{P}}) = I_{z_2}(\eta_{\mathcal{M}})$ and so it is invertible. $\square$

3C. *The isomorphism* $\mathscr{H}_R(J,\eta) \cong \mathscr{H}_R(K_B, K^1_B)$. We now prove that the subalgebra $\mathscr{H}_R(K_B, K^1_B)$ of $\mathscr{H}_R(B^\times, K^1_B)$ is isomorphic to the subalgebra $\mathscr{H}_R(J, \eta_{\mathcal{P}})$ of $\mathscr{H}_R(G, \eta_{\mathcal{P}})$ and so to $\mathscr{H}_R(J,\eta)$.

In accordance with Chapter 2 of [Chinello 2017], we denote by $f_x \in \mathscr{H}_R(B^\times, K^1_B)$ the characteristic function of $K^1_B x K^1_B$ for every $x \in B^\times$ and we write $\Phi_1\Phi_2 = \Phi_1 * \Phi_2$ for every Φ_1 and Φ_2 in $\mathscr{H}_R(B^\times, K^1_B)$, in $\mathscr{H}_R(G,\eta)$ or in $\mathscr{H}_R(G,\eta_{\mathcal{P}})$.

We observe that every element in $\mathscr{H}_R(J,\eta_{\mathcal{P}})$ has support in $J \cap J^1_{\mathcal{P}} B^\times J^1_{\mathcal{P}} = J^1_{\mathcal{P}}(J \cap B^\times) J^1_{\mathcal{P}} = J^1_{\mathcal{P}} K_B J^1_{\mathcal{P}}$ and so its image by ζ^{-1} has support in $J^1 K_B J^1$. This implies that ζ induces an algebra isomorphism from $\mathscr{H}_R(J,\eta)$ to $\mathscr{H}_R(J,\eta_{\mathcal{P}})$. We also remark that $\mathscr{H}_R(K_B, K^1_B)$ is isomorphic to the group algebra $R[K_B/K^1_B] \cong R[J/J^1]$, then we can identify every $\Phi \in \mathscr{H}_R(K_B, K^1_B)$ with a function $\Phi \in \mathscr{H}_R(J, J^1)$.

From now on we fix a β-extension κ of η. We recall that $\mathrm{res}^J_{J^1} \kappa = \eta$, $I_G(\eta) = I_G(\kappa) = J^1 B^\times J^1$ and for every $y \in B^\times$ we have $I_y(\eta) = I_y(\kappa)$ which is an R-vector space of dimension 1. Then V_η is also the R-vector space of κ and $\kappa(j) \in I_j(\eta)$ for every $j \in J$.

Lemma 3.11. *The map $\Theta' : \mathscr{H}_R(K_B, K^1_B) \to \mathscr{H}_R(J,\eta)$ defined by $\Phi \mapsto \Phi \otimes \kappa$ for every $\Phi \in \mathscr{H}_R(K_B, K^1_B)$ is an algebra isomorphism.*

Proof. The map is well defined since for every $\Phi \in \mathscr{H}_R(K_B, K^1_B)$ we have $\Phi \otimes \kappa : J \to \mathrm{End}_R(V_\eta)$ and $(\Phi \otimes \kappa)(j_1 j j'_1) = \Phi(j)\kappa(j_1 j j'_1) = \eta(j_1) \circ (\Phi(j)\kappa(j)) \circ \eta(j'_1)$ for every $j \in J$ and $j_1, j'_1 \in J^1$. It is clearly R-linear and

$$\begin{aligned}\Theta'(\Phi_1 * \Phi_2)(j) &= \sum_{x \in J/J^1} \Phi_1(x)\Phi_2(x^{-1}j)\kappa(j) = \sum_{x \in J/J^1} \Phi_1(x)\Phi_2(x^{-1}j)\kappa(x) \circ \kappa(x^{-1}j) \\ &= \sum_{x \in J/J^1} (\Phi_1(x)\kappa(x)) \circ (\Phi_2(x^{-1}j)\kappa(x^{-1}j)) = (\Theta'(\Phi_1) * \Theta'(\Phi_2))(j)\end{aligned}$$

for every $\Phi_1,\Phi_2\in\mathscr{H}_R(K_B,K_B^1)$ and $j\in J$. Hence, Θ' is an R-algebra homomorphism. It is injective because $\kappa(j)\in\mathrm{GL}(V_\eta)$ for every $j\in J$. Let $\tilde f\in\mathscr{H}_R(J,\eta)$ and $j\in J$. Since $\tilde f(j)\in I_j(\eta)=\mathrm{Hom}_{J^1}(\eta,\eta^j)$, which is of dimension 1, we have $\tilde f(j)\in R\kappa(j)$ and then we can write $\tilde f(j)=\Phi(j)\kappa(j)$ with $\Phi:J\to R$. Since $\tilde f\in\mathscr{H}_R(J,\eta)$, for every $j_1\in J^1$ we have

$$\Phi(j_1j)\kappa(j_1j)=\tilde f(j_1j)=\eta(j_1)\tilde f(j)=\eta(j_1)\Phi(j)\kappa(j)=\Phi(j)\kappa(j_1j)$$

and so $\Phi\in\mathscr{H}_R(J,J^1)$. We conclude that Θ' is surjective and then it is an algebra isomorphism. □

Composing the restriction of ζ to $\mathscr{H}_R(J,\eta)$ with Θ' we obtain an algebra isomorphism $\mathscr{H}_R(K_B,K_B^1)\to\mathscr{H}_R(J,\eta_{\mathcal{P}})$. For every $x\in K_B$ let $\tilde f_x=\Theta'(f_x)\in\mathscr{H}_R(J,\eta)$ which is given by $\tilde f_x(y)=\kappa(y)$ for every $y\in J^1xJ^1=J^1x$ and let $\hat f_x=\zeta(\tilde f_x)\in\mathscr{H}_R(J,\eta_{\mathcal{P}})$ which is given by $\hat f_x(z)=\boldsymbol{p}\circ\kappa(z)\circ\iota$ for every $z\in J^1_{\mathcal{P}}xJ^1_{\mathcal{P}}$.

3D. *Generators and relations of $\mathscr{H}_R(B^\times,K_B^1)$*. In this section we introduce some notation and recall the presentation by generators and relations of the algebra $\mathscr{H}_R(B^\times,K_B^1)$ presented in [Chinello 2017].

We set $\Omega=K_B\cup\{\tau_0,\tau_0^{-1}\}\cup\{\tau_\alpha\mid\alpha\in\Sigma\}$ and $\mathbf{\Omega}=\{f_\omega\mid\omega\in\Omega\}$ which is a finite set. We now define some subgroups of G, through its identification with $\mathrm{GL}_{m'}(A(E))$. For every $\alpha=\alpha_{ij}\in\mathbf{\Phi}$ we denote by $\mathcal{U}_\alpha$ the subgroup of matrices $(a_{hk})\in G$ with $a_{hh}=1$ for every $h\in\{1,\dots,m'\}$, $a_{ij}\in A(E)$ and $a_{hk}=0$ if $h\neq k$ and $(h,k)\neq(i,j)$. For every $P\subset\Sigma$ we denote by $\mathcal{M}_P$ the standard Levi subgroup associated to P and by $\mathcal{U}_P^+$ and $\mathcal{U}_P^-$ the unipotent radical of, respectively, upper and lower standard parabolic subgroups with Levi component $\mathcal{M}_P$. We remark that $\mathcal{M}=\mathcal{M}_\varnothing$, $\mathcal{U}=\mathcal{U}_\varnothing$ and $\mathcal{U}^-=\mathcal{U}_\varnothing^-$. Thus, we have $\mathcal{U}_P^+=\prod_{\alpha\in\Psi_P^+}\mathcal{U}_\alpha$ and $\mathcal{U}_P^-=\prod_{\alpha\in\Psi_P^-}\mathcal{U}_\alpha$. Furthermore, if $P_1\subset P_2\subset\Sigma$ then $\mathcal{U}_{P_2}^+$ is a subgroup of $\mathcal{U}_{P_1}^+$ and $\mathcal{U}_{P_2}^-$ a subgroup of $\mathcal{U}_{P_1}^-$.

Remark 3.12. By Proposition 3.4, if we take $\alpha=\alpha_{ij}\in\mathbf{\Phi}$ and (a_{hk}) in $\mathcal{U}_\alpha\cap J^1$ or $\mathcal{U}_\alpha\cap H^1$ then a_{ij} is in $\mathfrak{J}_0^1$ or $\mathfrak{H}_0^1$, respectively.

Remark 3.13. In accordance with Section 2.2 of [Chinello 2017] we set $M_P=\mathcal{M}_P\cap K_B$, $U_P^+=\mathcal{U}_P^+\cap K_B$ and $U_P^-=\mathcal{U}_P^-\cap K_B$ for every $P\subset\Sigma$ and $U_\alpha=\mathcal{U}_\alpha\cap K_B$ for every $\alpha\in\mathbf{\Phi}$.

As in Section 2.3 of [Chinello 2017], for every $\alpha=\alpha_{i,i+1}\in\Sigma$ and $w\in W$ we consider the following sets: $A(w,\alpha)=\{w(j)\mid i+1\le j\le m'\}$, $B(w,\alpha)=\{w(j)-1\mid i+1\le j\le m'\}$, $P'(w,\alpha)=A(w,\alpha)\setminus B(w,\alpha)$, $P(w,\alpha)=\{\alpha_{i,i+1}\in\Sigma\mid i\in P'(w,\alpha)\}$ and $Q(w,\alpha)=B(w,\alpha)\setminus A(w,\alpha)$. We remark that $\tau_{P'(w,\alpha)}=\tau_{P(w,\alpha)}$ because $0\notin P'(w,\alpha)$ and $\tau_{m'}=\mathbb{I}_{m'}$. Moreover, if $\alpha=\alpha_{i,i+1}\in\Sigma$, $w'\in W$ and w is of minimal length in $w'W_{\hat\alpha}\in W/W_{\hat\alpha}$ then we have

$$w'\tau_iw'^{-1}=w\tau_iw^{-1}=\prod_{h=i+1}^{m'}w\tau_{h-1}\tau_h^{-1}w^{-1}=\prod_{h=i+1}^{m'}\tau_{w(h)-1}\tau_{w(h)}^{-1}=\tau_{P(w,\alpha)}^{-1}\tau_{Q(w,\alpha)}.$$

Lemma 3.14. *The algebra $\mathscr{H}_R(B^\times,K_B^1)$ is the R-algebra generated by $\mathbf{\Omega}$ subject to the following relations*:

(1) $f_k=1$ *for every* $k\in K^1$ *and* $f_{k_1}f_{k_2}=f_{k_1k_2}$ *for every* $k_1,k_2\in K$.

(2) $f_{\tau_0} f_{\tau_0^{-1}} = 1$ *and* $f_{\tau_0^{-1}} f_\omega = f_{\tau_0^{-1}\omega\tau_0} f_{\tau_0^{-1}}$ *for every* $\omega \in \Omega$.

(3) $f_{\tau_\alpha} f_x = f_{\tau_\alpha x \tau_\alpha^{-1}} f_{\tau_\alpha}$ *for every* $\alpha \in \Sigma$ *and* $x \in M_{\hat\alpha}$.

(4) $f_u f_{\tau_\alpha} = f_{\tau_\alpha}$ *if* $u \in U_{\alpha'}$ *with* $\alpha' \in \Psi^+_{\hat\alpha}$, *for every* $\alpha \in \Sigma$.

(5) $f_{\tau_\alpha} f_u = f_{\tau_\alpha}$ *if* $u \in U_{\alpha'}$ *with* $\alpha' \in \Psi^-_{\hat\alpha}$, *for every* $\alpha \in \Sigma$.

(6) $f_{\tau_\alpha} f_{\tau_{\alpha'}} = f_{\tau_{\alpha'}} f_{\tau_\alpha}$ *for every* $\alpha, \alpha' \in \Sigma$.

(7) $\big(\prod_{\alpha' \in P(w,\alpha)} f_{\tau_{\alpha'}}\big) f_w f_{\tau_\alpha} f_{w^{-1}} = q^{\ell(w)} \big(\prod_{\alpha'' \in Q(w,\alpha)} f_{\tau_{\alpha''}}\big)\big(\sum_u f_u\big)$ *for every* $\alpha \in \Sigma$ *and* w *of minimal length in* $wW_{\hat\alpha} \in W/W_{\hat\alpha}$ *and where* u *runs over a system of representatives of* $(U \cap wU^- w^{-1})K_B^1/K_B^1$ *in* $U \cap wU^- w^{-1}$.

Proof. The only difference between this presentation and that in [Chinello 2017] is relation 3 which is equivalent to relations 3, 4 and 7 of Definition 2.21 of [Chinello 2017] because $\mathcal{M} \cap K_B$, $U_{\alpha'}$ with $\alpha' \in \mathbf{\Phi}_{\hat\alpha}$ and $W_{\hat\alpha}$ generate $M_{\hat\alpha}$. □

Hence, to define an algebra homomorphism from $\mathscr{H}_R(B^\times, K_B^1)$ to $\mathscr{H}_R(G, \eta_{\mathcal{P}})$, it is sufficient to choose elements $\hat f_\omega \in \mathscr{H}_R(G, \eta_{\mathcal{P}})$ for every $\omega \in \Omega$ such that the $\hat f_\omega$ respect the relations of Lemma 3.14. We remark that we can take $\hat f_\omega \in \mathscr{H}_R(G, \eta_{\mathcal{P}})_{J^1_{\mathcal{P}}\omega J^1_{\mathcal{P}}}$ for every $\omega \in \Omega$ and we recall that in Section 3C we have defined $\hat f_k$ for every $k \in K_B$ as the image of f_k by $\zeta \circ \Theta'$.

3E. ***Some decompositions of $J^1_{\mathcal{P}}$-double cosets.*** In this section we introduce some notation and some tools that we will use to construct elements in $\mathscr{H}_R(G, \eta_{\mathcal{P}})_{J^1_{\mathcal{P}}\tau_i J^1_{\mathcal{P}}}$ with $i \in \{0, \dots, m'-1\}$.

Lemma 3.15. *Let* $\tau \in \mathbf{\Delta}$ *and* $P = P(\tau)$.

(1) *We have* $J^1_{\mathcal{P}} = (J^1_{\mathcal{P}} \cap \mathcal{U}^-_P)(J^1_{\mathcal{P}} \cap \mathcal{M}_P)(J^1_{\mathcal{P}} \cap \mathcal{U}^+_P) = (J^1_{\mathcal{P}} \cap \mathcal{U}^+_P)(J^1_{\mathcal{P}} \cap \mathcal{M}_P)(J^1_{\mathcal{P}} \cap \mathcal{U}^-_P)$.

(2) *We have* $(J^1_{\mathcal{P}} \cap \mathcal{U}^+_P)^\tau \subset H^1 \cap \mathcal{U}^+_P \subset J^1_{\mathcal{P}} \cap \mathcal{U}^+_P$, $(J^1_{\mathcal{P}} \cap \mathcal{U}^-_P)^{\tau^{-1}} \subset (J^1 \cap \mathcal{U}^-_P)^{\tau^{-1}} \subset H^1 \cap \mathcal{U}^-_P = J^1_{\mathcal{P}} \cap \mathcal{U}^-_P$ *and* $(J^1_{\mathcal{P}} \cap \mathcal{M}_P)^\tau = J^1_{\mathcal{P}} \cap \mathcal{M}_P$.

(3) *We have* $(J^1_{\mathcal{P}} \cap \mathcal{U})^\tau \subset J^1_{\mathcal{P}} \cap \mathcal{U}$, $(J^1_{\mathcal{P}} \cap \mathcal{U}^-)^{\tau^{-1}} \subset J^1_{\mathcal{P}} \cap \mathcal{U}^-$ *and* $(J^1_{\mathcal{M}})^\tau = J^1_{\mathcal{M}}$.

Proof. The first point follows from Remark 3.5. To prove the second point we observe that Remark 3.12 implies that $(J^1_{\mathcal{P}} \cap \mathcal{U}^+_P)^\tau = (J^1 \cap \prod_{\alpha \in \Psi^+_P} \mathcal{U}_\alpha)^\tau$ is contained in $(\mathbb{I}_{m'} + \varpi \mathfrak{J}^1) \cap \mathcal{U}^+_P$ which is in $H^1 \cap \mathcal{U}^+_P \subset J^1_{\mathcal{P}} \cap \mathcal{U}^+_P$. Similarly we prove $(J^1 \cap \mathcal{U}^-_P)^{\tau^{-1}} \subset H^1 \cap \mathcal{U}^-_P$. Moreover, since $\varpi^{-1}\mathfrak{J}^1_0 \varpi = \mathfrak{J}^1_0$ and $\varpi^{-1}\mathfrak{H}^1_0 \varpi = \mathfrak{H}^1_0$, we have $(J^1_{\mathcal{P}} \cap \mathcal{M}_P)^\tau = J^1_{\mathcal{P}} \cap \mathcal{M}_P$. To prove the third point, we observe that $(J^1_{\mathcal{P}} \cap \mathcal{U})^\tau \subset ((J^1_{\mathcal{P}} \cap \mathcal{M}_P)(J^1_{\mathcal{P}} \cap \mathcal{U}^+_P))^\tau \cap \mathcal{U}$ which is in $(J^1_{\mathcal{P}} \cap \mathcal{M}_P)(J^1_{\mathcal{P}} \cap \mathcal{U}^+_P) \cap \mathcal{U} = J^1_{\mathcal{P}} \cap \mathcal{U}$. Similarly we prove $(J^1_{\mathcal{P}} \cap \mathcal{U}^-)^{\tau^{-1}} \subset J^1_{\mathcal{P}} \cap \mathcal{U}^-$. Finally, since $\varpi^{-1}\mathfrak{J}^1_0 \varpi = \mathfrak{J}^1_0$ we obtain $(J^1_{\mathcal{M}})^\tau = J^1_{\mathcal{M}}$. □

Lemma 3.16. *Let* $\tau, \tau' \in \mathbf{\Delta}$ *and* $w \in W$.

(1) *We have* $J^1_{\mathcal{P}} \tau J^1_{\mathcal{P}} = (J^1_{\mathcal{P}} \cap \mathcal{U}^-_{P(\tau)})\tau J^1_{\mathcal{P}} = J^1_{\mathcal{P}} \tau (J^1_{\mathcal{P}} \cap \mathcal{U}^+_{P(\tau)})$ *and* $J^1_{\mathcal{P}} \tau^{-1} J^1_{\mathcal{P}} = (J^1_{\mathcal{P}} \cap \mathcal{U}^+_{P(\tau)})\tau^{-1} J^1_{\mathcal{P}} = J^1_{\mathcal{P}} \tau^{-1} (J^1_{\mathcal{P}} \cap \mathcal{U}^-_{P(\tau)})$.

(2) *We have* $(J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} = (J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-) J^1_{\mathcal{P}}$.

(3) *We have* $J^1_{\mathcal{P}} \mathcal{U}^- J^1_{\mathcal{P}} \cap \mathcal{U} = J^1_{\mathcal{P}} \cap \mathcal{U}$ *and* $J^1_{\mathcal{P}} \mathcal{U} J^1_{\mathcal{P}} \cap \mathcal{U}^- = J^1_{\mathcal{P}} \cap \mathcal{U}^-$.

(4) *We have* $J^1_{\mathcal{P}}\tau J^1_{\mathcal{P}}\tau' J^1_{\mathcal{P}} = J^1_{\mathcal{P}}\tau\tau' J^1_{\mathcal{P}}$ *and* $(J^1_{\mathcal{P}})^{\tau} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{\tau'^{-1}} J^1_{\mathcal{P}} = J^1_{\mathcal{P}}$.

Proof. Let $P = P(\tau)$.

(1) By Lemma 3.15 we have $J^1_{\mathcal{P}} = (J^1_{\mathcal{P}} \cap \mathcal{U}^-_P)(J^1_{\mathcal{P}} \cap \mathcal{M}_P)(J^1_{\mathcal{P}} \cap \mathcal{U}^+_P)$ and so we obtain $J^1_{\mathcal{P}}\tau J^1_{\mathcal{P}} = (J^1_{\mathcal{P}} \cap \mathcal{U}^-_P)\tau(J^1_{\mathcal{P}} \cap \mathcal{M}_P)^{\tau}(J^1_{\mathcal{P}} \cap \mathcal{U}^+_P)^{\tau} J^1_{\mathcal{P}}$ which is equal to $(J^1_{\mathcal{P}} \cap \mathcal{U}^-_P)\tau J^1_{\mathcal{P}}$ by Lemma 3.15. We prove the other equalities similarly.

(2) Since $(H^1 \cap \mathcal{U}^-)^w \subset J^1_{\mathcal{P}}$ and $(J^1_{\mathcal{M}})^w = J^1_{\mathcal{M}}$ we obtain $(J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} = (J^1 \cap \mathcal{U})^w J^1_{\mathcal{P}}$. Moreover, we have $(J^1 \cap \mathcal{U})^w \cap \mathcal{U} \subset J^1_{\mathcal{P}}$ and so $(J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} = (J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-)J^1_{\mathcal{P}}$.

(3) We have $J^1_{\mathcal{P}}\mathcal{U}^- J^1_{\mathcal{P}} \cap \mathcal{U} = (J^1_{\mathcal{P}} \cap \mathcal{U})\big((J^1_{\mathcal{P}} \cap \mathcal{M})\mathcal{U}^-(J^1_{\mathcal{P}} \cap \mathcal{M}) \cap \mathcal{U}\big)(J^1_{\mathcal{P}} \cap \mathcal{U})$ which is contained in $(J^1_{\mathcal{P}} \cap \mathcal{U})(\mathcal{P}^- \cap \mathcal{U})(J^1_{\mathcal{P}} \cap \mathcal{U}) = J^1_{\mathcal{P}} \cap \mathcal{U}$. We prove the second statement similarly.

(4) By point 1, we have $J^1_{\mathcal{P}}\tau J^1_{\mathcal{P}}\tau' J^1_{\mathcal{P}} = J^1_{\mathcal{P}}\tau(J^1_{\mathcal{P}} \cap \mathcal{U}^+_{P(\tau)})\tau' J^1_{\mathcal{P}}$ which is equal to $J^1_{\mathcal{P}}\tau\tau'(J^1_{\mathcal{P}} \cap \mathcal{U}^+_{P(\tau)})^{\tau'} J^1_{\mathcal{P}}$. By Lemma 3.15 it is in $J^1_{\mathcal{P}}\tau\tau'(J^1_{\mathcal{P}} \cap \mathcal{U})^{\tau'} J^1_{\mathcal{P}} \subset J^1_{\mathcal{P}}\tau\tau' J^1_{\mathcal{P}}$ and so we have $J^1_{\mathcal{P}}\tau J^1_{\mathcal{P}}\tau' J^1_{\mathcal{P}} = J^1_{\mathcal{P}}\tau\tau' J^1_{\mathcal{P}}$. By point 1, $(J^1_{\mathcal{P}})^{\tau} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{\tau'^{-1}} J^1_{\mathcal{P}}$ is contained in $(J^1_{\mathcal{P}} \cap \mathcal{U}^-)^{\tau} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}} \cap \mathcal{U})^{\tau'^{-1}} J^1_{\mathcal{P}} = \big((J^1_{\mathcal{P}} \cap \mathcal{U}^-)^{\tau} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}} \cap \mathcal{U})^{\tau'^{-1}}\big) J^1_{\mathcal{P}}$ which is contained in $(\mathcal{U}^- J^1_{\mathcal{P}} \cap \mathcal{U})J^1_{\mathcal{P}}$ and so it is equal to $J^1_{\mathcal{P}}$ by point 3. □

Remark 3.17. We can prove results similar to Lemmas 3.15 and 3.16 with J^1 in place of $J^1_{\mathcal{P}}$.

Lemma 3.18. *Let* $\alpha = \alpha_{i,i+1} \in \Sigma$, $w \in W$ *and* $P = P(w, \alpha)$. *Then* $\Psi^+_{\hat{P}} \cap w\Psi^-_{\hat{\alpha}} = \mathbf{\Phi}^+ \cap w\Psi^-_{\hat{\alpha}}$ *and* $\Psi^-_{\hat{P}} \cap w\Psi^+_{\hat{\alpha}} = \mathbf{\Phi}^- \cap w\Psi^+_{\hat{\alpha}}$. *If in addition* w *is of minimal length in* $wW_{\hat{\alpha}} \in W/W_{\hat{\alpha}}$ *then* $\mathbf{\Phi}^+ \cap w\Psi^-_{\hat{\alpha}} = \mathbf{\Phi}^+ \cap w\mathbf{\Phi}^-$ *and* $\mathbf{\Phi}^- \cap w\Psi^+_{\hat{\alpha}} = \mathbf{\Phi}^- \cap w\mathbf{\Phi}^+$.

Proof. This follows from Lemma 2.19 of [Chinello 2017]. □

From now on, we set $\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0) = [\mathfrak{J}^1_0 : \mathfrak{H}^1_0]$ and $\delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0) = [\mathfrak{H}^1_0 : \varpi\mathfrak{H}^1_0]$.

Remark 3.19. By Remark 3.12, for every $\alpha \in \mathbf{\Phi}$, $\alpha' \in \mathbf{\Phi}^+$ and $\alpha'' \in \mathbf{\Phi}^-$ we have $\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0) = [J^1 \cap \mathcal{U}_{\alpha} : H^1 \cap \mathcal{U}_{\alpha}]$ and $\delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0) = [H^1 \cap \mathcal{U}_{\alpha'} : (H^1 \cap \mathcal{U}_{\alpha'})^{\tau_{\alpha'}}] = [H^1 \cap \mathcal{U}_{\alpha''} : (H^1 \cap \mathcal{U}_{\alpha''})^{\tau^{-1}_{\alpha''}}]$. In particular $\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)$ and $\delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0)$ are powers of p and so they are invertible in R.

From now on we fix $1 \le i \le m' - 1$ and we consider $\alpha = \alpha_{i,i+1}$, w of minimal length in $wW_{\hat{\alpha}}$, $P = P(w, \alpha)$ and $Q = Q(w, \alpha)$.

Remark 3.20. Lemma 3.18 implies that $w\mathcal{U}^-_{\hat{\alpha}}w^{-1} \cap \mathcal{U}^+_{\hat{P}} = w\mathcal{U}^- w^{-1} \cap \mathcal{U}^+$ and $w\mathcal{U}^+_{\hat{\alpha}}w^{-1} \cap \mathcal{U}^-_{\hat{P}} = w\mathcal{U}w^{-1} \cap \mathcal{U}^-$. Moreover, we have $\ell(w) = |\Psi^+_{\hat{P}} \cap w\Psi^-_{\hat{\alpha}}| = |\Psi^-_{\hat{P}} \cap w\Psi^+_{\hat{\alpha}}|$ by Remark 3.1.

We define

$$\mathcal{V}(w, \alpha) = (J^1_{\mathcal{P}} \cap w\mathcal{U}^+_{\hat{\alpha}}w^{-1} \cap \mathcal{U}^-_{\hat{P}})^{w\tau^{-1}_{\alpha}w^{-1}} \tag{4}$$

which is a pro-p-group. We remark that it is equal to $(J^1_{\mathcal{P}} \cap w\mathcal{U}w^{-1} \cap \mathcal{U}^-)^{w\tau^{-1}_{\alpha}w^{-1}}$ by Remark 3.20 and to $(H^1 \cap w\mathcal{U}^+_{\hat{\alpha}}w^{-1} \cap \mathcal{U}^-_{\hat{P}})^{w\tau^{-1}_{\alpha}w^{-1}}$ since $J^1_{\mathcal{P}} \cap \mathcal{U}^-_{\hat{P}} = H^1 \cap \mathcal{U}^-_{\hat{P}}$. Then $\mathcal{V}(w, \alpha)$ is equal to

$$\prod_{\alpha' \in w\mathbf{\Psi}^+_{\hat{\alpha}} \cap \mathbf{\Psi}^-_{\hat{P}}} (H^1 \cap \mathcal{U}_{\alpha'})^{w\tau^{-1}_{\alpha}w^{-1}} = \prod_{\alpha'' \in \mathbf{\Psi}^+_{\hat{\alpha}} \cap w^{-1}\mathbf{\Psi}^-_{\hat{P}}} (H^1 \cap \mathcal{U}_{\alpha''})^{\tau^{-1}_{\alpha}w^{-1}} = \prod_{\alpha' \in w\mathbf{\Psi}^+_{\hat{\alpha}} \cap \mathbf{\Psi}^-_{\hat{P}}} (\mathbb{I}_{m'} + \varpi^{-1}\mathfrak{H}^1) \cap \mathcal{U}_{\alpha'}$$

which is $(\mathbb{I}_{m'}+\varpi^{-1}\mathfrak{H}^1)\cap w\mathcal{U}_{\hat{\alpha}}^+w^{-1}\cap\mathcal{U}_{\hat{P}}^-$. We remark that $\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1=J_{\mathcal{P}}^1\cap w\mathcal{U}_{\hat{\alpha}}^+w^{-1}\cap\mathcal{U}_{\hat{P}}^-$ which is equal to $H^1\cap w\mathcal{U}_{\hat{\alpha}}^+w^{-1}\cap\mathcal{U}_{\hat{P}}^-$ since $J_{\mathcal{P}}^1\cap\mathcal{U}^-=H^1\cap\mathcal{U}^-$.

Lemma 3.21. *The group $wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-$ is in $\mathcal{V}(w,\alpha)$, it normalizes $\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1$ and*

$$(wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-)\cap(\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1)=wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-\cap K_B^1.$$

Proof. We recall that by Remark 3.13 we have $wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-=w\mathcal{U}_{\hat{\alpha}}^+w^{-1}\cap\mathcal{U}_{\hat{P}}^-\cap K_B$. Since $U_{\alpha'}=\tau_\alpha(K_B^1\cap U_{\alpha'})\tau_\alpha^{-1}$ for every $\alpha'\in\mathbf{\Psi}_{\hat{\alpha}}^+$ (see Lemma 2.9 of [Chinello 2017]), then we have $wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-=(K_B^1\cap wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-)^{w\tau_\alpha^{-1}w^{-1}}$ which is contained in $\mathcal{V}(w,\alpha)$. Moreover, the group $wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-$ normalizes $\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1=\mathcal{V}(w,\alpha)\cap H^1$ because we have $wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-\subset K_B$ and K_B normalizes H^1. Finally, since $\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1=H^1\cap w\mathcal{U}_{\hat{\alpha}}^+w^{-1}\cap\mathcal{U}_{\hat{P}}^-$, we have $wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-\cap\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1=wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-\cap H^1$ and, since $K_B\cap H^1=K_B^1$, it is equal to $wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-\cap K_B\cap H^1=wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-\cap K_B^1$. ☐

By Lemma 3.21 the group $\mathcal{V}'=(wU_{\hat{\alpha}}^+w^{-1}\cap U_{\hat{P}}^-)(\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1)$ is a subgroup of $\mathcal{V}(w,\alpha)$. We set

$$d(w,\alpha)=[\mathcal{V}(w,\alpha):\mathcal{V}']\in R$$

which is nonzero because it is a power of p.

Remark 3.22. We have $\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1=H^1\cap w\mathcal{U}_{\hat{\alpha}}^+w^{-1}\cap\mathcal{U}_{\hat{P}}^-=\prod_{\alpha'\in w\mathbf{\Psi}_{\hat{\alpha}}^+\cap\mathbf{\Psi}_{\hat{P}}^-}H^1\cap\mathcal{U}_{\alpha'}$. Hence, by Remarks 3.19 and 3.20 we have

$$[\mathcal{V}(w,\alpha):\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1]=[\varpi^{-1}\mathfrak{H}_0^1:\mathfrak{H}_0^1]^{\ell(w)}=\delta(\mathfrak{H}_0^1,\varpi\mathfrak{H}_0^1)^{\ell(w)}.$$

On the other hand we have $[\mathcal{V}(w,\alpha):\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1]=d(w,\alpha)[\mathcal{V}':\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1]$ which is equal to $d(w,\alpha)[(wU^+w^{-1}\cap U^-)(\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1):\mathcal{V}(w,\alpha)\cap J_{\mathcal{P}}^1]$ and by Remark 3.20 to $d(w,\alpha)[wUw^{-1}\cap U^-:wUw^{-1}\cap U^-\cap K_B^1]=d(w,\alpha)q^{\ell(w)}$ where q is the cardinality of $\mathfrak{k}_{D'}$. So, if we denote $\partial=\delta(\mathfrak{H}_0^1,\varpi\mathfrak{H}_0^1)/q\in R^\times$ then $d(w,\alpha)=\partial^{\ell(w)}$.

Lemma 3.23. *We have $(J_{\mathcal{P}}^1)^{\tau_P}J_{\mathcal{P}}^1\cap(J_{\mathcal{P}}^1)^{w\tau_\alpha^{-1}w^{-1}}J_{\mathcal{P}}^1=\mathcal{V}(w,\alpha)J_{\mathcal{P}}^1$.*

Proof. We have $(J_{\mathcal{P}}^1)^{w\tau_\alpha^{-1}w^{-1}}=(H^1\cap w^{-1}\mathcal{U}^-w)^{\tau_\alpha^{-1}w^{-1}}(J_{\mathcal{M}}^1)^{w\tau_\alpha^{-1}w^{-1}}(J^1\cap w^{-1}\mathcal{U}w)^{\tau_\alpha^{-1}w^{-1}}$. Now we consider the decompositions $H^1\cap w^{-1}\mathcal{U}^-w=(H^1\cap w^{-1}\mathcal{U}^-w\cap\mathcal{U})(H^1\cap w^{-1}\mathcal{U}^-w\cap\mathcal{U}^-)$ and $J^1\cap w^{-1}\mathcal{U}w=(J^1\cap w^{-1}\mathcal{U}w\cap\mathcal{U}^-)(J^1\cap w^{-1}\mathcal{U}w\cap\mathcal{U})$. By Lemma 3.18 we have $J^1\cap w^{-1}\mathcal{U}w\cap\mathcal{U}^-=J^1\cap w^{-1}\mathcal{U}w\cap\mathcal{U}_{\hat{\alpha}}^-$ and so $(J^1\cap w^{-1}\mathcal{U}w\cap\mathcal{U}^-)^{\tau_\alpha^{-1}w^{-1}}$ is contained in $(J^1\cap\mathcal{U}_{\hat{\alpha}}^-)^{\tau_\alpha^{-1}w^{-1}}\subset(H^1\cap\mathcal{U}_{\hat{\alpha}}^-)^{w^{-1}}\subset J_{\mathcal{P}}^1$ and, by Lemma 3.15, $(H^1\cap w^{-1}\mathcal{U}^-w\cap\mathcal{U}^-)^{\tau_\alpha^{-1}w^{-1}}$ is contained in $(H^1\cap\mathcal{U}^-)^{\tau_\alpha^{-1}w^{-1}}\subset(H^1\cap\mathcal{U}^-)^{w^{-1}}\subset J_{\mathcal{P}}^1$. Then, since $(J_{\mathcal{M}}^1)^{w\tau_\alpha^{-1}w^{-1}}=J_{\mathcal{M}}^1$ by Lemma 3.15 and since $(H^1\cap\mathcal{U}^-\cap w\mathcal{U}w^{-1})^{w\tau_\alpha^{-1}w^{-1}}=\mathcal{V}(w,\alpha)$, we obtain $(J_{\mathcal{P}}^1)^{w\tau_\alpha^{-1}w^{-1}}\subset\mathcal{V}(w,\alpha)J_{\mathcal{P}}^1(J^1\cap\mathcal{U}\cap w\mathcal{U}w^{-1})^{w\tau_\alpha^{-1}w^{-1}}$. By Lemma 3.16 and by previous calculations we have

$$(J_{\mathcal{P}}^1)^{\tau_P}J_{\mathcal{P}}^1\cap(J_{\mathcal{P}}^1)^{w\tau_\alpha^{-1}w^{-1}}J_{\mathcal{P}}^1=\big((J_{\mathcal{P}}^1\cap\mathcal{U}_{\hat{P}}^-)^{\tau_P}\cap\mathcal{V}(w,\alpha)J_{\mathcal{P}}^1(J^1\cap\mathcal{U}\cap w\mathcal{U}w^{-1})^{w\tau_\alpha^{-1}w^{-1}}J_{\mathcal{P}}^1\big)J_{\mathcal{P}}^1.$$

Now, since $w\tau_\alpha^{-1}w^{-1}=\tau_Q^{-1}\tau_P$, the group $\mathcal{V}(w,\alpha)$ is contained both in $(\mathcal{U}_{\hat{P}}^-)^{\tau_Q^{-1}\tau_P}=(\mathcal{U}_{\hat{P}}^-)^{\tau_P}$ and in $(J_{\mathcal{P}}^1\cap\mathcal{U}^-)^{\tau_Q^{-1}\tau_P}\subset(J_{\mathcal{P}}^1\cap\mathcal{U}^-)^{\tau_P}\subset(J_{\mathcal{P}}^1)^{\tau_P}$ by Lemma 3.15. This implies $\mathcal{V}(w,\alpha)\subset(J_{\mathcal{P}}^1\cap\mathcal{U}_{\hat{P}}^-)^{\tau_P}$ and so

$(J^1_{\mathcal{P}})^{\tau_P} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{w\tau_\alpha^{-1}w^{-1}} J^1_{\mathcal{P}} = \mathcal{V}(w,\alpha)\big((J^1_{\mathcal{P}} \cap \mathcal{U}^-_{\hat{P}})^{\tau_P} \cap J^1_{\mathcal{P}}(J^1 \cap \mathcal{U} \cap w\mathcal{U}w^{-1})^{w\tau_\alpha^{-1}w^{-1}} J^1_{\mathcal{P}}\big) J^1_{\mathcal{P}}$. Now we have $(J^1_{\mathcal{P}} \cap \mathcal{U}^-_{\hat{P}})^{\tau_P} \cap J^1_{\mathcal{P}}(J^1 \cap \mathcal{U} \cap w\mathcal{U}w^{-1})^{w\tau_\alpha^{-1}w^{-1}} J^1_{\mathcal{P}} \subset \mathcal{U}^- \cap J^1_{\mathcal{P}}\mathcal{U}J^1_{\mathcal{P}}$ that is in $J^1_{\mathcal{P}}$ by point 3 of Lemma 3.16. □

3F. *The group $\tilde{W}$*. In this section we use a presentation by generators and relations of $\tilde{W}$ to find a subgroup of $\mathrm{Aut}_R(V_{\mathcal{M}})$ isomorphic to a quotient of $\tilde{W}$.

Remark 3.24. We know that the Iwahori–Hecke algebra (see I.3.14 of [Vignéras 1996]) is a deformation of the R-algebra $R[\tilde{W}]$ and so it is not difficult to show that $\tilde{W}$ is the group generated by $s_1, \dots, s_{m'-1}$ and $\tau_{m'-1}$ subject to relations $s_i s_j = s_j s_i$ for every i and j such that $|i-j| > 1$, $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$ for every $i \neq m'-1$, $s_i^2 = 1$ for every i, $\tau_{m'-1} s_i = s_i \tau_{m'-1}$ for every $i \neq m'-1$ and $\tau_{m'-1} s_{m'-1} \tau_{m'-1} s_{m'-1} = s_{m'-1} \tau_{m'-1} s_{m'-1} \tau_{m'-1}$.

Lemma 3.25. *Let $i \in \{1, \dots, m'-1\}$, $\alpha = \alpha_{i,i+1}$, $w \in W$ be of minimal length in $wW_{\hat{\alpha}}$ and $\Phi \in \mathscr{H}_R(G, \eta_{\mathcal{P}})_{J^1_{\mathcal{P}}\tau_i J^1_{\mathcal{P}}}$. Then the support of $\hat{f}_w \Phi \hat{f}_{w^{-1}}$ is in $J^1_{\mathcal{P}} w\tau_i w^{-1} J^1_{\mathcal{P}}$ and*

$$(\hat{f}_w \Phi \hat{f}_{w^{-1}})(w\tau_i w^{-1}) = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{\ell(w)} \hat{f}_w(w) \circ \Phi(\tau_i) \circ \hat{f}_{w^{-1}}(w^{-1}).$$

Proof. Since w and w^{-1} normalize J^1, by Lemma 3.9 the support of $\hat{f}_w \Phi \hat{f}_{w^{-1}}$ is in $J^1_{\mathcal{P}} w\tau_i w^{-1} J^1_{\mathcal{P}}$. We recall that

$$(\hat{f}_w \Phi \hat{f}_{w^{-1}})(w\tau_i w^{-1}) = \sum_{x \in G/J^1_{\mathcal{P}}} (\hat{f}_w \Phi)(w\tau_i x) \hat{f}_{w^{-1}}(x^{-1}w^{-1}).$$

By point 2 of Lemma 3.16, the support of the function $x \mapsto (\hat{f}_w \Phi)(w\tau_i x) \hat{f}_{w^{-1}}(x^{-1}w^{-1})$ is contained in $(J^1_{\mathcal{P}})^{w\tau_i} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} = (J^1_{\mathcal{P}})^{w\tau_i} J^1_{\mathcal{P}} \cap (J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-) J^1_{\mathcal{P}}$. Since w is of minimal length in $wW_{\hat{\alpha}}$, by Lemma 3.18 we have $J^1 \cap \mathcal{U}^w \cap \mathcal{U}^- = J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-_{\hat{\alpha}}$ which is included in $(J^1_{\mathcal{P}})^{w\tau_i}$ because $(J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-_{\hat{\alpha}})^{\tau_i^{-1}w^{-1}} = ((J^1 \cap \mathcal{U}^-_{\hat{\alpha}})^{\tau_i^{-1}} \cap \mathcal{U}^w)^{w^{-1}}$ that by Lemma 3.15 is included in $(H^1 \cap \mathcal{U}^-_{\hat{\alpha}})^{w^{-1}} \cap \mathcal{U}$ and so in $J^1_{\mathcal{P}}$. Hence, we obtain $(J^1_{\mathcal{P}})^{w\tau_i} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} = (J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-) J^1_{\mathcal{P}}$. Now, since $(J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-)^{w^{-1}}$ and $(J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-)^{\tau_i^{-1}w^{-1}}$ are contained in $J^1 \cap \mathcal{U}$ and so in the kernel of $\eta_{\mathcal{P}}$ and since we have $[(J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-) J^1_{\mathcal{P}} : J^1_{\mathcal{P}}] = [J^1 \cap \mathcal{U}^w \cap \mathcal{U}^- : H^1 \cap \mathcal{U}^w \cap \mathcal{U}^-] = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{\ell(w)}$ we obtain $(\hat{f}_w \Phi \hat{f}_{w^{-1}})(w\tau_i w^{-1}) = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{\ell(w)} (\hat{f}_w \Phi)(w\tau_i) \circ \hat{f}_{w^{-1}}(w^{-1})$. To conclude we observe that by Lemma 3.9 the support of $\hat{f}_w \Phi$ is contained in $J^1_{\mathcal{P}} w\tau_i J^1_{\mathcal{P}}$ and by points 1 and 2 of Lemma 3.16 the support of $x \mapsto (\hat{f}_w)(wx)\Phi(x^{-1}\tau_i)$ is in $(J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{\tau_i^{-1}} J^1_{\mathcal{P}} = (J^1 \cap \mathcal{U}^w \cap \mathcal{U}^-) J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}} \cap \mathcal{U}^+_{P(\tau_i)})^{\tau_i^{-1}} J^1_{\mathcal{P}}$, which is contained in $(\mathcal{U} J^1_{\mathcal{P}} \cap \mathcal{U}^-) J^1_{\mathcal{P}} = J^1_{\mathcal{P}}$ by point 3 of Lemma 3.16. Hence, $(\hat{f}_w \Phi)(w\tau_i) = \hat{f}_w(w) \circ \Phi(\tau_i)$. □

Lemma 3.26. *Let $w \in W$ and $\alpha \in \Sigma$. Then*

$$\boldsymbol{p} \circ \kappa(w) \circ \iota \circ \boldsymbol{p} \circ \kappa(s_\alpha) \circ \iota = \begin{cases} \boldsymbol{p} \circ \kappa(ws_\alpha) \circ \iota & \text{if } w\alpha > 0, \\ \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{-1} \boldsymbol{p} \circ \kappa(ws_\alpha) \circ \iota & \text{if } w\alpha < 0. \end{cases}$$

Proof. By Lemma 3.11 we have $\hat{f}_w \hat{f}_{s_\alpha} = \hat{f}_{ws_\alpha}$ and then $(\hat{f}_w \hat{f}_{s_\alpha})(ws_\alpha) = \boldsymbol{p} \circ \kappa(ws_\alpha) \circ \iota$. On the other hand we have

$$(\hat{f}_w \hat{f}_{s_\alpha})(ws_\alpha) = \sum_{x \in G/J^1_{\mathcal{P}}} (\hat{f}_w)(wx) \hat{f}_{s_\alpha}(x^{-1}s_\alpha).$$

Moreover, by point 2 of Lemma 3.16, the support of the function $x \mapsto \hat{f}_w(wx)\hat{f}_{s_\alpha}(x^{-1}s_\alpha)$ is contained in $(J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{s_\alpha} J^1_{\mathcal{P}} = (J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} \cap (J^1 \cap \mathcal{U}^{s_\alpha} \cap \mathcal{U}^{-1}) J^1_{\mathcal{P}} = ((J^1_{\mathcal{P}})^w J^1_{\mathcal{P}} \cap J^1 \cap \mathcal{U}_{-\alpha}) J^1_{\mathcal{P}}$ which is equal to $J^1_{\mathcal{P}}$ if $w(-\alpha) < 0$ and to $(J^1 \cap \mathcal{U}_{-\alpha}) J^1_{\mathcal{P}}$ if $w(-\alpha) > 0$. Hence, if $w\alpha > 0$ we obtain $(\hat{f}_w \hat{f}_{s_\alpha})(ws_\alpha) = \boldsymbol{p} \circ \kappa(w) \circ \iota \circ \boldsymbol{p} \circ \kappa(s_\alpha) \circ \iota$ while if $w\alpha < 0$, since $(J^1 \cap \mathcal{U}_{-\alpha})^{w^{-1}}$ and $(J^1 \cap \mathcal{U}_{-\alpha})^{s_\alpha}$ are contained in $J^1 \cap \mathcal{U}$ and so in the kernel of $\eta_{\mathcal{P}}$ and since we have $[(J^1 \cap \mathcal{U}_{-\alpha})J^1_{\mathcal{P}} : J^1_{\mathcal{P}}] = [J^1 \cap \mathcal{U}_{-\alpha} : H^1 \cap \mathcal{U}_{-\alpha}] = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)$, we obtain $(\hat{f}_w \hat{f}_{s_\alpha})(ws_\alpha) = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)\boldsymbol{p} \circ \kappa(w) \circ \iota \circ \boldsymbol{p} \circ \kappa(s_\alpha) \circ \iota$. □

From now on we fix a nonzero element $\gamma \in I_{\tau_{m'-1}}(\eta_{\mathcal{P}})$, which is invertible by Lemma 3.10, and a square root $\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{1/2}$ of $\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)$ in R. We consider the function $\hat{f}_{\tau_{m'-1}} \in \mathscr{H}_R(G, \eta_{\mathcal{P}})_{J^1_{\mathcal{P}}\tau_{m'-1}J^1_{\mathcal{P}}}$ defined by $\hat{f}_{\tau_{m'-1}}(j_1\tau_{m'-1}j_2) = \eta_{\mathcal{P}}(j_1) \circ \gamma \circ \eta_{\mathcal{P}}(j_2)$ for every $j_1, j_2 \in J^1_{\mathcal{P}}$ and the subgroup $\widetilde{\mathcal{W}}$ of $\mathrm{Aut}_R(V_{\mathcal{M}})$ generated by γ and by $\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{1/2}\boldsymbol{p} \circ \kappa(s_i) \circ \iota$ with $i \in \{1, \ldots, m'-1\}$.

Lemma 3.27. *The function that maps s_i to $\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{1/2}\boldsymbol{p} \circ \kappa(s_i) \circ \iota$ for every $i \in \{1, \ldots, m'-1\}$ and $\tau_{m'-1}$ to γ extends to a surjective group homomorphism $\varepsilon : \tilde{W} \to \widetilde{\mathcal{W}}$.*

Proof. Let $\delta = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)$. To prove that ε is a group homomorphism we use the presentation of $\tilde{W}$ given in Remark 3.24. For every $i, j \in \{1, \ldots, m'-1\}$ such that $|i - j| > 1$ we have $\varepsilon(s_i)\varepsilon(s_j) = \delta\boldsymbol{p} \circ \kappa(s_i) \circ \iota \circ \boldsymbol{p} \circ \kappa(s_j) \circ \iota$ which, by Lemma 3.26, is equal to $\delta\boldsymbol{p} \circ \kappa(s_is_j) \circ \iota = \delta\boldsymbol{p} \circ \kappa(s_js_i) \circ \iota = \varepsilon(s_j)\varepsilon(s_i)$. For every $i \neq m'-1$ we have $\varepsilon(s_i)\varepsilon(s_{i+1})\varepsilon(s_i) = \delta^{3/2}\boldsymbol{p} \circ \kappa(s_i) \circ \iota \circ \boldsymbol{p} \circ \kappa(s_{i+1}) \circ \iota \circ \boldsymbol{p} \circ \kappa(s_i) \circ \iota$ which, by Lemma 3.26, is equal to $\delta^{3/2}\boldsymbol{p} \circ \kappa(s_is_{i+1}s_i) \circ \iota = \delta^{3/2}\boldsymbol{p} \circ \kappa(s_{i+1}s_is_{i+1}) \circ \iota = \varepsilon(s_{i+1})\varepsilon(s_i)\varepsilon(s_{i+1})$. For every i we have $\varepsilon(s_i)^2 = \delta\boldsymbol{p} \circ \kappa(s_i) \circ \iota \circ \boldsymbol{p} \circ \kappa(s_i) \circ \iota$ which, by Lemma 3.26, is equal to $\boldsymbol{p} \circ \kappa(s_is_i) \circ \iota$ which is the identity of $\mathrm{Aut}_R(V_{\mathcal{M}})$. Let $\tau = \tau_{m'-1}$ and $\hat{f}_\tau = \hat{f}_{\tau_{m'-1}}$. For every $i \neq m'-1$ we have $\varepsilon(\tau)\varepsilon(s_i) = \delta^{1/2}\gamma \circ \boldsymbol{p} \circ \kappa(s_i) \circ \iota$ which is equal to $\delta^{1/2}(\hat{f}_\tau \hat{f}_{s_i})(\tau s_i)$ since the support of $x \mapsto \hat{f}_\tau(\tau x)\hat{f}_{s_i}(x^{-1}s_i)$ is contained in $(J^1_{\mathcal{P}})^\tau J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{s_i} J^1_{\mathcal{P}} = ((J^1_{\mathcal{P}} \cap \mathcal{U}^-_{P(\tau)})^\tau J^1_{\mathcal{P}} \cap J^1_{\mathcal{P}} \cap \mathcal{U}_{\alpha_{i+1,i}}) J^1_{\mathcal{P}} = J^1_{\mathcal{P}}$. Hence, by Lemma 3.9 we have $\varepsilon(\tau)\varepsilon(s_i) = \delta^{1/2}\boldsymbol{p} \circ \zeta^{-1}(\hat{f}_\tau)(\tau) \circ \kappa(s_i) \circ \iota$. Since $\zeta^{-1}(\hat{f}_\tau)(\tau) \in I_\tau(\eta) = I_\tau(\kappa)$ and $s_i \in J \cap J^\tau$ we obtain $\varepsilon(\tau)\varepsilon(s_i) = \delta^{1/2}\boldsymbol{p} \circ \kappa(s_i) \circ \zeta^{-1}(\hat{f}_\tau)(\tau) \circ \iota = \delta^{1/2}(\hat{f}_{s_i}\hat{f}_\tau)(s_i\tau)$, which is equal to $\delta^{1/2}\boldsymbol{p} \circ \kappa(s_i) \circ \iota \circ \gamma = \varepsilon(s_i)\varepsilon(\tau)$ since the support of $x \mapsto \hat{f}_{s_i}(s_ix)\hat{f}_\tau(x^{-1}\tau)$ is contained in $(J^1_{\mathcal{P}})^{s_i} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{\tau^{-1}} J^1_{\mathcal{P}} = (J^1_{\mathcal{P}} \cap \mathcal{U}_{\alpha_{i+1,i}} \cap (J^1_{\mathcal{P}} \cap \mathcal{U}^+_{P(\tau)})^{\tau^{-1}} J^1_{\mathcal{P}}) J^1_{\mathcal{P}} = J^1_{\mathcal{P}}$. It remains to prove the last relation. Let $s = s_{m'-1}$ and $\tau = \tau_{m'-1}$. Then $\tau s \tau s = \tau_{m'-2} = s \tau s \tau$ and by Lemma 3.9 we have $(\hat{f}_\tau \hat{f}_s \hat{f}_\tau \hat{f}_s)(\tau s \tau s) = \boldsymbol{p} \circ \zeta^{-1}(\hat{f}_\tau \hat{f}_s \hat{f}_\tau)(\tau s \tau) \circ \kappa(s) \circ \iota$. Now, since $\zeta^{-1}(\hat{f}_\tau \hat{f}_s \hat{f}_\tau)(\tau s \tau) \in I_{\tau s \tau}(\kappa)$ and $s = s^{\tau s \tau} \in J \cap J^{\tau s \tau}$, we obtain $(\hat{f}_\tau \hat{f}_s \hat{f}_\tau \hat{f}_s)(\tau_{m'-2}) = \boldsymbol{p} \circ \kappa(s) \circ \zeta^{-1}(\hat{f}_\tau \hat{f}_s \hat{f}_\tau)(\tau s \tau) \circ \iota = (\hat{f}_s \hat{f}_\tau \hat{f}_s \hat{f}_\tau)(\tau_{m'-2})$. On the other hand we have

$$(\hat{f}_\tau \hat{f}_s \hat{f}_\tau \hat{f}_s)(\tau_{m'-2}) = (\hat{f}_\tau \hat{f}_s \hat{f}_\tau \hat{f}_s)(\tau s \tau s) = \sum_{x \in G/J^1_{\mathcal{P}}} \hat{f}_\tau(\tau x)(\hat{f}_s \hat{f}_\tau \hat{f}_s)(x^{-1}s\tau s).$$

The support of $x \mapsto \hat{f}_\tau(\tau x)(\hat{f}_s \hat{f}_\tau \hat{f}_s)(x^{-1}s\tau s)$ is in $(H^1 \cap \mathcal{U}_{\alpha'})^\tau J^1_{\mathcal{P}}$ with $\alpha' = \alpha_{m',m'-1}$ by Lemma 3.23. For every $x \in (H^1 \cap \mathcal{U}_{\alpha'})^\tau$ the elements $x^{\tau^{-1}}$ and $(x^{-1})^{s\tau s}$ are in $H^1 \cap \mathcal{U}$ and so in the kernel of $\eta_{\mathcal{P}}$. Then $(\hat{f}_\tau \hat{f}_s \hat{f}_\tau \hat{f}_s)(\tau_{m'-2}) = (\hat{f}_s \hat{f}_\tau \hat{f}_s \hat{f}_\tau)(\tau_{m'-2})$ is equal to $\delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0)\gamma \circ (\hat{f}_s \hat{f}_\tau \hat{f}_s)(s\tau s)$ and by Lemma 3.25 it is also equal to $\delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0)\varepsilon(\tau)\varepsilon(s)\varepsilon(\tau)\varepsilon(s)$. Now, if $\alpha'' = \alpha_{m'-2,m'-1}$ then $\alpha' \notin \boldsymbol{\Psi}^+_{\widehat{\alpha''}} \cup \boldsymbol{\Psi}^-_{\widehat{\alpha''}}$ and so we have $(J^1_{\mathcal{P}})^s J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{\tau^{-1}_{m'-2}} J^1_{\mathcal{P}} = J^1_{\mathcal{P}} = (J^1_{\mathcal{P}})^{\tau_{m'-2}} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^s J^1_{\mathcal{P}}$. Hence, $(\hat{f}_s \hat{f}_\tau \hat{f}_s \hat{f}_\tau \hat{f}_s)(s\tau s\tau s)$ is equal both to

$$\hat{f}_s(s) \circ (\hat{f}_\tau \hat{f}_s \hat{f}_\tau \hat{f}_s)(\tau_{m'-2}) = \delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0)\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{-1/2}\varepsilon(s)\varepsilon(\tau)\varepsilon(s)\varepsilon(\tau)\varepsilon(s)$$

and also to

$$\begin{aligned}(\hat{f}_s \hat{f}_\tau \hat{f}_s \hat{f}_\tau)(\tau_{m'-2}) \circ \hat{f}_s(s) &= \delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0)\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{-1/2}\varepsilon(\tau)\varepsilon(s)\varepsilon(\tau)\varepsilon(s)^2 \\ &= \delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0)\delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{-1/2}\varepsilon(\tau)\varepsilon(s)\varepsilon(\tau).\end{aligned}$$

This implies $\varepsilon(\tau)\varepsilon(s)\varepsilon(\tau)\varepsilon(s) = \varepsilon(s)\varepsilon(\tau)\varepsilon(s)\varepsilon(\tau)$ since both $\delta(\mathfrak{H}^1_0, \varpi\mathfrak{H}^1_0)$ and $\delta^{-1/2}$ are invertible in R. We conclude that ε is a group homomorphism and it is clearly surjective. □

Remark 3.28. For every $w \in W$ we have $\varepsilon(w) = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{\ell(w)/2} \boldsymbol{p} \circ \kappa(w) \circ \iota$.

Lemma 3.29. *For every $\tilde{w} \in \tilde{W}$ we have $\varepsilon(\tilde{w}) \in I_{\tilde{w}}(\eta_{\mathcal{P}})$.*

Proof. Since $\eta_{\mathcal{M}}$ is the restriction of $\eta_{\mathcal{P}}$ to the group $J^1_{\mathcal{M}}$, we have $\varepsilon(w) = \delta(\mathfrak{J}^1_0, \mathfrak{H}^1_0)^{\ell(w)/2} \hat{f}_w(w) \in I_w(\eta_{\mathcal{M}})$ for every $w \in W$ and $\gamma \in I_{\tau_{m'-1}}(\eta_{\mathcal{M}})$. Then, since every $w \in W$ and $\tau_{m'-1}$ normalize $J^1_{\mathcal{M}}$, we have $\varepsilon(\tilde{w}) \in I_{\tilde{w}}(\eta_{\mathcal{M}})$ for every $\tilde{w} \in \tilde{W}$ and so $\varepsilon(\tilde{w}) \in I_{\tilde{w}}(\eta_{\mathcal{P}})$ by Lemma 3.10. □

Lemma 3.30. *For every $\tau', \tau'' \in \boldsymbol{\Delta}$, $\gamma' \in I_{\tau'}(\eta_{\mathcal{P}})$ and $\gamma'' \in I_{\tau''}(\eta_{\mathcal{P}})$ we have $\gamma' \circ \gamma'' = \gamma'' \circ \gamma'$.*

Proof. We recall that for every $\tau \in \boldsymbol{\Delta}$ the vector space $I_\tau(\eta_{\mathcal{P}})$ is 1-dimensional and so there exist elements $c', c'' \in R$ such that $\gamma' = c'\varepsilon(\tau')$ and $\gamma'' = c''\varepsilon(\tau'')$. We obtain $\gamma' \circ \gamma'' = c'c''\varepsilon(\tau') \circ \varepsilon(\tau'') = c'c''\varepsilon(\tau'\tau'') = c'c''\varepsilon(\tau''\tau') = \gamma'' \circ \gamma'$. □

3G. *The isomorphisms $\mathscr{H}_R(G, \eta_P) \cong \mathscr{H}_R(B^\times, K^1_B)$.* In this section we define the elements $\hat{f}_{\tau_i} \in \mathscr{H}_R(G, \eta_{\mathcal{P}})_{J^1_{\mathcal{P}}\tau_i J^1_{\mathcal{P}}}$ for every $i \in \{0, \dots, m'-1\}$ and we prove that $\hat{f}_\omega$ with $\omega \in \Omega$ respect the relations of Lemma 3.14 obtaining an algebra homomorphism from $\mathscr{H}_R(B^\times, K^1_B)$ to $\mathscr{H}_R(G, \eta_{\mathcal{P}})$.

For every $i \in \{0, \dots, m'-1\}$ we put $\gamma_i = \partial^{(m'-i)(m'-i-1)/2}\varepsilon(\tau_i)$ where ∂ is the power of p defined in Remark 3.22. Then γ_i is an invertible element in $I_{\tau_i}(\eta_{\mathcal{P}})$ and $\gamma_{m'-1} = \gamma$.

Lemma 3.31. *We have, for every $i \in \{1, \dots, m'-1\}$,*

$$\gamma_{i-1} \circ \gamma_i^{-1} = \partial^{m'-i}\varepsilon(\tau_{i-1}\tau_i^{-1}) \quad \textit{and} \quad \gamma_i = \prod_{h=i+1}^{m'} \partial^{m'-h}\varepsilon(\tau_i).$$

Proof. Since $((m'-(i-1))(m'-(i-1)-1) - (m'-i)(m'-i-1))/2 = m'-i$ we have that $\gamma_{i-1} \circ \gamma_i^{-1} = \partial^{m'-i}\varepsilon(\tau_{i-1})\varepsilon(\tau_i)^{-1} = \partial^{m'-i}\varepsilon(\tau_{i-1}\tau_i^{-1})$. The second statement is true because

$$\sum_{h=i+1}^{m'} m'-h = \sum_{j=0}^{m'-i-1} j = \frac{(m'-i)(m'-i-1)}{2}. \qquad □$$

For every $i \in \{0, \dots, m'-1\}$ we consider the function $\hat{f}_{\tau_i} \in \mathscr{H}_R(G, \eta_{\mathcal{P}})_{J^1_{\mathcal{P}}\tau_i J^1_{\mathcal{P}}}$ defined by $\hat{f}_{\tau_i}(j_1\tau_i j_2) = \eta_{\mathcal{P}}(j_1) \circ \gamma_i \circ \eta_{\mathcal{P}}(j_2)$ for every $j_1, j_2 \in J^1_{\mathcal{P}}$. We remark that in general $\hat{f}_{\tau_i}$ is not invertible but since τ_0 normalizes $J^1_{\mathcal{P}}$ the function $\hat{f}_{\tau_0}$ is invertible in $\mathscr{H}_R(G, \eta_{\mathcal{P}})$ with inverse $\hat{f}_{\tau_0^{-1}} : \tau_0^{-1}J^1_{\mathcal{P}} \to \mathrm{End}_R(V_{\mathcal{M}})$ defined by $\hat{f}_{\tau_0^{-1}}(\tau_0^{-1}j) = \gamma_0^{-1} \circ \eta_{\mathcal{P}}(j)$ for every $j \in J^1_{\mathcal{P}}$.

Lemma 3.32. *The map $\Theta'' : \mathbf{\Omega} \to \mathscr{H}_R(G, \eta_P)$ given by $f_\omega \mapsto \hat{f}_\omega$ for every $f_\omega \in \mathbf{\Omega}$ is well defined.*

Proof. The map is well defined on f_k with $k \in K_B$ because Θ' is a homomorphism and it is well defined on τ_i with $i \in \{0, \dots, m'-1\}$ because $K_B^1 \tau_i K_B^1 = K_B^1 \tau_j K_B^1$ implies $i = j$. □

Lemma 3.33. *The function $\hat{f}_{\tau_i} \hat{f}_{\tau_j}$ is in $\mathscr{H}_R(G, \eta_{\mathcal{P}})_{J^1_{\mathcal{P}} \tau_i \tau_j J^1_{\mathcal{P}}}$ and $(\hat{f}_{\tau_i} \hat{f}_{\tau_j})(\tau_i \tau_j) = \gamma_i \circ \gamma_j$, for every $i, j \in \{0, \dots, m'-1\}$*

Proof. If i or j is 0 then the result follows from Lemma 3.9 since τ_0 normalizes $J^1_{\mathcal{P}}$. Otherwise, by point 4 of Lemma 3.16 the support of $\hat{f}_{\tau_i} \hat{f}_{\tau_j}$ is contained in $J^1_{\mathcal{P}} \tau_i J^1_{\mathcal{P}} \tau_j J^1_{\mathcal{P}} = J^1_{\mathcal{P}} \tau_i \tau_j J^1_{\mathcal{P}}$ and the support of $x \mapsto \hat{f}_{\tau_i}(\tau_i x) \hat{f}_{\tau_j}(x^{-1} \tau_j)$ is contained in $(J^1_{\mathcal{P}})^{\tau_i} J^1_{\mathcal{P}} \cap (J^1_{\mathcal{P}})^{\tau_j^{-1}} J^1_{\mathcal{P}} = J^1_{\mathcal{P}}$. Hence, we obtain $(\hat{f}_{\tau_i} \hat{f}_{\tau_j})(\tau_i \tau_j) = \sum_{x \in G/J^1_{\mathcal{P}}} \hat{f}_{\tau_i}(\tau_i x) \hat{f}_{\tau_j}(x^{-1} \tau_j) = \hat{f}_{\tau_i}(\tau_i) \circ \hat{f}_{\tau_j}(\tau_j) = \gamma_i \circ \gamma_j$. □

By Lemmas 3.33 and 3.30 we obtain $\hat{f}_{\tau_i} \hat{f}_{\tau_j} = \hat{f}_{\tau_j} \hat{f}_{\tau_i}$ for every $i, j \in \{0, \dots, m'-1\}$. So, if $P \subset \{0, \dots, m'-1\}$ we denote by γ_P the composition of γ_i with $i \in P$, which is well defined by Lemma 3.30, and by $\hat{f}_{\tau_P}$ the product of $\hat{f}_{\tau_i}$ with $i \in P$, which is well defined because the $\hat{f}_{\tau_i}$ commute. Furthermore, by point 4 of Lemma 3.16 we obtain that the support of $\hat{f}_{\tau_P}$ is $J^1_{\mathcal{P}} \tau_P J^1_{\mathcal{P}}$ and by Lemma 3.33 we have $\hat{f}_{\tau_P}(\tau_P) = \gamma_P$.

Lemma 3.34. *We have $\hat{f}_{\tau_i} \hat{f}_x = \hat{f}_{\tau_i x \tau_i^{-1}} \hat{f}_{\tau_i}$ for every $i \in \{0, \dots, m'-1\}$ and every $x \in M_{\widehat{\alpha_{i,i+1}}} = K_B \cap \mathcal{M}_{\widehat{\alpha_{i,i+1}}}$ if $i \neq 0$ or $x \in K_B$ if $i = 0$.*

Proof. Since x normalizes J^1, by Lemma 3.9 the supports of $\hat{f}_{\tau_i} \hat{f}_x$ and of $\hat{f}_{\tau_i x \tau_i^{-1}} \hat{f}_{\tau_i}$ are contained in $J^1_{\mathcal{P}} \tau_i x J^1_{\mathcal{P}}$ and $(\hat{f}_{\tau_i} \hat{f}_x)(\tau_i x) = \boldsymbol{p} \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \kappa(x) \circ \iota$, which is equal to $\boldsymbol{p} \circ \kappa(\tau_i x \tau_i^{-1}) \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \iota = (\hat{f}_{\tau_i x \tau_i^{-1}} \hat{f}_{\tau_i})(\tau_i x)$ because $\zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \in I_{\tau_i}(\kappa)$ and $x \in J \cap J^{\tau_i}$. □

Lemma 3.35. *Let $i \in \{1, \dots, m'-1\}$ and $\alpha \in \Psi^+_{\widehat{\alpha_{i,i+1}}}$. Then for every $u \in U_\alpha$ and $u' \in U_{-\alpha}$ we have $\hat{f}_u \hat{f}_{\tau_i} = \hat{f}_{\tau_i}$ and $\hat{f}_{\tau_i} \hat{f}_{u'} = \hat{f}_{\tau_i}$.*

Proof. The elements $\tau_i^{-1} u \tau_i$ and $\tau_i u' \tau_i^{-1}$ are in $K_B^1 \subset J^1_{\mathcal{P}}$ and so, since u and u' normalize J^1, by Lemma 3.9 the supports of $\hat{f}_u \hat{f}_{\tau_i}$ and of $\hat{f}_{\tau_i} \hat{f}_{u'}$ are in $J^1_{\mathcal{P}} u \tau_i J^1_{\mathcal{P}} = J^1_{\mathcal{P}} \tau_i J^1_{\mathcal{P}} = J^1_{\mathcal{P}} \tau_i u' J^1_{\mathcal{P}}$. Now since $\zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \in I_{\tau_i}(\eta) = I_{\tau_i}(\kappa)$ and $u \in J \cap J^{\tau_i^{-1}}$, by Lemma 3.9 we have $(\hat{f}_u \hat{f}_{\tau_i})(u \tau_i) = \boldsymbol{p} \circ \kappa(u) \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \iota = \boldsymbol{p} \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \eta(\tau_i^{-1} u \tau_i) \circ \iota$. By Lemma 3.7 we obtain $(\hat{f}_u \hat{f}_{\tau_i})(u \tau_i) = \boldsymbol{p} \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \iota \circ \eta_{\mathcal{P}}(\tau_i^{-1} u \tau_i) = \hat{f}_{\tau_i}(\tau_i) \circ \eta_{\mathcal{P}}(\tau_i^{-1} u \tau_i) = \hat{f}_{\tau_i}(u \tau_i)$. Similarly we have $\hat{f}_{\tau_i}(\tau_i u') = \boldsymbol{p} \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \kappa(u') \circ \iota = \boldsymbol{p} \circ \eta(\tau_i u' \tau_i^{-1}) \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \iota$ which is equal to $\eta_{\mathcal{P}}(\tau_i u' \tau_i^{-1}) \circ \boldsymbol{p} \circ \zeta^{-1}(\hat{f}_{\tau_i})(\tau_i) \circ \iota = \eta_{\mathcal{P}}(\tau_i u' \tau_i^{-1}) \circ \hat{f}_{\tau_i}(\tau_i) = \hat{f}_{\tau_i}(\tau_i u')$. □

We introduce some subgroups of G, through its identification with $\mathrm{GL}_{m'}(A(E))$, in order to find the support of $\hat{f}_{\tau_P} \hat{f}_w \hat{f}_{\tau_\alpha} \hat{f}_{w^{-1}}$. We recall that $\mathfrak{A}(E)$ is the unique hereditary order normalized by $E^\times$ in $A(E)$ and $\mathfrak{P}(E)$ is its radical.

- Let $\mathcal{Z}$ be the set of matrices (z_{ij}) such that $z_{ii} = 1$, $z_{ij} \in \varpi^{-1}\mathfrak{P}(E)$ if $i < j$ and $z_{ij} = 0$ if $i > j$.
- Let $\mathcal{V}$ be the group $(J^1 \cap w\mathcal{U}^-_{\hat{\alpha}} w^{-1} \cap \mathcal{U}^+_{\hat{P}})^{w\tau_\alpha w^{-1}} = \prod_{\alpha' \in w\Psi^-_{\hat{\alpha}} \cap \Psi^+_{\hat{P}}} (\mathbb{I}_{m'} + \varpi^{-1}\mathfrak{J}^1) \cap \mathcal{U}_{\alpha'} \subset \mathcal{Z}$. We remark that it is different from $\mathcal{V}(w, \alpha)$ defined by (4).
- Let $\tilde{I}^1$ be the group of matrices (m_{ij}) such that $m_{ii} \in 1 + \mathfrak{P}(E)$, $m_{ij} \in \mathfrak{A}(E)$ if $i < j$ and $m_{ij} \in \mathfrak{P}(E)$ if $i > j$.

- Let $\boldsymbol{W} = W \ltimes M$ be the subgroup of $B^\times$ of monomial matrices with coefficients in $\mathcal{O}_{D'}^\times$. Then $B^\times$ is the disjoint union of $I_B(1) w I_B(1)$ with $w \in \boldsymbol{W}$, where $I_B(1) = K^1 U$ is the standard *pro-p-Iwahori subgroup of K_B*, i.e., the pro-p-radical of I_B.

Lemma 3.36. *We have $J^1_{\mathcal{P}} \tau_P J^1_{\mathcal{P}} w \tau_\alpha w^{-1} J^1_{\mathcal{P}} = J^1_{\mathcal{P}} \tau_Q \mathcal{V} J^1_{\mathcal{P}}$.*

Proof. We proceed in a similar way to the beginning of the proof of Lemma 3.23: we can prove that $J^1_{\mathcal{P}} w \tau_\alpha w^{-1} J^1_{\mathcal{P}} = (J^1_{\mathcal{P}} \cap w \mathcal{U}^-_{\hat{\alpha}} w^{-1}) w \tau_\alpha w^{-1} J^1_{\mathcal{P}}$. Now we consider the decomposition of the group $(J^1_{\mathcal{P}} \cap w \mathcal{U}^-_{\hat{\alpha}} w^{-1})$ into the product $(J^1_{\mathcal{P}} \cap w \mathcal{U}^-_{\hat{\alpha}} w^{-1} \cap \mathcal{U}^-)(J^1_{\mathcal{P}} \cap w \mathcal{U}^-_{\hat{\alpha}} w^{-1} \cap \mathcal{U})$. By Lemma 3.15 we have $(J^1_{\mathcal{P}} \cap w \mathcal{U}^-_{\hat{\alpha}} w^{-1} \cap \mathcal{U}^-)^{\tau_P^{-1}} \subset J^1_{\mathcal{P}}$ and by Lemma 3.18 we have $J^1_{\mathcal{P}} \cap w \mathcal{U}^-_{\hat{\alpha}} w^{-1} \cap \mathcal{U} = J^1_{\mathcal{P}} \cap w \mathcal{U}^-_{\hat{\alpha}} w^{-1} \cap \mathcal{U}^+_{\hat{P}}$. □

Lemma 3.37. *Let $\tau \in \boldsymbol{\Delta}$. If $z \in \mathcal{Z}$ is such that $\tilde{I}^1 \tau z \tilde{I}^1 \cap \boldsymbol{W} \neq \varnothing$ then $\tilde{I}^1 \tau z \tilde{I}^1 \cap \boldsymbol{W} = \{\tau\}$.*

Proof. For every $r \in \{1, \dots, m'\}$ we denote by $\boldsymbol{\Delta}_{(r)}$, $\mathcal{Z}_{(r)}$, $\tilde{I}^1_{(r)}$ and $\boldsymbol{W}_{(r)}$ the subsets of $\mathrm{GL}_r(A(E))$ similar to those defined for $\mathrm{GL}_{m'}(A(E))$. We prove the statement of the lemma by induction on r. If $r = 1$ we have $\boldsymbol{\Delta}_{(1)} = \varpi^{\mathbb{Z}}$, $\mathcal{Z}_{(1)} = \{1\}$, $\tilde{I}^1_{(1)} = 1 + \mathfrak{P}(E)$ and $\boldsymbol{W}_{(1)} = \varpi^{\mathbb{Z}}$ and we have $(1 + \mathfrak{P}(E)) \varpi^a (1 + \mathfrak{P}(E)) \cap \varpi^{\mathbb{Z}} = \varpi^a (1 + \mathfrak{P}(E)) \cap \varpi^{\mathbb{Z}} = \{\varpi^a\}$ for every $a \in \mathbb{Z}$. Now we suppose the statement true for every $r < m'$. Let $x, y \in \tilde{I}^1$ such that $x \tau z y \in \boldsymbol{W}$. We proceed by steps.

First step: We consider the decomposition $\tilde{I}^1 = (\tilde{I}^1 \cap \mathcal{U}^-)(\tilde{I}^1 \cap \mathcal{U})(\tilde{I}^1 \cap \mathcal{M})$ and we write $x = x_1 x_2 x_3$ with $x_1 \in \tilde{I}^1 \cap \mathcal{U}^-$, $x_2 \in \tilde{I}^1 \cap \mathcal{U}$ and $x_3 \in \tilde{I}^1 \cap \mathcal{M}$. Then we have

$$x \tau z y = x_1 \tau ((\tau^{-1} x_2 \tau)(\tau^{-1} x_3 \tau) z (\tau^{-1} x_3^{-1} \tau))(\tau^{-1} x_3 \tau) y.$$

We observe that $\tau^{-1} x_3 \tau$ is a diagonal matrix with coefficients in $1 + \mathfrak{P}(E)$ and the conjugate of z by this element is in $\mathcal{Z}$. Moreover, $\tau^{-1} x_2 \tau$ is in $\tilde{I}^1 \cap \mathcal{U}$ and if we multiply it by an element of $\mathcal{Z}$ we obtain another element of $\mathcal{Z}$. If we set $z_1 = \tau^{-1} x_2 x_3 \tau z \tau^{-1} x_3^{-1} \tau \in \mathcal{Z}$ then $\tilde{I}^1 \tau z \tilde{I}^1 = \tilde{I}^1 \tau z_1 \tilde{I}^1$ and $(\tilde{I}^1 \cap \mathcal{U}^-) \tau z_1 \tilde{I}^1 \cap \boldsymbol{W} \neq \varnothing$. Hence, we can suppose $x \in \tilde{I}^1 \cap \mathcal{U}^-$.

Second step: Let $a_1 \leq \cdots \leq a_{m'} \in \mathbb{N}$ such that $\tau = \mathrm{diag}(\varpi^{a_i})$ and let $s \in \mathbb{N}^*$ such that $a_1 = \cdots = a_s$ and $a_1 < a_{s+1}$. We want to prove $z_{ij} \in \mathfrak{A}(E)$ for every $i \in \{1, \dots, s\}$ so we assume the opposite and we look for a contradiction. Let v be the valuation on $A(E)$ associated to $\mathfrak{P}(E)$ and let

$$\begin{aligned} b &= \min\{\mathrm{v}(\varpi^{a_1} z_{ij}) \mid 1 \leq i \leq s, 1 \leq j \leq m'\}, \\ k &= \min\{1 \leq j \leq m' \mid \text{there exists } z_{ij} \text{ with } 1 \leq i \leq s \text{ such that } \mathrm{v}(\varpi^{a_1} z_{ij}) = b\}. \end{aligned}$$

Let $1 \leq h \leq s$ be such that $\mathrm{v}(\varpi^{a_1} z_{hk}) = b$. By hypothesis the element z_{hk} is not in $\mathfrak{A}(E)$ and so $h < k$ and

$$(a_1 - 1)\mathrm{v}(\varpi) < b < a_1 \mathrm{v}(\varpi). \tag{5}$$

We observe that for every $i \in \{1, \dots, m'\}$ and $j > i$ we have $\mathrm{v}(\varpi^{a_i} z_{ij}) \geq b$: if $i \leq s$ by definition of b and if $i > s$ because $\mathrm{v}(\varpi^{a_i} z_{ij}) = a_i \mathrm{v}(\varpi) + \mathrm{v}(z_{ij}) > (a_i - 1)\mathrm{v}(\varpi) \geq a_1 \mathrm{v}(\varpi) > b$. We consider the coefficient at position (h,k) of $x\tau zy$ which is equal to

$$\sum_{e=1}^{m'}\sum_{f=1}^{m'} x_{he}\varpi^{a_e} z_{ef} y_{fk} = \sum_{e=1}^{h}\sum_{f=e}^{m'} x_{he}\varpi^{a_1} z_{ef} y_{fk},$$

since $x_{he}=0$ if $e>h$ and $z_{ef}=0$ if $f<e$. Now,

- if $e=h$ and $f=k$ then $\mathrm{v}(x_{hh}\varpi^{a_1} z_{hk} y_{kk}) = b$ because $x_{hh}=1$, and $y_{kk}\in 1+\mathfrak{P}(E)$;
- if $e=h$ and $f<k$ then $\mathrm{v}(x_{hh}\varpi^{a_1} z_{hf} y_{fk}) > b$ by definition of k;
- if $e=h$ and $f>k$ then $\mathrm{v}(x_{hh}\varpi^{a_1} z_{hf} y_{fk}) > b$ because $y_{fk}\in\mathfrak{P}(E)$;
- if $e<h$ then $\mathrm{v}(x_{he}\varpi^{a_1} z_{ef} y_{fk}) > b$ because $x_{he}\in\mathfrak{P}(E)$.

We obtain an element of valuation b. Then b must be a multiple of $\mathrm{v}(\varpi)$ because $x\tau zy\in\boldsymbol{W}$ but this in contradiction with (5). Hence, $z_{ij}\in\mathfrak{A}(E)$ for every $i\in\{1,\dots,s\}$. Now, we can write $z=z'z''$ with $z'_{ii}=1$ for all i, $z'_{ij}=z_{ij}$ if $i\in\{s+1,\dots,m'\}$ and $j>i$ and $z'_{ij}=0$ otherwise and $z''_{ii}=1$ for all i, $z''_{ij}=z_{ij}$ if $i\in\{1,\dots,s\}$ and $j>i$ and $z''_{ij}=0$ otherwise. Then $z''\in\tilde{I}^1$ and so $\tilde{I}^1\tau z\tilde{I}^1=\tilde{I}^1\tau z'\tilde{I}^1$ and $(\tilde{I}^1\cap\mathcal{U}^-)\tau z'\tilde{I}^1\cap\boldsymbol{W}\neq\varnothing$. Then we can suppose z of the form $\left(\begin{smallmatrix}\mathbb{I}_s & 0\\ 0 & \hat{z}\end{smallmatrix}\right)$ with $\hat{z}\in\mathcal{Z}_{(m'-s)}$.

Third step: We write $x=x'x''$ with $x'_{ii}=1$ for all i, $x'_{ij}=x_{ij}$ if $i\in\{s+1,\dots,m'\}$ and $j<i$ and $x'_{ij}=0$ otherwise and $x''_{ii}=1$ for all i, $x''_{ij}=x_{ij}$ if $i\in\{1,\dots,s\}$ and $j<i$ and $x''_{ij}=0$ otherwise. Then $\tau^{-1}x''\tau\in\tilde{I}^1$ and it commutes with z. Then we can suppose x is of the form $\left(\begin{smallmatrix}\mathbb{I}_s & 0\\ x''' & \hat{x}\end{smallmatrix}\right)$ with $x'''\in M_{(m'-s)\times s}(\mathfrak{P}(E))$ and $\hat{x}\in\tilde{I}^1_{(m'-s)}$.

Fourth step: Let $\tau=\left(\begin{smallmatrix}\varpi^{a_1}\mathbb{I}_s & 0\\ 0 & \hat{\tau}\end{smallmatrix}\right)$ with $\hat{\tau}\in\boldsymbol{\Delta}_{(m'-s)}$ and $y=\left(\begin{smallmatrix}y_1 & y_2\\ y_3 & \hat{y}\end{smallmatrix}\right)$ with $y_1\in\tilde{I}^1_{(s)}$, $y_2\in M_{s\times(m'-s)}(\mathfrak{A}(E))$, $y_3\in M_{(m'-s)\times s}(\mathfrak{P}(E))$ and $\hat{y}\in\tilde{I}^1_{(m'-s)}$. Then the product $x\tau zy$ is

$$\begin{pmatrix}\mathbb{I}_s & 0\\ x''' & \hat{x}\end{pmatrix}\begin{pmatrix}\varpi^{a_1}\mathbb{I}_s & 0\\ 0 & \hat{\tau}\end{pmatrix}\begin{pmatrix}\mathbb{I}_s & 0\\ 0 & \hat{z}\end{pmatrix}\begin{pmatrix}y_1 & y_2\\ y_3 & \hat{y}\end{pmatrix}=\begin{pmatrix}\varpi^{a_1}y_1 & \varpi^{a_1}y_2\\ t & x'''\varpi^{a_1}y_2+\hat{x}\hat{\tau}\hat{z}\hat{y}\end{pmatrix}$$

where $t=x'''\varpi^{a_1}y_1+\hat{x}\hat{\tau}\hat{z}y_3$. Since $x\tau zy$ is in $\boldsymbol{W}$ and since $y_1\in\tilde{I}^1_{(s)}$ is invertible then $\varpi^{a_1}y_1$ must be in $\boldsymbol{W}_{(s)}$ and so $y_1=\mathbb{I}_s$. This also implies $\varpi^{a_1}y_2=t=0$ since $x\tau zy$ is a monomial matrix and so $x\tau zy=\left(\begin{smallmatrix}\varpi^{a_1}\mathbb{I}_s & 0\\ 0 & \hat{x}\hat{\tau}\hat{z}\hat{y}\end{smallmatrix}\right)$ with $\hat{x}\hat{\tau}\hat{z}\hat{y}\in\boldsymbol{W}_{(m'-s)}$. Now, since $\tilde{I}^1_{(m'-s)}\hat{\tau}\hat{z}\tilde{I}^1_{(m'-s)}\cap\boldsymbol{W}_{(m'-s)}\neq\varnothing$, by the inductive hypothesis we have $\hat{x}\hat{\tau}\hat{z}\hat{y}=\hat{\tau}$ and so $x\tau zy=\tau$. □

Lemma 3.38. *We have* $J^1_{\mathcal{P}}\tau_P J^1_{\mathcal{P}} w\tau_\alpha w^{-1} J^1_{\mathcal{P}}\cap J^1_{\mathcal{P}}B^\times J^1_{\mathcal{P}} = J^1_{\mathcal{P}}\tau_Q(U\cap wU^-w^{-1})J^1_{\mathcal{P}}$.

Proof. By Lemma 3.36 we have $J^1_{\mathcal{P}}\tau_P J^1_{\mathcal{P}} w\tau_\alpha w^{-1} J^1_{\mathcal{P}} = J^1_{\mathcal{P}}\tau_Q\mathcal{V}J^1_{\mathcal{P}}$. Now, since $\mathfrak{J}^1\subset M_{m'}(\mathfrak{P}(E))$ we have $\mathcal{V}\subset\mathcal{Z}$ and $J^1_{\mathcal{P}}\subset\tilde{I}^1$ and so we obtain

$$\begin{aligned} J^1_{\mathcal{P}}\tau_P J^1_{\mathcal{P}} w\tau_\alpha w^{-1} J^1_{\mathcal{P}}\cap B^\times &\subset \tilde{I}^1\tau_Q\mathcal{Z}\tilde{I}^1\cap K^1_B U\boldsymbol{W}UK^1_B = K^1_B U(\tilde{I}^1\tau_Q\mathcal{Z}\tilde{I}^1\cap\boldsymbol{W})UK^1_B\\ &\text{(Lemma 3.37)} = K^1_B U\tau_Q UK^1_B = K^1_B\tau_Q UK^1_B. \end{aligned}$$

This implies $J^1_{\mathcal{P}}\tau_P J^1_{\mathcal{P}} w\tau_\alpha w^{-1} J^1_{\mathcal{P}}\cap B^\times = J^1_{\mathcal{P}}\tau_Q\mathcal{V}J^1_{\mathcal{P}}\cap K^1_B\tau_Q UK^1_B$. Let now $v\in\mathcal{V}$ be such that $J^1_{\mathcal{P}}\tau_Q vJ^1_{\mathcal{P}}\cap K^1_B\tau_Q UK^1_B\neq\varnothing$. Then $v\in\tau_Q^{-1}J^1_{\mathcal{P}}K^1_B\tau_Q UK^1_B J^1_{\mathcal{P}}\cap\mathcal{V}\subset\tau_Q^{-1}J^1_{\mathcal{P}}\tau_Q UJ^1_{\mathcal{P}}\cap\mathcal{U}$. Now $U=K_B\cap\mathcal{U}\subset J\cap\mathcal{P}$ normalizes $J^1_{\mathcal{P}}$ and so $v\in\tau_Q^{-1}J^1_{\mathcal{P}}\tau_Q J^1_{\mathcal{P}}U\cap\mathcal{U}$ which is in $(\tau_Q^{-1}(J^1_{\mathcal{P}}\cap\mathcal{U}^-_{\hat{Q}})\tau_Q J^1_{\mathcal{P}}\cap\mathcal{U})U$ by point 1 of Lemma 3.16. Hence, by point 3 of Lemma 3.16 we obtain $v\in UJ^1_{\mathcal{P}}\cap\mathcal{V}\subset UJ^1\cap\mathcal{V}$. By Lemma 3.18 we have $U\cap wU^-w^{-1}=U^+_{\hat{P}}\cap wU^-_{\hat{\alpha}}w^{-1}$ and proceeding in a way similar to the proof of Lemma 3.21 we can prove $U^+_{\hat{P}}\cap wU^-_{\hat{\alpha}}w^{-1}\subset\mathcal{V}$. We obtain

$$\begin{aligned}UJ^1\cap\mathcal{V}&=(U\cap wU^-w^{-1})(U\cap wUw^{-1})J^1\cap\mathcal{V}\\&=(U\cap wU^-w^{-1})(J^1(U\cap wUw^{-1})\cap\mathcal{V})\\&=(U\cap wU^-w^{-1})(J^1(w^{-1}Uw\cap U)\cap\mathcal{V}^w)^{w^{-1}}.\end{aligned}$$

By the definition of $\mathcal{V}$ we have $\mathcal{V}^w=(J^1_{\mathcal{P}}\cap w\mathcal{U}^-_{\hat{\alpha}}w^{-1}\cap\mathcal{U}^+_{\hat{P}})^{w\tau_\alpha}\subset(\mathcal{U}^-_{\hat{\alpha}})^{\tau_\alpha}\subset\mathcal{U}^-$ and then $UJ^1\cap\mathcal{V}\subset(U\cap wU^-w^{-1})(J^1\mathcal{U}\cap\mathcal{U}^-)^{w^{-1}}$ which, by Remark 3.17, is equal to $(U\cap wU^-w^{-1})J^1$. Hence v is in $(U\cap wU^-w^{-1})J^1\cap UJ^1_{\mathcal{P}}=(U\cap wU^-w^{-1})(J^1\cap U)J^1_{\mathcal{P}}$ which is contained in $(U\cap wU^-w^{-1})K^1_BJ^1_{\mathcal{P}}=(U\cap wU^-w^{-1})J^1_{\mathcal{P}}$ and so $J^1\tau_PJ^1_{\mathcal{P}}w\tau_\alpha w^{-1}J^1_{\mathcal{P}}\cap J^1_{\mathcal{P}}B^\times J^1_{\mathcal{P}}=J^1_{\mathcal{P}}\tau_Q(U\cap wU^-w^{-1})J^1_{\mathcal{P}}$. □

Lemma 3.39. *For every $u\in U\cap wU^-w^{-1}$ we have*

$$(\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(\tau_Qu)=q^{\ell(w)}d(w,\alpha)\delta(\mathfrak{J}^1_0,\mathfrak{H}^1_0)^{\ell(w)}\gamma_P\circ\boldsymbol{p}\circ\kappa(w)\circ\iota\circ\gamma_i\circ\boldsymbol{p}\circ\kappa(w^{-1})\circ\iota\circ\boldsymbol{p}\circ\kappa(u)\circ\iota.$$

Proof. By Lemma 3.38 the support of $\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}}$ is contained in $J^1_{\mathcal{P}}\tau_Q(U\cap wU^-w^{-1})J^1_{\mathcal{P}}$. Let $u\in U\cap wU^-w^{-1}$. By Lemma 3.18 we have $U\cap wU^-w^{-1}=U^+_{\hat{P}}\cap wU^-_{\hat{\alpha}}w^{-1}$, by Lemma 3.35 we have $\hat{f}_{\tau_\alpha}=\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}uw}$ and by Lemma 3.11 we have $\hat{f}_{w^{-1}uw}\hat{f}_{w^{-1}}=\hat{f}_{w^{-1}}\hat{f}_u$. Since u is in $U=K_B\cap\mathcal{U}\subset J\cap\mathcal{P}$, it normalizes $J^1_{\mathcal{P}}$ and then by Lemma 3.9 we obtain $(\hat{f}_{\tau_P}\hat{f}_w\tilde{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(\tau_Qu)=(\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}}\hat{f}_u)(\tau_Qu)=(\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(\tau_Q)\circ\boldsymbol{p}\circ\kappa(u)\circ\iota$. It remains to calculate

$$(\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(\tau_Q)=\sum_{x\in G/J^1_{\mathcal{P}}}\hat{f}_{\tau_P}(\tau_Px)(\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(x^{-1}w\tau_\alpha w^{-1}).$$

By Lemma 3.23 the support of the function $x\mapsto\hat{f}_{\tau_P}(\tau_Px)(\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(x^{-1}w\tau_\alpha w^{-1})$ is in $\mathcal{V}(w,\alpha)J^1_{\mathcal{P}}$. Now, since for every $x\in\mathcal{V}(w,\alpha)=(J^1_{\mathcal{P}}\cap w\mathcal{U}^+_{\hat{\alpha}}w^{-1}\cap\mathcal{U}^-_{\hat{P}})^{w\tau_\alpha^{-1}w^{-1}}$ we have $(x^{-1})^{w\tau_\alpha w^{-1}}\in J^1_{\mathcal{P}}\cap\mathcal{U}^-$ and $x^{\tau_{\mathcal{P}}^{-1}}\in(J^1_{\mathcal{P}}\cap w\mathcal{U}^+_{\hat{\alpha}}w^{-1}\cap\mathcal{U}^-_{\hat{P}})^{\tau_{\mathcal{Q}}^{-1}}\subset(J^1_{\mathcal{P}}\cap\mathcal{U}^-)^{\tau_{\mathcal{Q}}^{-1}}$ which is in $J^1_{\mathcal{P}}\cap\mathcal{U}^-$ by Lemma 3.15, then $(x^{-1})^{w\tau_\alpha w^{-1}}$ and $x^{\tau_{\mathcal{P}}^{-1}}$ are in the kernel of $\eta_{\mathcal{P}}$. We obtain

$$\begin{aligned}(\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(\tau_Q)&=[\mathcal{V}(w,\alpha):\mathcal{V}(w,\alpha)\cap H^1]\hat{f}_{\tau_P}(\tau_P)\circ(\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(w\tau_\alpha w^{-1})\\ \text{(Remark 3.22)}&=d(w,\alpha)q^{\ell(w)}\hat{f}_{\tau_P}(\tau_P)\circ(\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(w\tau_\alpha w^{-1})\\ \text{(Lemma 3.25)}&=d(w,\alpha)q^{\ell(w)}\delta(\mathfrak{J}^1_0,\mathfrak{H}^1_0)^{\ell(w)}\gamma_P\circ\boldsymbol{p}\circ\kappa(w)\circ\iota\circ\gamma_i\circ\boldsymbol{p}\circ\kappa(w^{-1})\circ\iota.\end{aligned}$$

The result follows. □

Lemma 3.40. *We have $\gamma_Q=d(w,\alpha)\delta(\mathfrak{J}^1_0,\mathfrak{H}^1_0)^{\ell(w)}\gamma_P\circ\boldsymbol{p}\circ\kappa(w)\circ\iota\circ\gamma_i\circ\boldsymbol{p}\circ\kappa(w^{-1})\circ\iota$.*

Proof. By the definition of $P=P(w,\alpha)$ and $Q=Q(w,\alpha)$ (see Section 3D) we have

$$\tau_P^{-1}\tau_Q=w\tau_iw^{-1}=\prod_{h=i+1}^{m'}\tau^{-1}_{w(h)}\tau_{w(h)-1}$$

and so

$$\begin{aligned}
\gamma_P^{-1}\gamma_Q &= \prod_{h=i+1}^{m'} \gamma_{w(h)}^{-1}\gamma_{w(h)-1} \\
\text{(Lemma 3.31)} &= \prod_{h=i+1}^{m'} \partial^{m'-w(h)}\varepsilon(\tau_{w(h)}^{-1}\tau_{w(h)-1}) \\
&= \biggl(\prod_{h=i+1}^{m'} \partial^{m'-w(h)}\biggr)\varepsilon(w\tau_i w^{-1}) \\
\text{(Lemma 3.31)} &= \biggl(\prod_{h=i+1}^{m'} \partial^{m'-w(h)}\biggr)\biggl(\prod_{h=i+1}^{m'} \partial^{h-m'}\biggr)\varepsilon(w)\circ\gamma_i\circ\varepsilon(w^{-1}) \\
\text{(Remark 3.28)} &= \biggl(\prod_{h=i+1}^{m'} \partial^{m'-w(h)}\biggr)\biggl(\prod_{h=i+1}^{m'} \partial^{h-m'}\biggr)\delta(\mathfrak{J}_0^1,\mathfrak{H}_0^1)^{\ell(w)}\boldsymbol{p}\circ\kappa(w)\circ\iota\circ\gamma_i\circ\boldsymbol{p}\circ\kappa(w^{-1})\circ\iota \\
&= \biggl(\prod_{h=i+1}^{m'} \partial^{h-w(h)}\biggr)\delta(\mathfrak{J}_0^1,\mathfrak{H}_0^1)^{\ell(w)}\boldsymbol{p}\circ\kappa(w)\circ\iota\circ\gamma_i\circ\boldsymbol{p}\circ\kappa(w^{-1})\circ\iota.
\end{aligned}$$

It remains to prove that $d(w,\alpha)=\prod_{h=i+1}^{m'}\partial^{h-w(h)}$. Since by Remark 3.22 we have $d(w,\alpha)=\partial^{\ell(w)}$, it is sufficient to prove $\sum_{h=i+1}^{m'} h-w(h)=\ell(w)$. We prove this statement by induction on $\ell(w)$. If $\ell(w)=1$, since w is of minimal length in $wW_{\hat{\alpha}}$, we have $w=s_\alpha=(i,i+1)$ and

$$\sum_{h=i+1}^{m'} h-w(h) = i+1-w(i+1)+\sum_{h=i+2}^{m'} h-w(h) = i+1-i+0=1.$$

Let now w be of length $\ell(w)=n>1$. By Lemma 2.12 of [Chinello 2017] there exists $\alpha_{j,j+1}\in P$ and $w'\in W$ of length $n-1$ such that $w=s_jw'$. Then w' is of minimal length in $w'W_{\hat{\alpha}}$ and so we can use the inductive hypothesis. Moreover, by definition of P, there exist $\hat{h}\in\{i+1,\dots,m'\}$ such that $j=w(\hat{h})$ and $j+1\neq w(h)$ for every $h\in\{i+1,\dots,m'\}$ and then $w(h)=w'(h)$ for every $h\in\{i+1,\dots,m'\}$ different from $\hat{h}$. We obtain $\sum_{h=i+1}^{m'} h-w(h)=\sum_{h\neq\hat{h}}(h-w(h))+\hat{h}-w'(\hat{h})+w'(\hat{h})-w(\hat{h})$ which is equal to

$$\sum_{h\neq\hat{h}}(h-w'(h))+\hat{h}-w'(\hat{h})+(s_j(j))-j=\sum_{h=i+1}^{m'} h-w'(h)+j+1-j=\ell(w')+1=\ell(w). \qquad \square$$

Lemma 3.41. *We have* $\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}}=q^{\ell(w)}\hat{f}_{\tau_Q}\sum_u\hat{f}_u$ *where u runs over a system of representatives of* $(U\cap wU^-w^{-1})K^1/K^1$ *in* $U\cap wU^-w^{-1}$.

Proof. By Lemma 3.38 the support of $\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}}$ is contained in $J_{\mathcal{P}}^1\tau_Q(U\cap wU^-w^{-1})J_{\mathcal{P}}^1$. For every $u'\in U\cap wU^-w^{-1}$, by Lemmas 3.39 and 3.40, $(\hat{f}_{\tau_P}\hat{f}_w\hat{f}_{\tau_\alpha}\hat{f}_{w^{-1}})(\tau_Qu')$ is equal to

$$q^{\ell(w)}d(w,\alpha)\delta(\mathfrak{J}_0^1,\mathfrak{H}_0^1)^{\ell(w)}\gamma_P\circ\boldsymbol{p}\circ\kappa(w)\circ\iota\circ\gamma_i\circ\boldsymbol{p}\circ\kappa(w^{-1})\circ\iota\circ\boldsymbol{p}\circ\kappa(u')\circ\iota=q^{\ell(w)}\gamma_Q\circ\boldsymbol{p}\circ\kappa(u')\circ\iota.$$

To conclude we observe that $\big(\hat{f}_{\tau_Q}\sum_u \hat{f}_u\big)(\tau_Q u')=(\hat{f}_{\tau_Q}\hat{f}_{u'})(\tau_Q u')=\gamma_Q\circ \boldsymbol{p}\circ\kappa(u')\circ\iota$ □

Proposition 3.42. *The map Θ'' of Lemma 3.32 respect the relations of Lemma 3.14.*

Proof. By Lemma 3.11 the map Θ'' respects relation 1. By Lemma 3.34 it respects relation 3 and $\hat{f}_{\tau_0^{-1}}\hat{f}_k=\hat{f}_{\tau_0^{-1}k\tau_0}\hat{f}_{\tau_0^{-1}}$ for every $k\in K_B$ and by Lemmas 3.33 and 3.30 it respects relations 2 and 6. Moreover, it respects relations 4 and 5 by Lemma 3.35 and relation 7 by Lemma 3.41. □

Theorem 3.43. *For every nonzero $\gamma\in I_{\tau_{m'-1}}(\eta)$ and every β-extension κ of η there exists an algebra isomorphism $\Theta_{\gamma,\kappa}:\mathscr{H}_R(B^\times,K_B^1)\to\mathscr{H}_R(G,\eta)$.*

Proof. By Proposition 3.42 and by Lemma 3.8 there exists an algebra homomorphism from $\mathscr{H}_R(B^\times,K_B^1)$ to $\mathscr{H}_R(G,\eta)$ which depends on the choice of a β-extension of η and of an element in $I_{\tau_{m'-1}}(\eta_{\mathcal{P}})$, which is isomorphic to $I_{\tau_{m'-1}}(\eta)$ by Lemma 3.8. Let Ξ be a set of representatives of K_B^1-double cosets of $B^\times$. Then $\{f_x\mid x\in\Xi\}$ is a basis of $\mathscr{H}_R(B^\times,K_B^1)$ as an R-vector space and, since $I_G(\eta)=J^1B^\times J^1$ and $\dim_R(I_y(\eta))=1$ for every $y\in I_G(\eta)$, the set $\{\Theta_{\gamma,\kappa}(f_x)\mid x\in\Xi\}$ is a set of generators of $\mathscr{H}_R(G,\eta)$ as an R-vector space and so $\Theta_{\gamma,\kappa}$ is surjective. Moreover, the set $\{\Theta_{\gamma,\kappa}(f_x)\mid x\in\Xi\}$ is linearly independent and so $\Theta_{\gamma,\kappa}$ is also injective. □

Remark 3.44. Let κ and κ' be two β-extensions of η. By Section 2A there exists a character χ of $\mathcal{O}_E^\times$ trivial on $1+\wp_E$ such that $\kappa'=\kappa\otimes(\chi\circ N_{B/E})$. If we consider χ trivial on ϖ_E and we write $\tilde{\chi}=\chi\circ N_{B/E}$, which is a character of $B^\times$, then $\Theta_{\gamma,\kappa}^{-1}\circ\Theta_{\gamma,\kappa'}$ maps f_x to $\tilde{\chi}f_x=\tilde{\chi}(x)f_x$ for every $x\in B^\times$.

4. Semisimple types

Using the notation of Section 2, in this section we present the construction of semisimple types of G with coefficients in R. We refer to Sections 2.8–9 of [Mínguez and Sécherre 2014b] for more details.

Let $r\in\mathbb{N}^*$ and let $(m_1,\dots,m_r)$ be a family of strictly positive integers such that $\sum_{i=1}^r m_i=m$. For every $i\in\{1,\dots,r\}$ we fix a maximal simple type (J_i,λ_i) of $\mathrm{GL}_{m_i}(D)$ and a simple stratum $[\Lambda_i,n_i,0,\beta_i]$ of $A_i=M_{m_i}(D)$ such that $J_i=J(\beta_i,\Lambda_i)$. Then, the centralizer B_i of $E_i=F[\beta_i]$ in A_i is isomorphic to $M_{m_i'}(D_i')$ for a suitable E_i-division algebra D_i' of reduced degree d_i' and a suitable $m_i'\in\mathbb{N}^*$. Moreover, $U(\Lambda_i)\cap B_i^\times$ is a maximal compact open subgroup of $B_i^\times$ which we identify with $\mathrm{GL}_{m_i'}(\mathcal{O}_{D_i'})$.

Let M be the standard Levi subgroup of G of block diagonal matrices of sizes $m_1,\dots,m_r$. The pair (J_M,λ_M) with $J_M=\prod_{i=1}^r J_i$ and $\lambda_M=\bigotimes_{i=1}^r\lambda_i$ is called a *maximal simple type* of M.

For every $i\in\{1,\dots,r\}$ we fix a simple character $\theta_i\in\mathscr{C}_R(\Lambda_i,0,\beta_i)$ contained in λ_i and we observe that this choice does not depend on the choices of the β-extensions implicit in λ_i. Grouping θ_i according their endoclasses, we obtain a partition $\{1,\dots,r\}=\bigsqcup_{j=1}^l I_j$ with $l\in\mathbb{N}^*$. Up to renumbering the (J_i,λ_i) we can suppose that there exist integers $0=a_0<a_1<\cdots<a_l=r$ such that we have $I_j=\{i\in\mathbb{N}\mid a_{j-1}<i\le a_j\}$. For every $j\in\{1,\dots,l\}$ we denote $m^j=\sum_{i\in I_j}m_i$ and $m'^j=\sum_{i\in I_j}m_i'$ and we consider the standard Levi subgroup L of G containing M of block diagonal matrices of sizes $m^1,\dots,m^l$.

Let $j\in\{1,\dots,l\}$. We choose a simple stratum $[\Lambda^j,n^j,0,\beta^j]$ of $M_{m^j}(D)$ as in Section 2.8 of [Mínguez and Sécherre 2014b] (see also Section 6.2 of [Sécherre and Stevens 2016]); in particular we can assume that for every $i \in I_j$ there exist an embedding $\iota_i : F[\beta^j] \to A_i$ such that $\beta_i = \iota_i(\beta^j)$ and that the characters θ_i with $i \in I_j$ are related by the transfer maps. If we denote by B^j the centralizer of $E^j = F[\beta^j]$ in $M_{m^j}(D)$, there exist an E^j-division algebra D'^j and an isomorphism that identifies B^j to $M_{m'^j}(D'^j)$ and $U(\Lambda^j) \cap B^{j\times}$ to the standard parabolic subgroup of $\mathrm{GL}_{m'^j}(\mathcal{O}_{D'^j})$ associated to m'_i with $i \in I_j$. We denote by θ^j the transfer of θ_i with $i \in I_j$ to $\mathscr{C}_R(\Lambda^j, 0, \beta^j)$, which does not depend on i, and we fix a β-extension κ^j of θ^j. In Section 2.8 of [Mínguez and Sécherre 2014b] the authors define two compact open subgroups $\boldsymbol{J}_j \subset J(\beta^j, \Lambda^j)$ and $\boldsymbol{J}_j^1 \subset J^1(\beta^j, \Lambda^j)$ of G such that $\boldsymbol{J}_j/\boldsymbol{J}_j^1 \cong \prod_{i\in I_j} J_i/J_i^1$, and representations $\boldsymbol{\kappa}_j$ of $\boldsymbol{J}_j$ and $\boldsymbol{\eta}_j$ of $\boldsymbol{J}_j^1$ such that

$$\mathrm{ind}_{\boldsymbol{J}_j^1}^{J^1(\beta^j,\Lambda^j)} \boldsymbol{\eta}_j \cong \mathrm{res}^{J(\beta^j,\Lambda^j)}_{J^1(\beta^j,\Lambda^j)} \kappa^j, \quad \mathrm{ind}_{\boldsymbol{J}_j}^{J(\beta^j,\Lambda^j)} \boldsymbol{\kappa}_j \cong \kappa^j, \quad \boldsymbol{J}_j \cap M = \prod_{i\in I_j} J_i, \quad \mathrm{res}^{\boldsymbol{J}_j}_{\boldsymbol{J}_j\cap M} \boldsymbol{\kappa}_j = \bigotimes_{i\in I_j} \kappa_i,$$

where $\kappa_i \in \mathcal{B}(\theta_i)$ for every $i \in I_j$. We denote by η_i the restriction of κ_i to $J^1(\beta_i, \Lambda_i)$ for every $i \in I_j$. We obtain a decomposition $\lambda_i = \kappa_i \otimes \sigma_i$ for every $i \in I_j$ where σ_i is a representation of J_i trivial on J_i^1. We denote by $\boldsymbol{\sigma}_j$ the representation $\bigotimes_{i\in I_j} \sigma_i$ viewed as a representation of $\boldsymbol{J}_j$ trivial on $\boldsymbol{J}_j^1$ and we set $\boldsymbol{\lambda}_j = \boldsymbol{\kappa}_j \otimes \boldsymbol{\sigma}_j$. Then $(\boldsymbol{J}_j, \boldsymbol{\lambda}_j)$ is a cover of $\bigl(\prod_{i\in I_j} J_i, \bigotimes_{i\in I_j} \lambda_i\bigr)$ by Proposition 2.26 of [Mínguez and Sécherre 2014b], $(\boldsymbol{J}_j, \boldsymbol{\kappa}_j)$ is decomposed above $\bigl(\prod_{i\in I_j} J_i, \bigotimes_{i\in I_j} \kappa_i\bigr)$ and $(\boldsymbol{J}_j^1, \boldsymbol{\eta}_j)$ is a cover of $\bigl(\prod_{i\in I_j} J_i^1, \bigotimes_{i\in I_j} \eta_i\bigr)$ by Proposition 2.27 of the same reference.

We set

$$J_M^1 = \prod_{i=1}^r J_i^1, \quad \kappa_M = \bigotimes_{i=1}^r \kappa_i, \quad \eta_M = \bigotimes_{i=1}^r \eta_i, \quad \boldsymbol{J}_L = \prod_{j=1}^l \boldsymbol{J}_j, \quad \boldsymbol{J}_L^1 = \prod_{j=1}^l \boldsymbol{J}_j^1,$$
$$\boldsymbol{\lambda}_L = \bigotimes_{j=1}^l \boldsymbol{\lambda}_j, \quad \boldsymbol{\kappa}_L = \bigotimes_{j=1}^l \boldsymbol{\kappa}_j, \quad \boldsymbol{\eta}_L = \bigotimes_{j=1}^l \boldsymbol{\eta}_j, \quad \boldsymbol{\sigma}_L = \bigotimes_{j=1}^l \boldsymbol{\sigma}_j.$$

By construction $(\boldsymbol{J}_L, \boldsymbol{\lambda}_L)$ and $(\boldsymbol{J}_L^1, \boldsymbol{\eta}_L)$ are covers of (J_M, λ_M) and (J_M^1, η_M) respectively and $(\boldsymbol{J}_L, \boldsymbol{\kappa}_L)$ is decomposed above (J_M, κ_M).

Proposition 2.28 of [loc. cit.] defines a cover $(\boldsymbol{J}, \boldsymbol{\lambda})$ of $(\boldsymbol{J}_L, \boldsymbol{\lambda}_L)$ and so of (J_M, λ_M), that we call a *semisimple type* of G. If the (J_i, λ_i) are maximal simple supertypes, we call $(\boldsymbol{J}, \boldsymbol{\lambda})$ a *semisimple supertype* of G. The semisimple type $(\boldsymbol{J}, \boldsymbol{\lambda})$ is associated to a stratum $[\boldsymbol{\Lambda}, \boldsymbol{n}, 0, \boldsymbol{\beta}]$ of A, which is not necessarily simple (Section 2.9 of [loc. cit.]). We denote by B the centralizer of $\boldsymbol{\beta}$ in A, $B_L^\times = B^\times \cap L = \prod_{j=1}^l B^{j\times}$ and $\boldsymbol{J}^1 = \boldsymbol{J} \cap U_1(\boldsymbol{\Lambda})$. By Propositions 2.30 and 2.31 of [loc. cit.] there exists a unique pair $(\boldsymbol{J}^1, \boldsymbol{\eta})$ decomposed above $(\boldsymbol{J}_L^1, \boldsymbol{\eta}_L)$ and so above (J_M^1, η_M). Its intertwining set is $I_G(\boldsymbol{\eta}) = \boldsymbol{J} B_L^\times \boldsymbol{J}$ and for every $y \in B_L^\times$ the R-vector space $I_y(\boldsymbol{\eta})$ is 1-dimensional. We also have the isomorphisms

$$\boldsymbol{J}/\boldsymbol{J}^1 \cong \boldsymbol{J}_L/\boldsymbol{J}_L^1 \cong \prod_{i=1}^r J_i/J_i^1 \cong \prod_{i=1}^r \mathrm{GL}_{m'_i}(\mathfrak{k}_{D'_i}).$$

We can identify $\boldsymbol{\sigma}_L$ with an irreducible representation $\boldsymbol{\sigma}$ of $\boldsymbol{J}$ trivial on $\boldsymbol{J}^1$. By Proposition 2.33 of [loc. cit.] there exists a unique pair $(\boldsymbol{J}, \boldsymbol{\kappa})$ decomposed above $(\boldsymbol{J}_L, \boldsymbol{\kappa}_L)$ and so above (J_M, κ_M). Moreover, we have

$\boldsymbol{\eta} = \mathrm{res}^{\boldsymbol{J}}_{\boldsymbol{J}^1} \boldsymbol{\kappa}$, $\boldsymbol{\lambda} = \boldsymbol{\kappa} \otimes \boldsymbol{\sigma}$ and $I_G(\boldsymbol{\kappa}) = \boldsymbol{J} B_L^\times \boldsymbol{J}$. We denote by $\mathscr{M}$ the finite group $\prod_{i=1}^r \mathrm{GL}_{m'_i}(\mathfrak{k}_{D'_i})$. Then we can identify $\boldsymbol{\sigma}$ to a cuspidal (supercuspidal if $(\boldsymbol{J}, \boldsymbol{\lambda})$ is a semisimple supertype) representation of $\mathscr{M}$.

Remark 4.1. The choice of β-extensions $\kappa^j \in \mathcal{B}(\theta^j)$ for every $j \in \{1, \dots, l\}$ determines $\kappa_i \in \mathcal{B}(\theta_i)$ for every $i \in \{1, \dots, r\}$, $\boldsymbol{\kappa}^j$ for every $j \in \{1, \dots, l\}$, $\boldsymbol{\kappa}_L$ and $\boldsymbol{\kappa}$ and so the decompositions $\lambda_i = \kappa_i \otimes \sigma_i$, $\boldsymbol{\lambda}_j = \boldsymbol{\kappa}_j \otimes \boldsymbol{\sigma}_j$ and $\boldsymbol{\lambda} = \boldsymbol{\kappa} \otimes \boldsymbol{\sigma}$.

4A. *The representation $\boldsymbol{\eta}_{\max}$*. In this section we associate to every semisimple supertype $(\boldsymbol{J}, \boldsymbol{\lambda})$ of G an irreducible projective representation $\boldsymbol{\eta}_{\max}$ of a compact open subgroup of G and we prove that the algebra $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ is isomorphic to $\mathscr{H}_R(B_L^\times, K_L^1)$ where K_L^1 is the pro-p-radical of the maximal compact open subgroup of $B_L^\times$.

For every $j \in \{1, \dots, l\}$ we choose a simple stratum $[\Lambda_{\max,j}, n_{\max,j}, 0, \beta^j]$ of $M_{m^j}(D)$ such that $U(\Lambda_{\max,j}) \cap B^{j\times}$ is a maximal compact open subgroup of $B^{j\times}$ containing $U(\Lambda^j) \cap B^{j\times}$ as in Section 6.2 of [Sécherre and Stevens 2016]. Then we can identify $U(\Lambda_{\max,j}) \cap B^{j\times}$ to $\mathrm{GL}_{m'^j}(\mathcal{O}_{D'^j})$. Let $J_{\max,j} = J(\beta^j, \Lambda_{\max,j})$ and $J^1_{\max,j} = J^1(\beta^j, \Lambda_{\max,j})$. We can also choose $\theta_{\max,j} \in \mathscr{C}_R(\Lambda_{\max,j}, 0, \beta^j)$ such that its transfer to $\mathscr{C}_R(\Lambda^j, 0, \beta^j)$ is θ^j. We fix a β-extension $\kappa_{\max,j}$ of $\theta_{\max,j}$ and we denote by $\eta_{\max,j}$ its restriction to $J^1_{\max,j}$. By (5.2) of [Sécherre and Stevens 2016], there exists a unique $\kappa^j \in \mathcal{B}(\theta^j)$ such that

$$\mathrm{ind}_{J(\beta^j,\Lambda^j)}^{(U(\Lambda_j)\cap B^{j\times})U_1(\Lambda^j)} \kappa^j \cong \mathrm{ind}_{(U(\Lambda_j)\cap B^{j\times})J^1_{\max,j}}^{(U(\Lambda^j)\cap B^{j\times})U_1(\Lambda^j)} \kappa_{\max,j} \tag{6}$$

and so by Remark 4.1 the choice of $\kappa_{\max,j}$ determines $\boldsymbol{\kappa}_j$. We set

$$J_{\max} = \prod_{j=1}^{l} J_{\max,j}, \quad J^1_{\max} = \prod_{j=1}^{l} J^1_{\max,j}, \qquad \kappa_{\max} = \bigotimes_{j=1}^{l} \kappa_{\max,j},$$
$$\eta_{\max} = \bigotimes_{j=1}^{l} \eta_{\max,j}, \quad K_L = \prod_{j=1}^{l} U(\Lambda_{\max,j}) \cap B^{j\times}, \;\; K_L^1 = \prod_{j=1}^{l} U_1(\Lambda_{\max,j}) \cap B^{j\times}.$$

If we denote by $\mathscr{G}$ the finite group $\prod_{j=1}^l \mathrm{GL}_{m'^j}(\mathfrak{k}_{D'^j})$, we obtain $J_{\max}/J^1_{\max} \cong K_L/K_L^1 \cong \mathscr{G}$ and $(\mathscr{M}, \boldsymbol{\sigma})$ is a supercuspidal pair of $\mathscr{G}$.

As before in this section, by Propositions 2.30, 2.31 and 2.33 of [Mínguez and Sécherre 2014b] we can define two compact open subgroups $\boldsymbol{J}_{\max}$ and $\boldsymbol{J}^1_{\max}$ of G such that $\boldsymbol{J}_{\max}/\boldsymbol{J}^1_{\max} \cong J_{\max}/J^1_{\max} \cong \mathscr{G}$ and pairs $(\boldsymbol{J}_{\max}, \boldsymbol{\kappa}_{\max})$ and $(\boldsymbol{J}^1_{\max}, \boldsymbol{\eta}_{\max})$ decomposed above $(J_{\max}, \kappa_{\max})$ and $(J^1_{\max}, \eta_{\max})$ respectively. Then we have $I_G(\boldsymbol{\kappa}_{\max}) = I_G(\boldsymbol{\eta}_{\max}) = \boldsymbol{J}_{\max} B_L^\times \boldsymbol{J}_{\max}$ and the R-vector spaces $I_y(\boldsymbol{\eta}_{\max})$ and $I_y(\boldsymbol{\kappa}_{\max})$ have dimension 1 for every $y \in B_L^\times$.

Remark 4.2. Since for every $j \in \{1, \dots, l\}$ the choice of $\kappa_{\max,j} \in \mathcal{B}(\theta_{\max,j})$ determines $\boldsymbol{\kappa}_j$, the choice of $\kappa_{\max}$ determines $\boldsymbol{\kappa}$ and $\boldsymbol{\kappa}_{\max}$ and so the decomposition $\boldsymbol{\lambda} = \boldsymbol{\kappa} \otimes \boldsymbol{\sigma}$. On the other hand $\boldsymbol{\eta}_{\max}$, the group $\mathscr{G}$ and the conjugacy class of $\mathscr{M}$ are uniquely determined by the semisimple supertype $(\boldsymbol{J}, \boldsymbol{\lambda})$, independently by the choice of $\kappa_{\max}$ or of $\boldsymbol{\kappa}$.

Proposition 4.3. *The algebras $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ and $\bigotimes_{j=1}^l \mathscr{H}_R(\mathrm{GL}_{m^j}(D), \eta_{\max,j})$ are isomorphic.*

Proof. By Lemma 1.3 and by Lemma 2.4 and Proposition 2.5 of [Guiraud 2013] there exists an algebra isomorphism $\bigotimes_{j=1}^{l} \mathscr{H}_R(\mathrm{GL}_{m^j}(D), \eta_{\max,j}) \to \mathscr{H}_R(L, \eta_{\max})$. Now, since $I_G(\boldsymbol{\eta}_{\max}) \subset \boldsymbol{J}_{\max} L \boldsymbol{J}_{\max}$ the subalgebra $\mathscr{H}_R(\boldsymbol{J}_{\max} L \boldsymbol{J}_{\max}, \boldsymbol{\eta}_{\max})$ of $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ of functions with support in $\boldsymbol{J}_{\max} L \boldsymbol{J}_{\max}$ is equal to $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ and so by Sections II.6–8 of [Vignéras 1998] there exists an algebra isomorphism $\mathscr{H}_R(L, \eta_{\max}) \to \mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ which preserves the support. □

Corollary 4.4. *The R-algebras $\mathscr{H}_R(B_L^\times, K_L^1)$ and $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ are isomorphic.*

Proof. By Remark 1.5 of [Chinello 2017] (see also Theorem 6.3 of [Krieg 1990]) the algebra $\mathscr{H}_R(B_L^\times, K_L^1)$ is isomorphic to $\bigotimes_{j=1}^{l} \mathscr{H}_R(B^{j\times}, U_1(\Lambda_{\max,j}) \cap B^{j\times})$ and then by Theorem 3.43 we obtain, for every $j \in \{1, \dots, l\}$,

$$\mathscr{H}_R(B^{j\times}, U_1(\Lambda_{\max,j}) \cap B^{j\times}) \cong \mathscr{H}_R(\mathrm{GL}_{m^j}(D), \eta_{\max,j}).$$ □

Remark 4.5. By Theorem 3.43 the isomorphism of Corollary 4.4 depends on the choice of a β-extension $\kappa_{\max,j}$ of $\eta_{\max,j}$ and of an intertwining element of $\eta_{\max,j}$ for every $j \in \{1, \dots, l\}$. Using Proposition 4.3, the tensor product of these intertwining elements becomes an intertwining element of $\boldsymbol{\eta}_{\max}$.

Remark 4.6. The procedure that associates $\boldsymbol{\eta}_{\max}$ to $(\boldsymbol{J}, \boldsymbol{\lambda})$ depends on several noncanonical choices, for example the choice of the isomorphism $B_L^\times \to \prod \mathrm{GL}_{m'^j}(D'^j)$. To obtain a canonical correspondence, we denote by $\boldsymbol{\Theta}_i$ the endoclass of θ_i with $i \in \{1, \dots, r\}$ and we canonically associate to $(\boldsymbol{J}, \boldsymbol{\lambda})$ the formal sum

$$\boldsymbol{\Theta}(\boldsymbol{J}, \boldsymbol{\lambda}) = \boldsymbol{\Theta} = \sum_{i=1}^{r} \frac{m_i d}{[E_i : F]} \boldsymbol{\Theta}_i.$$

Furthermore, the group $\mathscr{G}$ and the $\mathscr{G}$-conjugacy class of $\mathscr{M}$ depend only on $(\boldsymbol{J}, \boldsymbol{\lambda})$ and actually the group $\mathscr{G}$ depends only on $\boldsymbol{\Theta}$ because $m'^j[\mathfrak{k}_{D'^j} : \mathfrak{k}_{E^j}] = m^j d/[E^j : F] = \sum_{i \in I_j} m_i d/[E_i : F]$ which is the coefficient of $\boldsymbol{\Theta}_i$ in $\boldsymbol{\Theta}$. We refer to Section 6.3 of [Sécherre and Stevens 2016] for more details.

5. The category equivalence $\mathscr{R}(G, \boldsymbol{\eta}_{\max}) \simeq \mathscr{R}(B_L^\times, K_L^1)$

Using the notation of Section 4, in this section we prove that there exists an equivalence of categories between $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and $\mathscr{R}(B_L^\times, K_L^1)$. This allows to reduce the description of a positive-level block of $\mathscr{R}_R(G)$ to the description of a level-0 block of $\mathscr{R}_R(B_L^\times)$.

5A. ***The category $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$.*** In this section we associate to a semisimple supertype $(\boldsymbol{J}, \boldsymbol{\lambda})$ of G a subcategory of $\mathscr{R}_R(G)$. We refer to [Sécherre and Stevens 2016] for more details.

From now on we fix an extension $\boldsymbol{\kappa}_{\max}$ of $\boldsymbol{\eta}_{\max}$ to $\boldsymbol{J}_{\max}$, as in Section 4A. This uniquely determines a decomposition $\boldsymbol{\lambda} = \boldsymbol{\kappa} \otimes \boldsymbol{\sigma}$ where $\boldsymbol{\kappa}$ is an irreducible representation of $\boldsymbol{J}$ and $\boldsymbol{\sigma}$ is a supercuspidal representation of $\mathscr{M}$ viewed as an irreducible representation of $\boldsymbol{J}$ trivial on $\boldsymbol{J}^1$. We consider the functor $\mathsf{K}_{\boldsymbol{\kappa}_{\max}} : \mathscr{R}_R(G) \to \mathscr{R}(\boldsymbol{J}_{\max}/\boldsymbol{J}_{\max}^1) = \mathscr{R}_R(\mathscr{G})$ given by $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(\pi) = \mathrm{Hom}_{\boldsymbol{J}_{\max}^1}(\boldsymbol{\eta}_{\max}, \pi)$ for every representation π of G, with $\boldsymbol{J}_{\max}$ acting on $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(\pi)$ by

$$x.\varphi = \pi(x) \circ \varphi \circ \boldsymbol{\kappa}_{\max}(x)^{-1} \tag{7}$$

for every $x \in \boldsymbol{J}_{\max}$. We denote by $\pi(\boldsymbol{\kappa}_{\max})$ this representation of $\mathscr{G}$. We remark that if V_1 and V_2 are representations of G and $\phi \in \operatorname{Hom}_G(V_1, V_2)$ then $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(\phi)$ maps φ to $\phi \circ \varphi$ for every $\varphi \in \operatorname{Hom}_G(\rho, V_1)$. For more details on this functor see Section 5 of [Mínguez and Sécherre 2014b] and [Sécherre and Stevens 2016].

We recall that we have $\boldsymbol{\sigma} = \bigotimes_{i=1}^{r} \sigma_i$ where σ_i is a supercuspidal representation of $\mathrm{GL}_{m'_i}(\mathfrak{k}_{D'_i})$. We put $\Gamma_{\mathscr{M}} = \prod_{j=1}^{l} \operatorname{Gal}(\mathfrak{k}_{D'^j}/\mathfrak{k}_{E^j})^{|I_j|}$. The equivalence class of $(\mathscr{M}, \boldsymbol{\sigma})$ (see Definition 1.14 of [Sécherre and Stevens 2016]) is the set, denoted by $[\mathscr{M}, \boldsymbol{\sigma}]$, of supercuspidal pairs $(\mathscr{M}', \boldsymbol{\sigma}')$ of $\mathscr{G}$ such that there exists $\epsilon \in \Gamma_{\mathscr{M}}$ such that $(\mathscr{M}', \boldsymbol{\sigma}')$ is $\mathscr{G}$-conjugate to $(\mathscr{M}, \boldsymbol{\sigma}^{\epsilon})$.

Let $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\boldsymbol{J}, \boldsymbol{\lambda})$. For every representation V of G let $V[\boldsymbol{\Theta}, \boldsymbol{\sigma}]$ be the subrepresentation of V generated by the maximal subspace of $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(V)$ such that every irreducible subquotient has supercuspidal support in $[\mathscr{M}, \boldsymbol{\sigma}]$ and let $V[\boldsymbol{\Theta}]$ be the subrepresentation of V generated by $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(V)$ (see Section 9.1 of [Sécherre and Stevens 2016]).

Definition 5.1. Let $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ be the full subcategory of $\mathscr{R}_R(G)$ of representations V such that $V = V[\boldsymbol{\Theta}, \boldsymbol{\sigma}]$. This does not depend on the choice of $\boldsymbol{\kappa}_{\max}$ (see Section 10.1 of [loc. cit.]).

Remark 5.2. For every representation V of G we have $V[\boldsymbol{\Theta}, \boldsymbol{\sigma}][\boldsymbol{\Theta}, \boldsymbol{\sigma}] = V[\boldsymbol{\Theta}, \boldsymbol{\sigma}]$ (see Lemma 9.2 of [loc. cit.]) and so $V[\boldsymbol{\Theta}, \boldsymbol{\sigma}]$ is an object of $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$.

We define the *equivalence class of* $(\boldsymbol{J}, \boldsymbol{\lambda})$ to be the set $[\boldsymbol{J}, \boldsymbol{\lambda}]$ of semisimple supertypes $(\tilde{\boldsymbol{J}}, \tilde{\boldsymbol{\lambda}})$ of G such that $\operatorname{ind}_{\tilde{\boldsymbol{J}}}^{G}(\tilde{\boldsymbol{\lambda}}) \cong \operatorname{ind}_{\boldsymbol{J}}^{G}(\boldsymbol{\lambda})$.

Theorem 5.3. *The category $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ depends only on the class $[\boldsymbol{J}, \boldsymbol{\lambda}]$ and it is a block of $\mathscr{R}_R(G)$.*

Proof. This follows from Propositions 10.2 and 10.5 and Theorem 10.4 of [Sécherre and Stevens 2016]. □

Remark 5.4. The proof in [loc. cit.] of Theorem 5.3 uses the notions of inertial class of a supercuspidal pair of G and of supercuspidal support (see 1.1.3, 2.1.2 and 2.1.3 of [Mínguez and Sécherre 2014a]). These notions are very important in the study of representations of $\mathrm{GL}_m(D)$ but in this article they are not used explicitly.

5B. ***The category equivalence.*** Let $(\boldsymbol{J}, \boldsymbol{\lambda})$ be a semisimple supertype of G and let $\boldsymbol{\Theta} = \boldsymbol{\Theta}(\boldsymbol{J}, \boldsymbol{\lambda})$ be the formal sum of endoclasses associated to it. In general there exist several semisimple supertypes of G associated to $\boldsymbol{\Theta}$. We put $\boldsymbol{X} = \boldsymbol{X}_{\boldsymbol{\Theta}} = \{[\boldsymbol{J}', \boldsymbol{\lambda}'] \mid \boldsymbol{\Theta}(\boldsymbol{J}', \boldsymbol{\lambda}') = \boldsymbol{\Theta}\}$. In this section we prove that the sum $\bigoplus_{[\boldsymbol{J}', \boldsymbol{\lambda}'] \in \boldsymbol{X}} \mathscr{R}(\boldsymbol{J}', \boldsymbol{\lambda}')$ is equivalent to the level-0 subcategory of $\mathscr{R}_R(B_L^{\times})$.

Let $\boldsymbol{Y} = \boldsymbol{Y}_{\boldsymbol{\Theta}}$ be the set of equivalence classes of supercuspidal pairs of $\mathscr{G}$, that is uniquely determined by $\boldsymbol{\Theta}$ by Remark 4.6. Let $\boldsymbol{\kappa}_{\max}$ be a fixed extension of $\boldsymbol{\eta}_{\max}$ to $\boldsymbol{J}_{\max}$ as in Section 4A and let $\mathsf{K} = \mathsf{K}_{\boldsymbol{\kappa}_{\max}}$. By Proposition 10.7 of [Sécherre and Stevens 2016] there exists a bijection

$$\phi_{\boldsymbol{\kappa}_{\max}} : \boldsymbol{X} \to \boldsymbol{Y} \tag{8}$$

given by $\phi_{\boldsymbol{\kappa}_{\max}}([\boldsymbol{J}', \boldsymbol{\lambda}']) = [\mathscr{M}, \boldsymbol{\sigma}]$ if the supercuspidal supports of irreducible subquotients of $\mathsf{K}(V)$ are in $[\mathscr{M}, \boldsymbol{\sigma}]$ for every (or equivalently for one) object V of $\mathscr{R}(\boldsymbol{J}', \boldsymbol{\lambda}')$. This is equivalent to saying that there exists $\boldsymbol{\kappa}$ as in Section 4 (which depends on $\boldsymbol{\kappa}_{\max}$) such that $\boldsymbol{\lambda}' = \boldsymbol{\kappa} \otimes \boldsymbol{\sigma}'$ with $(\mathscr{M}, \boldsymbol{\sigma}') \in [\mathscr{M}, \boldsymbol{\sigma}]$.

Proposition 5.5 [Sécherre and Stevens 2016, Corollary 9.4]. *For every representation V of G we have*

$$V[\boldsymbol{\Theta}] = \bigoplus_{[\mathscr{M}', \boldsymbol{\sigma}'] \in \boldsymbol{Y}} V[\boldsymbol{\Theta}, \boldsymbol{\sigma}']. \tag{9}$$

Proposition 5.6 [loc. cit., Lemma 10.3]. *If $[\boldsymbol{J}', \boldsymbol{\lambda}'] \in \boldsymbol{X}$ and W is a simple object of $\mathscr{R}(\boldsymbol{J}', \boldsymbol{\lambda}')$ then $\mathsf{K}(W) \neq 0$.*

Since $\boldsymbol{J}^1_{\max}$ has invertible pro-order in R, the representation $\boldsymbol{\eta}_{\max}$ is projective and so we can use the notation and results of Section 1B. We have defined the functor

$$\boldsymbol{M}_{\boldsymbol{\eta}_{\max}} : \mathscr{R}_R(G) \to \text{Mod-}\,\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$$

by $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V) = \mathrm{Hom}_G(\mathrm{ind}^G_{\boldsymbol{J}^1_{\max}}(\boldsymbol{\eta}_{\max}), V)$ and $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\phi) : \varphi \mapsto \varphi \circ \phi$ for all representations V and V_1 of G, $\phi \in \mathrm{Hom}_G(V, V_1)$ and $\varphi \in \mathrm{Hom}_G(\mathrm{ind}^G_{\boldsymbol{J}^1_{\max}}(\boldsymbol{\eta}_{\max}), V)$.

Remark 5.7. Frobenius reciprocity induces a natural isomorphism between the functor $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}$ composed with the forgetful functor $\text{Mod-}\,\mathscr{H}_R(G, \boldsymbol{\eta}_{\max}) \to \text{Mod-}\,R$ and the functor $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}$ composed with the forgetful functor $\mathscr{R}_R(\mathscr{G}) \to \text{Mod-}\,R$. This implies that for every representation V of G the subrepresentation $V[\boldsymbol{\Theta}]$ of V is the subrepresentation $V[\boldsymbol{\eta}_{\max}]$ defined in Section 1B.

We have also defined the full subcategories $\mathscr{R}_{\boldsymbol{\eta}_{\max}}(G)$ and $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ of $\mathscr{R}_R(G)$. We recall that $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ is the category of V such that $V = V[\boldsymbol{\Theta}]$ and $\mathscr{R}_{\boldsymbol{\eta}_{\max}}(G)$ is the category of V such that $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V') \neq 0$ for every irreducible subquotient V' of V.

Lemma 5.8. *We have $\mathscr{R}(G, \boldsymbol{\eta}_{\max}) = \mathscr{R}_{\boldsymbol{\eta}_{\max}}(G)$.*

Proof. Thanks to Remark 1.8 it is sufficient to prove $\mathscr{R}(G, \boldsymbol{\eta}_{\max}) \subset \mathscr{R}_{\boldsymbol{\eta}_{\max}}(G)$. Let V be a representation in $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$. By Proposition 5.5 we have $V = \bigoplus_{\boldsymbol{Y}} V[\boldsymbol{\Theta}, \boldsymbol{\sigma}']$ and by Remark 5.2 the representation $V[\boldsymbol{\Theta}, \boldsymbol{\sigma}']$ is an object of $\mathscr{R}(\boldsymbol{J}', \boldsymbol{\lambda}')$ where $[\boldsymbol{J}', \boldsymbol{\lambda}'] = \phi^{-1}_{\boldsymbol{\kappa}_{\max}}([\mathscr{M}, \boldsymbol{\sigma}']) \in \boldsymbol{X}$. Hence, we obtain the inclusion $\mathscr{R}(G, \boldsymbol{\eta}_{\max}) \subset \bigoplus_{\boldsymbol{X}} \mathscr{R}(\boldsymbol{J}', \boldsymbol{\lambda}')$. Let now W be an object of $\bigoplus_{\boldsymbol{X}} \mathscr{R}(\boldsymbol{J}', \boldsymbol{\lambda}')$ and W' an irreducible subquotient of W. Then W' is an irreducible object of $\mathscr{R}(\boldsymbol{J}', \boldsymbol{\lambda}')$ for a $[\boldsymbol{J}', \boldsymbol{\lambda}'] \in \boldsymbol{X}$ and so by Proposition 5.6 we have $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(W) \neq 0$. Therefore, by Remark 5.7 we have $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(W') \neq 0$ which implies $\bigoplus_{\boldsymbol{X}} \mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda}') \subset \mathscr{R}_{\boldsymbol{\eta}_{\max}}(G)$. $\square$

Remark 5.9. We have proved that $\mathscr{R}(G, \boldsymbol{\eta}_{\max}) = \mathscr{R}_{\boldsymbol{\eta}_{\max}}(G) = \bigoplus_{[\boldsymbol{J}, \boldsymbol{\lambda}] \in \boldsymbol{X}} \mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$. Moreover, by Proposition 1.7, a representation V of G is in this category if and only if it satisfies one of the following equivalent conditions:

- $V = V[\boldsymbol{\Theta}]$.
- For every subquotient Z of V we have $Z = Z[\boldsymbol{\Theta}]$.
- For every irreducible subquotient U of V we have $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(U) \neq 0$.
- For every nonzero subquotient W of V we have $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(W) \neq 0$.

Theorem 5.10. *The functor $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}$ is an equivalence of categories between*

$$\mathscr{R}(G, \boldsymbol{\eta}_{\max}) \quad \textit{and} \quad \text{Mod-}\mathscr{H}_R(G, \boldsymbol{\eta}_{\max}).$$

Proof. We apply Theorem 1.9 with $\mathtt{G} = G$ and $\sigma = \boldsymbol{\eta}_{\max}$. □

Remark 5.11. We recall that a level-0 representation of $B_L^\times$ is a representation generated by its K_L^1-invariant vectors. It is equivalent to say that all irreducible subquotients have nonzero K_L^1-invariant vectors (see Section 3 of [Chinello 2017]). The category $\mathscr{R}(B_L^\times, K_L^1)$ is called the *level-*0 *subcategory* of $\mathscr{R}_R(B_L^\times)$. By Section 3 of [Chinello 2017] and Theorem 1.9, the K_L^1-invariant functor $\operatorname{inv}_{K_L^1}$ induces an equivalence of categories between $\mathscr{R}(B_L^\times, K_L^1)$ and Mod-$\mathscr{H}_R(B_L^\times, K_L^1)$ whose quasiinverse is

$$W \mapsto W \otimes_{\mathscr{H}_R(B_L^\times, K_L^1)} \operatorname{ind}_{K_L^1}^{B_L^\times}(1).$$

We recall that if (ϱ, Z) is a representation of $B_L^\times$ then the action of $\Phi \in \mathscr{H}_R(B_L^\times, K_L^1)$ on $z \in Z^{K_L^1}$ is given by $z.\Phi = \sum_{x \in K_L^1 \backslash B_L^\times} \Phi(x)\varrho(x^{-1})z$.

Corollary 5.12. *There exists an equivalence of categories between $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and $\mathscr{R}(B_L^\times, K_L^1)$.*

Proof. By Corollary 4.4 the algebras $\mathscr{H}_R(B_L^\times, K_L^1)$ and $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ are isomorphic. We obtain an equivalence of categories between Mod-$\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ and Mod-$\mathscr{H}_R(B_L^\times, K_L^1)$ and so between $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and $\mathscr{R}(B_L^\times, K_L^1)$ by Theorem 5.10 and Remark 5.11. □

Now we want to describe the functor that induces this equivalence of categories. We recall that we have fixed an isomorphism $B_L^\times \cong \prod \mathrm{GL}_{m'^j}(D'^j)$ and an extension $\boldsymbol{\kappa}_{\max}$ of $\boldsymbol{\eta}_{\max}$. We also fix a nonzero intertwining element γ of $\boldsymbol{\eta}_{\max}$ as in Remark 4.5. By Corollary 4.4 we have an isomorphism $\Theta_{\gamma,\boldsymbol{\kappa}_{\max}}: \mathscr{H}_R(B_L^\times, K_L^1) \to \mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ which induces an equivalence of categories $\Theta^*_{\gamma,\boldsymbol{\kappa}_{\max}}:$ Mod-$\mathscr{H}_R(G, \boldsymbol{\eta}_{\max}) \to$ Mod-$\mathscr{H}_R(B_L^\times, K_L^1)$. We obtain the diagram

$$\begin{tikzcd}
\mathscr{R}(G, \boldsymbol{\eta}_{\max}) \arrow[r, "\text{Corollary 5.12}"] \arrow[d, "\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}"] & \mathscr{R}(B_L^\times, K_L^1) \\
\text{Mod-}\mathscr{H}_R(G, \boldsymbol{\eta}_{\max}) \arrow[r, "\Theta^*_{\gamma,\boldsymbol{\kappa}_{\max}}"] & \text{Mod-}\mathscr{H}_R(B_L^\times, K_L^1). \arrow[u, "\text{Remark 5.11}"']
\end{tikzcd} \tag{10}$$

The functor $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}: \mathscr{R}(G, \boldsymbol{\eta}_{\max}) \to$ Mod-$\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ is an equivalence of categories by Theorem 5.10. By Lemma 1.3 the right action of $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ on $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V)$ is given by $(m.\Psi)(f) = m(\Psi * f)$ for every $m \in \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V)$, $\Psi \in \mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ and $f \in \operatorname{ind}_{\boldsymbol{J}^1_{\max}}^G(\boldsymbol{\eta}_{\max})$. The right-action of $\Phi \in \mathscr{H}_R(B_L^\times, K_L^1)$ on a $\mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$-module N is given by $N.\Phi = N.\Theta_{\gamma,\boldsymbol{\kappa}_{\max}}(\Phi)$. By Remark 5.11 the functor $W \mapsto W \otimes_{\mathscr{H}_R(B_L^\times, K_L^1)} \operatorname{ind}_{K_L^1}^{B_L^\times}(1)$ is an equivalence of categories between Mod-$\mathscr{H}_R(B_L^\times, K_L^1)$ and $\mathscr{R}(B_L^\times, K_L^1)$ where, by Lemma 1.3, the left-action of $\Phi \in \mathscr{H}_R(B_L^\times, K_L^1)$ on $f \in \operatorname{ind}_{K_L^1}^{B_L^\times}(1)$ is given by $\Phi.f = \Phi * f$. Moreover, the left-action of $x \in B_L^\times$ on $w \otimes f \in W \otimes_{\mathscr{H}_R(B_L^\times, K_L^1)} \operatorname{ind}_{K_L^1}^{B_L^\times}(1)$ is given by $x.(w \otimes f) = w \otimes (x.f)$.

Composing these three functors we obtain the equivalence of categories of Corollary 5.12 which we denote by $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ and is given by

$$\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(\pi, V) = \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\pi, V) \otimes_{\mathscr{H}_R(B_L^\times, K_L^1)} \mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1}) \tag{11}$$

for every (π, V) in $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$, where the right-action of $\Phi \in \mathscr{H}_R(B_L^\times, K_L^1)$ on $m \in \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\pi, V)$ is given by $(m.\Phi)(f) = m(\Theta_{\gamma,\boldsymbol{\kappa}_{\max}}(\Phi) * f)$ for every $f \in \mathrm{ind}_{\boldsymbol{J}_{\max}^1}^G(\boldsymbol{\eta}_{\max})$. We remark that if V_1 and V_2 are in $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and $\phi \in \mathrm{Hom}_G(V_1, V_2)$ then $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(\phi)$ maps $m \otimes f$ to $(\phi \circ m) \otimes f$ for every $m \in \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V_1)$ and $f \in \mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1})$.

5C. ***Correspondence between blocks.*** In this section we discuss the correspondence among blocks of $\mathscr{R}(B_L^\times, K_L^1)$ and those of $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ induced by the equivalence of categories $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ defined in (11).

We consider the functor $\mathsf{K}_{K_L} : \mathscr{R}(B_L^\times, K_L^1) \to \mathscr{R}_R(K_L/K_L^1) = \mathscr{R}_R(\mathscr{G})$ given by $\mathsf{K}_{K_L}(Z) = Z^{K_L^1}$ and $\mathsf{K}_{K_L}(\phi) = \phi_{|Z^{K_L^1}}$ for all representations (ϱ, Z) and (ϱ_1, Z_1) of $B_L^\times$ and every $\phi \in \mathrm{Hom}_{B_L^\times}(Z, Z_1)$, where $x \in K_L$ acts on $z \in Z^{K_L^1}$ by $x.z = \varrho(x)z$. It is the functor presented in Section 5A when we replace G by $B_L^\times$ and $\boldsymbol{\kappa}_{\max}$ by the trivial representation of K_L. We also consider the functor $\boldsymbol{H} : \text{Mod-}\mathscr{H}_R(B_L^\times, K_L^1) \to \mathscr{R}_R(K_L/K_L^1)$ given by $\boldsymbol{H}(W) = (\varrho', W)$ and $\boldsymbol{H}(\phi) = \phi$ for all $\mathscr{H}_R(B_L^\times, K_L^1)$-modules W and W_1 and every $\phi \in \mathrm{Hom}_{\mathscr{H}_R(B_L^\times, K_L^1)}(W, W_1)$, where $\varrho'(k)w = w.f_{k^{-1}}$ for every $k \in K_L$ and $w \in W$.

Remark 5.13. The functor K_{K_L} is the composition of $\mathrm{inv}_{K_L^1}$ (see Remark 5.11) and the functor $\boldsymbol{H}$. Actually if (ϱ, Z) is an object of $\mathscr{R}(B_L^\times, K_L^1)$ then $\boldsymbol{H}(\mathrm{inv}_{K_L^1}(Z)) = \boldsymbol{H}(Z^{K_L^1}) = (\varrho', Z^{K_L^1})$ where $\varrho'(k)z = z.f_{k^{-1}} = \sum_{x \in K_L^1 \backslash B_L^\times} f_{k^{-1}}(x)\varrho(x^{-1})z = \varrho(k)z$ for every $z \in Z^{K_L^1}$ and $k \in K_L$.

We obtain the diagram

$$\begin{tikzcd}
\mathscr{R}(G, \boldsymbol{\eta}_{\max}) \arrow[rr, "\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}"] \arrow[dr, "\Theta^*_{\gamma,\boldsymbol{\kappa}_{\max}} \circ \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}"] \arrow[ddr, bend right, "\mathsf{K}_{\boldsymbol{\kappa}_{\max}}"'] & & \mathscr{R}(B_L^\times, K_L^1) \arrow[dl, "\mathrm{inv}_{K_L^1}"'] \arrow[ddl, bend left, "\mathsf{K}_{K_L}"] \\
& \text{Mod-}\mathscr{H}_R(B_L^\times, K_L^1) \arrow[d, "\boldsymbol{H}"] & \\
& \mathscr{R}_R(\mathscr{G}) &
\end{tikzcd} \tag{12}$$

Proposition 5.14. *There exists a natural isomorphism between $\mathsf{K}_{K_L} \circ \boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ and $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}$.*

Proof. By Remark 5.13 we have $\mathsf{K}_{K_L} \circ \boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}} = \boldsymbol{H} \circ \mathrm{inv}_{K_L^1} \circ \boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ and by (10) we have a natural isomorphism between $\mathrm{inv}_{K_L^1} \circ \boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ and $\Theta^*_{\gamma,\boldsymbol{\kappa}_{\max}} \circ \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}$ so it is sufficient to find a natural isomorphism $\mathfrak{z} : \boldsymbol{H} \circ \Theta^*_{\gamma,\boldsymbol{\kappa}_{\max}} \circ \boldsymbol{M}_{\boldsymbol{\eta}_{\max}} \to \mathsf{K}_{\boldsymbol{\kappa}_{\max}}$. For every object (π, V) of $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$, let $\mathfrak{z}_V : \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V) \to \mathsf{K}_{\boldsymbol{\kappa}_{\max}}(V)$ be the isomorphism of R-modules given by Remark 5.7. The action of $x \in K_L/K_L^1 \cong \mathscr{G}$ on $m \in \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\pi, V)$ is given by $x.m = m.\Theta_{\gamma,\boldsymbol{\kappa}_{\max}}(f_{x^{-1}}) = m.\tilde{f}_{x^{-1}}$ where $\tilde{f}_{x^{-1}} \in \mathscr{H}_R(G, \boldsymbol{\eta}_{\max})$ has support $x^{-1}\boldsymbol{J}_{\max}^1$ and $\tilde{f}_{x^{-1}}(x^{-1}) = \boldsymbol{\kappa}_{\max}(x^{-1})$ while the action of $x \in \boldsymbol{J}_{\max}/\boldsymbol{J}_{\max}^1 \cong \mathscr{G}$ on $\varphi \in \mathsf{K}_{\boldsymbol{\kappa}_{\max}}(V)$ is given by (7). We have to prove that $\mathfrak{z}_V(x.m) = x.\mathfrak{z}_V(m)$ for every $m \in \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\pi, V)$ and $x \in \mathscr{G}$. We recall that in Section 1A we defined elements $i_v : \boldsymbol{J}^1_{\max} \to V_{\boldsymbol{\eta}_{\max}}$ with $v \in V_{\boldsymbol{\eta}_{\max}}$ such that $m(i_v) = \mathfrak{Z}_V(m)(v)$, which generate $\mathrm{ind}^G_{\boldsymbol{J}^1_{\max}}(\boldsymbol{\eta}_{\max})$ as a representation of G. Then for every $v \in V_{\boldsymbol{\eta}_{\max}}$ we have

$$\mathfrak{Z}_V(x.m)(v) = (x.m)(i_v) = (m.\tilde{f}_{x^{-1}})(i_v) = m(\tilde{f}_{x^{-1}} * i_v).$$

The support of $\tilde{f}_{x^{-1}} * i_v$ is $\boldsymbol{J}^1_{\max}x^{-1}$ and $(\tilde{f}_{x^{-1}} * i_v)(x^{-1}) = \tilde{f}_{x^{-1}}(x^{-1})v = \boldsymbol{\kappa}_{\max}(x^{-1})v$. We obtain $\mathfrak{Z}_V(x.m)(v) = m(x.i_{\boldsymbol{\kappa}_{\max}(x^{-1})v}) = \pi(x)(m(i_{\boldsymbol{\kappa}_{\max}(x^{-1})v})) = \pi(x)(\mathfrak{Z}_V(m)(\boldsymbol{\kappa}_{\max}(x^{-1})v)) = (x.\mathfrak{Z}_V(m))(v)$. Now, let V_1 and V_2 be two objects of $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and let $\phi \in \mathrm{Hom}_G(V_1, V_2)$. Then for every $m \in \boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V_1)$ and every $v \in V_{\boldsymbol{\eta}_{\max}}$ we have $\mathfrak{Z}_{V_2}(\boldsymbol{H}(\Theta^*_{\gamma,\boldsymbol{\kappa}_{\max}}(\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(\phi)))(m))(v) = \mathfrak{Z}_{V_2}(\phi \circ m)(v)$ which is equal to $(\phi \circ m)(i_v)$. On the other hand we have $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(\phi)(\mathfrak{Z}_{V_1}(m))(v) = \phi(\mathfrak{Z}_{V_1}(m)(v))$ which is equal to $\phi(m(i_v))$. This shows that $\mathfrak{Z}$ is a natural isomorphism. □

Now we look for a block decomposition of $\mathscr{R}(B_L^\times, K_L^1)$. Let $[\mathscr{M}, \boldsymbol{\sigma}] \in \boldsymbol{Y}$. Then $\mathscr{M} = \prod_{j=1}^l \mathscr{M}_j$ and $\boldsymbol{\sigma} = \bigotimes_{j=1}^l \boldsymbol{\sigma}_j$ where $\mathscr{M}_j \cong \boldsymbol{J}_j/\boldsymbol{J}_j^1$ and $[\mathscr{M}_j, \boldsymbol{\sigma}_j]$ is a class of supercuspidal pairs of $\mathrm{GL}_{m'^j}(\mathfrak{k}_{D'^j})$. For every $j \in \{1, \dots, l\}$, replacing G by $B^{j\times}$ and $\boldsymbol{\kappa}_{\max}$ by the trivial character of $U(\Lambda_{\max,j}) \cap B^{j\times}$ in Definition 5.1, we obtain an abelian full subcategory $\mathscr{R}(U(\Lambda_{\max,j}) \cap B^{j\times}, \boldsymbol{\sigma}_j)$ of $\mathscr{R}_R(B^{j\times})$ whose objects are representations V_j of $B^{j\times}$ generated by the maximal subspace of ${V_j}^{U_1(\Lambda_{\max,j})\cap B^{j\times}}$ for which every irreducible subquotient has supercuspidal support in $[\mathscr{M}_j, \boldsymbol{\sigma}_j]$. We obtain a full subcategory $\mathscr{R}(K_L, \boldsymbol{\sigma})$ of $\mathscr{R}_R(B_L^\times)$ (and of $\mathscr{R}(B_L^\times, K_L^1)$) whose objects are representations V of $B_L^\times$ generated by the maximal subspace of $V^{K_L^1}$ such that every irreducible subquotient has supercuspidal support in $[\mathscr{M}, \boldsymbol{\sigma}]$. Theorem 5.3 and Remark 5.9 give a block decomposition of $\mathscr{R}(B^{j\times}, U_1(\Lambda_{\max,j}) \cap B^{j\times})$ for every $j \in \{1, \dots, l\}$ and so we obtain a block decomposition

$$\mathscr{R}(B_L^\times, K_L^1) = \bigoplus_{[\mathscr{M},\boldsymbol{\sigma}]\in \boldsymbol{Y}} \mathscr{R}(K_L, \boldsymbol{\sigma}).$$

We recall that we have a block decomposition $\mathscr{R}(G, \boldsymbol{\eta}_{\max}) = \bigoplus_{[\boldsymbol{J},\boldsymbol{\lambda}]\in \boldsymbol{X}} \mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ by Remark 5.9 and a bijection $\phi_{\boldsymbol{\kappa}_{\max}} : \boldsymbol{X} \to \boldsymbol{Y}$ defined in (8) which depends on the choice of $\boldsymbol{\kappa}_{\max}$.

Theorem 5.15. *Let $[\boldsymbol{J}, \boldsymbol{\lambda}] \in \boldsymbol{X}$ and $[\mathscr{M}, \boldsymbol{\sigma}] = \phi_{\boldsymbol{\kappa}_{\max}}([\boldsymbol{J}, \boldsymbol{\lambda}]) \in \boldsymbol{Y}$. Then $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ induces an equivalence of categories between the block $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ of $\mathscr{R}_R(G)$ and the block $\mathscr{R}(K_L, \boldsymbol{\sigma})$ of $\mathscr{R}_R(B_L^\times)$.*

Proof. If V is an object of $\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$, by Proposition 5.14 there exists an isomorphism of representations of $\mathscr{G}$ between $\mathsf{K}_{K_L}(\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(V))$ and $\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(V)$. Then irreducible subquotients of $(\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(V))^{K_L^1}$ have supercuspidal support in $[\mathscr{M}, \boldsymbol{\sigma}]$ and so $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(V)$ is in $\mathscr{R}(K_L, \boldsymbol{\sigma})$. □

We remark that the matching of the blocks of $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and of $\mathscr{R}(B_L^\times, K_L^1)$ does not depend on the choice of the intertwining element γ of $\boldsymbol{\eta}_{\max}$ while the equivalence of categories between these blocks, induced by $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(V)$, depends on this choice.

5D. ***Dependence on the choice of $\boldsymbol{\kappa}_{\max}$.*** In this section we discuss the dependence of results of Sections 5A, 5B and 5C on the choice of the extension of $\boldsymbol{\eta}_{\max}$ to $\boldsymbol{J}_{\max}$.

Let $(\boldsymbol{J}, \boldsymbol{\lambda})$ be a semisimple supertype of G. We have just seen in Remark 4.6 that the group $\mathscr{G}$ depends only on $\boldsymbol{\Theta}(\boldsymbol{J}, \boldsymbol{\lambda})$ and by Remark 4.6 and Theorem 5.3 the $\mathscr{G}$-conjugacy class of $\mathscr{M}$ and the category

$\mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ do not depend on the choice of the extension of $\boldsymbol{\eta}_{\max}$ to $\boldsymbol{J}_{\max}$. Moreover, the sum (9) does not depend on this choice because a different one permutes the terms $V[\boldsymbol{\Theta}, \boldsymbol{\sigma}']$ in $V[\boldsymbol{\Theta}]$. Then $V[\boldsymbol{\Theta}]$, the equalities $\mathscr{R}(G, \boldsymbol{\eta}_{\max}) = \mathscr{R}_{\boldsymbol{\eta}_{\max}}(G) = \bigoplus_{[\boldsymbol{J}, \boldsymbol{\lambda}] \in X} \mathscr{R}(\boldsymbol{J}, \boldsymbol{\lambda})$ and the equivalence of Theorem 5.10 do not depend on the choice of the extension of $\boldsymbol{\eta}_{\max}$.

Let γ be a fixed nonzero intertwining element of $\boldsymbol{\eta}_{\max}$ as in Remark 4.5. Using notation of Section 4A let $\boldsymbol{\kappa}_{\max}$ and $\boldsymbol{\kappa}'_{\max}$ be two extensions of $\boldsymbol{\eta}_{\max}$ to $\boldsymbol{J}_{\max}$ and let $\kappa_{\max} = \bigotimes_{j=1}^{l} \kappa_{\max,j}$ and $\kappa'_{\max} = \bigotimes_{j=1}^{l} \kappa'_{\max,j}$ be the restrictions to $J_{\max}$ of $\boldsymbol{\kappa}_{\max}$ and $\boldsymbol{\kappa}'_{\max}$ respectively. Then, for every $j \in \{1, \dots, l\}$, $\kappa_{\max,j}$ and $\kappa'_{\max,j}$ are β-extensions of $\theta_{\max,j}$ and so by Section 2A there exists a character χ_j of $\mathcal{O}_{E^j}^{\times}$ trivial on $1 + \wp_{E^j}$ such that $\kappa'_{\max,j} = \kappa_{\max,j} \otimes (\chi_j \circ N_{B^j/E^j})$. Let χ and $\overline{\chi}$ be the character $\bigotimes_{j=1}^{l} (\chi_j \circ N_{B^j/E^j})$ viewed as characters of $\boldsymbol{J}_{\max}$ trivial on $\boldsymbol{J}^1_{\max}$ and of $\mathscr{G}$ respectively and, if we consider χ_j trivial on ϖ_{E^j} for every $j \in \{1, \dots, l\}$, let $\tilde{\chi} = \bigotimes_{j=1}^{l} (\chi_j \circ N_{B^j/E^j})$ viewed as a character of $B_L^{\times}$.

We consider the functors $\widetilde{\mathfrak{X}} : \mathscr{R}(B_L^{\times}, K_L^1) \to \mathscr{R}(B_L^{\times}, K_L^1)$ and $\overline{\mathfrak{X}} : \mathscr{R}_R(\mathscr{G}) \to \mathscr{R}_R(\mathscr{G})$ given by $\widetilde{\mathfrak{X}}(\varrho) = \varrho \otimes \tilde{\chi}^{-1}$, $\widetilde{\mathfrak{X}}(\tilde{\phi}) = \tilde{\phi}$, $\overline{\mathfrak{X}}(\tau) = \tau \otimes \overline{\chi}^{-1}$ and $\overline{\mathfrak{X}}(\bar{\phi}) = \bar{\phi}$ for every ϱ, ϱ_1 in $\mathscr{R}(B_L^{\times}, K_L^1)$, every $\tilde{\phi} \in \mathrm{Hom}_{B_L^{\times}}(\varrho, \varrho_1)$, all representations τ and τ_1 of $\mathscr{G}$ and every $\bar{\phi} \in \mathrm{Hom}_{\mathscr{G}}(\tau, \tau_1)$. We consider the following diagram.

$$
\begin{tikzcd}
\mathscr{R}(B_L^{\times}, K_L^1) \arrow[rr, "\mathsf{K}_{K_L}"] \arrow[dd, "\widetilde{\mathfrak{X}}"'] & & \mathscr{R}_R(\mathscr{G}) \arrow[dd, "\overline{\mathfrak{X}}"] \\
& \mathscr{R}(G, \boldsymbol{\eta}_{\max}) \arrow[ul, "F_{\gamma,\boldsymbol{\kappa}_{\max}}"'] \arrow[ur, "\mathsf{K}_{\boldsymbol{\kappa}_{\max}}"] \arrow[dl, "F_{\gamma,\boldsymbol{\kappa}'_{\max}}"] \arrow[dr, "\mathsf{K}_{\boldsymbol{\kappa}'_{\max}}"'] & \\
\mathscr{R}(B_L^{\times}, K_L^1) \arrow[rr, "\mathsf{K}_{K_L}"] & & \mathscr{R}_R(\mathscr{G}).
\end{tikzcd}
\tag{13}
$$

Lemma 5.16. *We have $\mathsf{K}_{\boldsymbol{\kappa}'_{\max}} = \overline{\mathfrak{X}} \circ \mathsf{K}_{\boldsymbol{\kappa}_{\max}}$ and so for every representation (π, V) in $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ we have $\pi(\boldsymbol{\kappa}'_{\max}) = \pi(\boldsymbol{\kappa}_{\max}) \otimes \overline{\chi}^{-1}$.*

Proof. The space of $\mathsf{K}_{\boldsymbol{\kappa}'_{\max}}(V)$ and $\overline{\mathfrak{X}}(\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(V))$ is $\mathrm{Hom}_{\boldsymbol{J}^1_{\max}}(\boldsymbol{\eta}_{\max}, V)$. Let φ be in this space and $x \in \boldsymbol{J}_{\max}$. Let Q be the standard parabolic subgroup of G with Levi component L, let N be the unipotent radical of Q such that $Q = LN$ and let N^- be the unipotent radical opposite to N. We choose $x_1 \in \boldsymbol{J}_{\max} \cap N^-$, $x_2 \in J_{\max}$ and $x_3 \in \boldsymbol{J}_{\max} \cap N$ such that $x = x_1 x_2 x_3$. Since $(\boldsymbol{\kappa}_{\max}, \boldsymbol{J}_{\max})$ and $(\boldsymbol{\kappa}'_{\max}, \boldsymbol{J}_{\max})$ are decomposed above $(\kappa_{\max}, J_{\max})$ and $(\kappa'_{\max}, J_{\max})$ respectively, we obtain $\pi(\boldsymbol{\kappa}'_{\max})(x)(\varphi) = \pi(x) \circ \varphi \circ \boldsymbol{\kappa}'_{\max}(x^{-1})$ which is equal to $\pi(x) \circ \varphi \circ \kappa'_{\max}(x_2^{-1}) = \pi(x) \circ \varphi \circ \kappa_{\max}(x_2^{-1}) \chi(x_2^{-1}) = \pi(\boldsymbol{\kappa}_{\max})(x)(\varphi) \chi(x_2)^{-1}$. Since $\boldsymbol{J}_{\max} \cap N = \boldsymbol{J}^1_{\max} \cap N$ and $\boldsymbol{J}_{\max} \cap N^- = \boldsymbol{J}^1_{\max} \cap N^-$ we obtain $\chi(x_2)^{-1} = \chi(x)^{-1}$. Now, let V_1 and V_2 be two objects of $\mathscr{R}(G, \boldsymbol{\eta}_{\max})$ and let $\phi \in \mathrm{Hom}_G(V_1, V_2)$. Then for every $\varphi \in \mathrm{Hom}_{\boldsymbol{J}^1_{\max}}(\boldsymbol{\eta}_{\max}, V_1)$ we have $\mathsf{K}_{\boldsymbol{\kappa}'_{\max}}(\phi)(\varphi) = \phi \circ \varphi = \overline{\mathfrak{X}}(\mathsf{K}_{\boldsymbol{\kappa}_{\max}}(\phi))(\varphi)$. $\square$

Lemma 5.17. *We have $\mathsf{K}_{K_L} \circ \widetilde{\mathfrak{X}} = \overline{\mathfrak{X}} \circ \mathsf{K}_{K_L}$.*

Proof. Let (ϱ, Z) be in $\mathscr{R}(B_L^{\times}, K_L^1)$. The space of $\mathsf{K}_{K_L}(\widetilde{\mathfrak{X}}(Z))$ and $\overline{\mathfrak{X}}(\mathsf{K}_{K_L}(Z))$ is $Z^{K_L^1}$. Let $x \in K_L$ and let $\bar{x}$ be the projection of x in $K_L/K_L^1 \cong \mathscr{G}$. For every $z \in Z^{K_L^1}$ we have $\mathsf{K}_{K_L}(\widetilde{\mathfrak{X}}(\varrho))(\bar{x})(z) = \tilde{\chi}(x^{-1}) \varrho(x) v$

while $\overline{\mathfrak{X}}(\mathsf{K}_{K_L}(\varrho))(\bar{x})(z)=\bar{\chi}(\bar{x}^{-1})\varrho(x)v$. Now, let Z_1 and Z_2 be two objects of $\mathscr{R}(B_L^\times, K_L^1)$ and let $\phi\in\mathrm{Hom}_{B_L^\times}(Z_1,Z_2)$. Then we have $\mathsf{K}_{K_L}(\widetilde{\mathfrak{X}}(\phi))=\phi_{|Z_1^{K_L^1}}=\overline{\mathfrak{X}}(\mathsf{K}_{K_L}(\phi))$. □

We remark that by Proposition 5.14 and Lemmas 5.16 and 5.17, the functor $\mathsf{K}_{K_L}\circ\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}'_{\max}}$ is naturally isomorphic to $\mathsf{K}_{\boldsymbol{\kappa}'_{\max}}$ which is equal to $\overline{\mathfrak{X}}\circ\mathsf{K}_{\boldsymbol{\kappa}_{\max}}$ which is naturally isomorphic to $\overline{\mathfrak{X}}\circ\mathsf{K}_{K_L}\circ\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$ which is equal to $\mathsf{K}_{K_L}\circ\widetilde{\mathfrak{X}}\circ\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$.

Proposition 5.18. *There exists a natural isomorphism between $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}'_{\max}}$ and $\widetilde{\mathfrak{X}}\circ\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}$.*

Proof. For every object (π,V) in $\mathscr{R}(G,\boldsymbol{\eta}_{\max})$, the space of $\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}'_{\max}}(V)$ and $\widetilde{\mathfrak{X}}(\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(V))$ is

$$\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V)\otimes_{\mathscr{H}_R(B_L^\times,K_L^1)}\mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1}).$$

If $m\in\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V)$ and $f\in\mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1})$, in the first case the right-action of $\Phi\in\mathscr{H}_R(B_L^\times,K_L^1)$ on m and the left-action of $x\in B_L^\times$ on $m\otimes f$ are given by $m\star'\Phi=m.\Theta_{\gamma,\boldsymbol{\kappa}'_{\max}}(\Phi)$ and $x\diamond'(m\otimes f)=m\otimes x.f$ while in the second case they are given by $m\star\Phi=m.\Theta_{\gamma,\boldsymbol{\kappa}_{\max}}(\Phi)$ and $x\diamond(m\otimes f)=\tilde{\chi}(x^{-1})m\otimes x.f$. Let $\mathfrak{Z}_V$ be the automorphism of $\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V)\otimes_{\mathscr{H}_R(B_L^\times,K_L^1)}\mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1})$ that maps $m\otimes f$ to $m\otimes\tilde{\chi}f$ for every $m\in\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V)$ and $f\in\mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1})$. By Remark 3.44 we have $m\star'\Phi=m\star\tilde{\chi}\Phi$ and then

$$\begin{aligned}\mathfrak{Z}_V(m\star'\Phi\otimes f)&=(m\star'\Phi)\otimes(\tilde{\chi}f)\\&=(m\star\tilde{\chi}\Phi)\otimes(\tilde{\chi}f)\\&=m\otimes((\tilde{\chi}\Phi)*(\tilde{\chi}f))\\&=m\otimes\tilde{\chi}(\Phi*f)\\&=\mathfrak{Z}_V(m\otimes(\Phi*f)).\end{aligned}$$

This implies that $\mathfrak{Z}_V$ is well defined as an R-linear automorphism. Moreover, for every $x\in B_L^\times$ we have $\mathfrak{Z}_V(x\diamond'(m\otimes f))=m\otimes\tilde{\chi}(x.f)=\tilde{\chi}(x^{-1})m\otimes x.(\tilde{\chi}f)=x\diamond\mathfrak{Z}_V(m\otimes f)$ and so $\mathfrak{Z}_V$ is an isomorphism of representations of $B_L^\times$. Now, let V_1 and V_2 be two objects of $\mathscr{R}(G,\boldsymbol{\eta}_{\max})$ and let $\phi\in\mathrm{Hom}_G(V_1,V_2)$. Then for every $m\in\boldsymbol{M}_{\boldsymbol{\eta}_{\max}}(V_1)$ and $f\in\mathrm{ind}_{K_L^1}^{B_L^\times}(1_{K_L^1})$ we have $\mathfrak{Z}_{V_2}(\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}'_{\max}}(\phi)(m\otimes f))=\mathfrak{Z}_{V_2}((\phi\circ m)\otimes f)=(\phi\circ m)\otimes\tilde{\chi}f$ which is equal to $\widetilde{\mathfrak{X}}(\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(\phi))(m\otimes\tilde{\chi}f)=\widetilde{\mathfrak{X}}(\boldsymbol{F}_{\gamma,\boldsymbol{\kappa}_{\max}}(\phi))(\mathfrak{Z}_{V_1}(m\otimes f))$. □

By Remark 4.2, the representations $\boldsymbol{\kappa}_{\max}$ and $\boldsymbol{\kappa}'_{\max}$ determine two decompositions $\boldsymbol{\lambda}=\boldsymbol{\kappa}\otimes\boldsymbol{\sigma}$ and $\boldsymbol{\lambda}=\boldsymbol{\kappa}'\otimes\boldsymbol{\sigma}'$ where $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}'$ are supercuspidal representations of $\mathscr{M}$ viewed as irreducible representations of $\boldsymbol{J}_L$ trivial on $\boldsymbol{J}_L^1$. Hence, the bijection $\phi_{\boldsymbol{\kappa}'_{\max}}\circ\phi_{\boldsymbol{\kappa}_{\max}}^{-1}$ permutes the elements of $\boldsymbol{Y}$ and it maps $[\mathscr{M},\boldsymbol{\sigma}]$ to $[\mathscr{M},\boldsymbol{\sigma}']$. Let $\boldsymbol{\kappa}_L$ and $\boldsymbol{\kappa}'_L$ be the restrictions to $\boldsymbol{J}_L$ of $\boldsymbol{\kappa}$ and $\boldsymbol{\kappa}'$ respectively. By (6) and by (2.20) of [Mínguez and Sécherre 2014b] we have $\boldsymbol{\kappa}'_L=\boldsymbol{\kappa}_L\otimes\chi$ and so $\boldsymbol{\sigma}'=\boldsymbol{\sigma}\otimes\bar{\chi}^{-1}$.

References

[Bernstein 1984] J. N. Bernstein, “Le ‘centre’ de Bernstein”, pp. 1–32 in *Représentations des groupes réductifs sur un corps local*, edited by P. Deligne, Hermann, Paris, 1984. MR Zbl

[Blondel 2005] C. Blondel, “Quelques propriétés des paires couvrantes”, *Math. Ann.* **331**:2 (2005), 243–257. MR Zbl

[Broussous et al. 2012] P. Broussous, V. Sécherre, and S. Stevens, "Smooth representations of $\mathrm{GL}_m(D)$, V: Endo-classes", *Doc. Math.* **17** (2012), 23–77. MR Zbl

[Bushnell and Henniart 1996] C. J. Bushnell and G. Henniart, "Local tame lifting for $\mathrm{GL}(N)$, I: Simple characters", *Inst. Hautes Études Sci. Publ. Math.* **83** (1996), 105–233. MR Zbl

[Bushnell and Kutzko 1993] C. J. Bushnell and P. C. Kutzko, *The admissible dual of* $\mathrm{GL}(N)$ *via compact open subgroups*, Annals of Math. Studies **129**, Princeton Univ. Press, 1993. MR Zbl

[Bushnell and Kutzko 1998] C. J. Bushnell and P. C. Kutzko, "Smooth representations of reductive p-adic groups: structure theory via types", *Proc. London Math. Soc.* (3) **77**:3 (1998), 582–634. MR Zbl

[Bushnell and Kutzko 1999] C. J. Bushnell and P. C. Kutzko, "Semisimple types in GL_n", *Compos. Math.* **119**:1 (1999), 53–97. MR Zbl

[Chinello 2015] G. Chinello, *Représentations ℓ-modulaires des groupes p-adiques: décomposition en blocs de la catégorie des représentations lisses de* $\mathrm{GL}(m, D)$*, groupe métaplectique et représentation de Weil*, Ph.D. thesis, Université de Versailles St-Quentin-en-Yvelines, 2015, Available at https://tinyurl.com/phdchiphd.

[Chinello 2017] G. Chinello, "Hecke algebra with respect to the pro-p-radical of a maximal compact open subgroup for $\mathrm{GL}(n, F)$ and its inner forms", *J. Algebra* **478** (2017), 296–317. MR Zbl

[Dat 2012] J.-F. Dat, "Théorie de Lubin–Tate non Abélienne ℓ-entière", *Duke Math. J.* **161**:6 (2012), 951–1010. MR Zbl

[Dat 2018] J.-F. Dat, "Equivalences of tame blocks for p-adic linear groups", *Math. Ann.* **371**:1-2 (2018), 565–613. MR Zbl

[Guiraud 2013] D.-A. Guiraud, "On semisimple ℓ-modular Bernstein-blocks of a p-adic general linear group", *J. Number Theory* **133**:10 (2013), 3524–3548. MR Zbl

[Helm 2016] D. Helm, "The Bernstein center of the category of smooth $W(k)[\mathrm{GL}_n(F)]$-modules", *Forum Math. Sigma* **4** (2016), art. id. e11. MR Zbl

[Krieg 1990] A. Krieg, *Hecke algebras*, Mem. Amer. Math. Soc. **435**, Amer. Math. Soc., Providence, RI, 1990. MR Zbl

[Mínguez and Sécherre 2014a] A. Mínguez and V. Sécherre, "Représentations lisses modulo ℓ de $\mathrm{GL}_m(D)$", *Duke Math. J.* **163**:4 (2014), 795–887. MR Zbl

[Mínguez and Sécherre 2014b] A. Mínguez and V. Sécherre, "Types modulo ℓ pour les formes intérieures de GL_n sur un corps local non archimédien", *Proc. Lond. Math. Soc.* (3) **109**:4 (2014), 823–891. MR Zbl

[Sécherre 2004] V. Sécherre, "Représentations lisses de $\mathrm{GL}(m, D)$, I: Caractères simples", *Bull. Soc. Math. France* **132**:3 (2004), 327–396. MR Zbl

[Sécherre 2005a] V. Sécherre, "Représentations lisses de $\mathrm{GL}(m, D)$, II: β-extensions", *Compos. Math.* **141**:6 (2005), 1531–1550. MR Zbl

[Sécherre 2005b] V. Sécherre, "Représentations lisses de $\mathrm{GL}_m(D)$, III: Types simples", *Ann. Sci. École Norm. Sup.* (4) **38**:6 (2005), 951–977. MR Zbl

[Sécherre and Stevens 2008] V. Sécherre and S. Stevens, "Représentations lisses de $\mathrm{GL}_m(D)$, IV: Représentations supercuspidales", *J. Inst. Math. Jussieu* **7**:3 (2008), 527–574. MR Zbl

[Sécherre and Stevens 2012] V. Sécherre and S. Stevens, "Smooth representations of $\mathrm{GL}_m(D)$, VI: Semisimple types", *Int. Math. Res. Not.* **2012**:13 (2012), 2994–3039. MR Zbl

[Sécherre and Stevens 2016] V. Sécherre and S. Stevens, "Block decomposition of the category of ℓ-modular smooth representations of GL_n(F) and its inner forms", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:3 (2016), 669–709. MR Zbl

[Vignéras 1996] M.-F. Vignéras, *Représentations l-modulaires d'un groupe réductif p-adique avec $l \neq p$*, Progress in Math. **137**, Birkhäuser, Boston, 1996. MR Zbl

[Vignéras 1998] M.-F. Vignéras, "Induced R-representations of p-adic reductive groups", *Selecta Math.* (*N.S.*) **4**:4 (1998), 549–623. MR Zbl

Communicated by Marie-France Vignéras
Received 2017-07-31 Revised 2018-05-08 Accepted 2018-06-12

gianmarco.chinello@unimib.it *Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, Milano, Italy*

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1715

Algebraic dynamics of the lifts of Frobenius

Junyi Xie

We study the algebraic dynamics of endomorphisms of projective spaces with coefficients in a p-adic field whose reduction in positive characteristic is the Frobenius. In particular, we prove a version of the dynamical Manin–Mumford conjecture and the dynamical Mordell–Lang conjecture for the coherent backward orbits of such endomorphisms. We also give a new proof of a dynamical version of the Tate–Voloch conjecture in this case. Our method is based on the theory of perfectoid spaces introduced by P. Scholze. In the appendix, we prove that under some technical condition on the field of definition, a dynamical system for a polarized lift of Frobenius on a projective variety can be embedded into a dynamical system for some endomorphism of a projective space.

1. Introduction

In this paper, we write $\mathbb{C}_p$ for the completion of the algebraically closure of $\mathbb{Q}_p$ with the induced norm. Denote by $\mathbb{C}_p^\circ$ its valuation ring and $\mathbb{C}_p^{\circ\circ}$ the maximal ideal of $\mathbb{C}_p^\circ$. Let $F : \mathbb{P}^N_{\mathbb{C}_p} \to \mathbb{P}^N_{\mathbb{C}_p}$ be an endomorphism taking form

$$F : [x_0 : \cdots : x_N] \mapsto [x_0^q + p' P_0(x_0, \ldots, x_N) : \cdots : x_N^q + p' P_N(x_0, \ldots, x_N)]$$

where q is a power of p, $p' \in \mathbb{C}_p^{\circ\circ}$, and $P_0, \ldots, P_N$ are homogeneous polynomials of degree q in $\mathbb{C}_p^\circ[x_0, \ldots, x_N]$. We say that F is *a lift of Frobenius on* $\mathbb{P}^N_{\mathbb{C}_p}$.

In this paper we present a new argument for studying the algebraic dynamics for such maps, which is based on the theory of perfectoid spaces introduced by Scholze. In particular, we study some dynamical analogues of diophantine geometry for such maps.

The author is supported by the labex CIMI.
MSC2010: primary 37P55; secondary 37P20, 37P35.
Keywords: algebraic dynamics, perfectoid space.

Dynamical Manin–Mumford conjecture. At first, we recall the dynamical Manin–Mumford conjecture proposed by Zhang [1995].

Dynamical Manin–Mumford Conjecture. Let $F : X_{\mathbb{C}} \to X_{\mathbb{C}}$ be an endomorphism of a quasiprojective variety defined over $\mathbb{C}$. Let V be a subvariety of X. If the Zariski closure of the set of preperiodic points[1] of F contained in V is Zariski dense in V, then V itself is preperiodic, and likewise for periodic points.[2]

This conjecture is a dynamical analogue of the Manin–Mumford conjecture on subvarieties of abelian varieties. More precisely, let V be an irreducible subvariety inside an abelian variety A over $\mathbb{C}$ such that the intersection of the set of torsion points of A and V is Zariski dense in V. Then the Manin–Mumford conjecture asserts that there exists an abelian subvariety V_0 of A and a torsion point $a \in A(\mathbb{C})$ such that $V = V_0 + a$.

The Manin–Mumford conjecture was first proved by Raynaud [1983a; 1983b]. Various versions of this conjecture were proved by Ullmo [1998], Zhang [1998], Buium [1996b], Hrushovski [2001] and Pink and Roessler[2002]. Observe that the dynamical Manin–Mumford conjecture for the map $x \mapsto 2x$ on A implies the classical Manin–Mumford conjecture.

The dynamical Manin–Mumford conjecture does not hold in full generality, as we have some counterexamples [Ghioca et al. 2011; Pazuki 2010; Pazuki 2013]. In particular, Pazuki [2013] shows that counterexamples can come from a lift of Frobenius crossed with a lift of its Verschiebung. This motivated the proposal of several modified versions of the conjecture [Ghioca et al. 2011; Yuan and Zhang 2017].

However, this conjecture is now known to hold in some special cases [Baker and Hsia 2005; Fakhruddin 2014; Medvedev and Scanlon 2014; Ghioca and Tucker 2010; Dujardin and Favre 2017; Ghioca et al. 2011; 2015; 2018]. It seems that the dynamical Manin–Mumford conjecture may be true except a few families of counterexamples.

In this paper, we prove the dynamical Manin–Mumford conjecture for periodic points of lifts of Frobenius on $\mathbb{P}^N$.

Theorem 1.1. *Let $F : \mathbb{P}^N_{\mathbb{C}_p} \to \mathbb{P}^N_{\mathbb{C}_p}$ be a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. Denote by* Per *the set of periodic closed points in $\mathbb{P}^N_{\mathbb{C}_p}$. Let V be any irreducible subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$ such that $V \cap$* Per *is Zariski dense in V. Then V is periodic i.e., there exists $\ell \geq 1$ such that $F^{\ell}(V) = V$.*

We note that Medvedev and Scanlon [2014] have proved Theorem 1.1 in the case

$$F : [x_0 : \cdots : x_N] \mapsto [x_0^q + pP(x_0, x_N) : \cdots : x_{N-1}^q + pP(x_{N-1}, x_N) : x_N^q],$$

where q is a power of p and $P \in \mathbb{Z}_p[x, y]$ is a homogenous polynomial of degree q. Pazuki [2013] studied the lifts of Frobenius on abelian varieties.

We should mention that, recently Scanlon gave a new proof of this theorem without using perfectoid spaces. Since this proof is unpublished and it is completely different from ours, we will discuss it briefly in Section 4 of this paper.

[1] A preperiodic point x is a point satisfying $F^m(x) = F^n(x)$ for some $m > n \geq 0$.

[2] A periodic point x is a point satisfying $F^n(x) = x$ for some $n > 0$.

Dynamical Tate–Voloch conjecture. Let V be an irreducible subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$. There are homogenous polynomials $H_i \in \mathbb{C}_p[x_0, \dots, x_N]$, $i = 1, \dots, m$ satisfying $\|H_i\| = 1$ which define V. For any point $y \in \mathbb{P}^N_{\mathbb{C}_p}(\mathbb{C}_p)$, we may write $y = [y_0 : \cdots : y_N]$, $\max\{|y_i|\}_{0\le i\le N} = 1$. Then we denote by $d(y, V) := \max\{|H_i(y_0, \dots, y_N)|\}_{1\le i\le m}$. Observe that $d(y, V)$ does not depend on the choice of $\{H_i\}_{1\le i\le m}$ or the coordinates $[y_0 : \cdots : y_N]$ of y. It can be viewed as the distance between y and V. Moreover for any quasiprojective variety X and subvariety V of X, by choosing an embedding $X \hookrightarrow \mathbb{P}^N_{\mathbb{C}_p}$, $d(\bullet, \overline{V})$ defines a distance between V and a point in X.

Tate and Voloch [1996] made the following conjecture:

Tate–Voloch Conjecture. Let A be a semiabelian variety over $\mathbb{C}_p$ and V a subvariety of A. Then there exists $c > 0$ such that for any torsion point $x \in A$, we have either $x \in V$ or $d(x, V) > c$.

This conjecture was proved by Scanlon [1999] when A is defined over a finite extension of $\mathbb{Q}_p$. Buium [1996a] proved a dynamical version of this conjecture for periodic points of lifts of Frobenius on any algebraic variety. Here we state it only for the lifts of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$.

Theorem 1.2. *Let $F : \mathbb{P}^N_{\mathbb{C}_p} \to \mathbb{P}^N_{\mathbb{C}_p}$ be a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. Let V be any irreducible subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$. Then there exists $\delta > 0$ such that for any point $x \in$* Per, *either $d(x, V) > \delta$ or $x \in V$.*

In this paper, we give a new proof of this theorem by using the theory of perfectoid spaces.

Dynamical Mordell–Lang conjecture. The Mordell–Lang conjecture on subvarieties of semiabelian varieties (now a theorem of Faltings [1994] and Vojta [1996]) says that if V is a subvariety of a semiabelian variety G defined over $\mathbb{C}$ and Γ is a finitely generated subgroup of $G(\mathbb{C})$, then $V(\mathbb{C}) \bigcap \Gamma$ is a union of at most finitely many translates of subgroups of Γ.

Inspired by this, Ghioca and Tucker proposed the following dynamical analogue of the Mordell–Lang conjecture.

Dynamical Mordell–Lang Conjecture [Ghioca and Tucker 2009]. Let X be a quasiprojective variety defined over $\mathbb{C}$, let $f : X \to X$ be an endomorphism, and V be any subvariety of X. For any point $x \in X(\mathbb{C})$ the set $\{n \in \mathbb{N} \mid f^n(x) \in V(\mathbb{C})\}$ is a union of at most finitely many arithmetic progressions.[3]

Observe that the dynamical Mordell–Lang conjecture implies the classical Mordell–Lang conjecture in the case $\Gamma \simeq (\mathbb{Z}, +)$.

The dynamical Mordell–Lang conjecture has been proved in many cases. For example, Bell, Ghioca and Tucker [2010] proved this conjecture for étale maps, and the author proved it for endomorphisms of $\mathbb{A}^2_{\overline{\mathbb{Q}}}$ [Xie 2017]. We refer to the book [Bell et al. 2016] for a good survey of this conjecture.

We note that the dynamical Mordell–Lang conjecture is not a full generalization of the Mordell–Lang conjecture. In particular, it considers only the forward orbit but not the backward orbit. In an informal seminar, Zhang asked me the following question:

[3]An arithmetic progression is a set of the form $\{an + b \mid n \in \mathbb{N}\}$ with $a, b \in \mathbb{N}$. In particular, when $a = 0$, it contains only one point.

Question 1.3. Let X be a quasiprojective variety over $\mathbb{C}$ and $F : X \to X$ be a finite endomorphism. Let x be a point in $X(\mathbb{C})$. Denote by $O^-(x) := \bigcup_{i=0}^{\infty} F^{-i}(x)$ the backward orbit of x. Let V be a positive dimensional irreducible subvariety of X. If $V \cap O^-(x)$ is Zariski dense in V, what can we say about V?

We note that if V is preperiodic, then $V \cap O^-(x)$ is Zariski dense in V. As with the dynamical Manin–Mumford conjecture, the converse is not true. Indeed, we have the following example. Let $X = \mathbb{A}^1_{\mathbb{C}} \times \mathbb{A}^1_{\mathbb{C}}$ and $f : X \to X$ be the endomorphism defined by $(x, y) \mapsto (x^4, y^6)$. Let V be the diagonal and $x = (1, 1)$. Then $V \cap O^-(x)$ is Zariski dense in V, but V is not preperiodic. We have counterexamples even when F is a polarized endomorphism.[4] The following example is given by Ghioca, which is similar to [Ghioca et al. 2011, Theorem 1.2].

Example 1.4. Let E be the elliptic curve over $\mathbb{C}$ defined by the lattice $\mathbb{Z}[i] \subseteq \mathbb{C}$. Let F_1 be the endomorphism on E defined by the multiplication by 10 and F_2 be the endomorphism on E defined by the multiplication by $6 + 8i$. Set $X := E \times E$, $F := (F_1, F_2)$ on X. Since $|10| = |6 + 8i|$, F is a polarized endomorphism on X. Let V be the diagonal in X and x be the origin. We may check that $V \cap O^-(x)$ is Zariski dense in V, but V is not preperiodic.

As a special case of Question 1.3, we propose the following conjecture.

Conjecture 1.5. Let X be a quasiprojective variety over $\mathbb{C}$ and $F : X \to X$ be a finite endomorphism. Let $\{b_i\}_{i\geq 0}$ be a sequence of points in $X(\mathbb{C})$ satisfying $f(b_i) = b_{i-1}$ for all $i \geq 1$. Let V be a positive dimensional irreducible subvariety of X. If the $\{b_i\}_{i\geq 0} \cap V$ is Zariski dense in V, then V is periodic under F.

Remark 1.6. This conjecture can be viewed as the dynamical Mordell–Lang conjecture for the coherent backward orbits. In fact, it is easy to see that Conjecture 1.5 is equivalent to the following:

Conjecture 1.5*. Let X be a quasiprojective variety over $\mathbb{C}$ and $F : X \to X$ be a finite endomorphism. Let $\{b_i\}_{i\geq 0}$ be a sequence of points in $X(\mathbb{C})$ satisfying $f(b_i) = b_{i-1}$ for all $i \geq 1$. Let V be a subvariety of X. Then the set $\{n \geq 0 \mid b_n \in V\}$ is a union of at most finitely many arithmetic progressions.

Conjecture 1.5⇒ Conjecture 1.5.* If $\{b_i\}_{i\geq 0}$ is finite, then the b_i are contained in a periodic circle. Then Conjecture 1.5* trivially holds. Now we assume that $\{b_i\}_{i\geq 0}$ is infinite. Set $W := \bigcap_{n\geq 0} \overline{\{b_i \mid b_i \in V, i \geq n\}}$. Then there exists $N \geq 0$ such that $W = \overline{\{b_i \mid b_i \in V, i \geq N\}}$. We note that $\{n \geq 0 \mid b_n \in V\} \setminus \{n \geq 0 \mid b_n \in W\} \subseteq \{0, \dots, N\}$ is finite. After replacing b_0 by b_N, we may assume that $N = 0$. If W is empty, then $\{n \geq 0 \mid b_n \in V\} = \{n \geq 0 \mid b_n \in W\} = \varnothing$. If W is not empty, then every irreducible component of W has positive dimension and $\{b_i\}_{i\geq N} \cap W$ is Zariski dense in W. Conjecture 1.5 implies that there exists $r \geq 1$ such that $F^r(W) = W$. If for some index $i \in \{0, \dots, r-1\}$, there exists $s \geq 0$ such that $b_{i+sr} \notin V$, then $b_{i+nr} \notin V$ for all $n \geq s$. Denote by T_i, $i = 0, \dots, r-1$, the set of $j \geq 0$ satisfying $b_j \in V$ and

[4] An endomorphism $F : X \to X$ on a projective variety is said to be polarized if there exists an ample line bundle L on X satisfying $F^*L = L^{\otimes d}$, $d \geq 2$.

$j = i \mod r$. Then T_i is either finite or equal to $\{i + rn \mid n \in \mathbb{N}\}$. It follows that

$$\{n \geq 0 \mid b_n \in V\} = \{n \geq 0 \mid b_n \in W\} = \bigcup_{i=0}^{r-1} T_i$$

is a union of at most finitely many arithmetic progressions. $\square$

*Conjecture 1.5** $\Rightarrow$*Conjecture 1.5.* Assume that V is a positive dimensional irreducible subvariety of X such that $\{b_i\}_{i\geq 0} \cap V$ is Zariski dense in V. Then $\{n \geq 0 \mid b_n \in V\}$ is infinite. Conjecture 1.5 shows that $\{n \geq 0 \mid b_n \in V\}$ takes the form $\{n \geq 0 \mid b_n \in V\} = F \cup \big(\bigcup_{j=1}^{s} T_j\big)$ where F is finite and T_j, $j = 1, \ldots, s$, are infinite arithmetic progressions. There exists $j \in \{1, \ldots, s\}$ such that $\{b_i \mid i \in T_j\}$ is Zariski dense in V. Write $T_j = a + r\mathbb{N}$ where $a \geq 0$, $r \geq 1$. Since $F(\{b_i \mid i \in T_j\}) \setminus \{b_i \mid i \in T_j\} = \{a\}$, we have $F^r(V) = V$. $\square$

In this paper, we prove Conjecture 1.5 for the lifts of Frobenius of $\mathbb{P}^N_{\mathbb{C}_p}$.

Theorem 1.7. *Let $F : \mathbb{P}^N_{\mathbb{C}_p} \to \mathbb{P}^N_{\mathbb{C}_p}$ be a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. Let $\{b_i\}_{i\geq 0}$ be a sequence of points in $\mathbb{P}^N_{\mathbb{C}_p}(\mathbb{C}_p)$ satisfying $f(b_i) = b_{i-1}$ for all $i \geq 1$. Let V be a positive dimensional irreducible subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$. If $\{b_i\}_{i\geq 0} \cap V$ is Zariski dense in V, then V is periodic under F.*

In fact, we prove a stronger statement.

Theorem 1.8. *Let $F : \mathbb{P}^N_{\mathbb{C}_p} \to \mathbb{P}^N_{\mathbb{C}_p}$ be a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. Let $\{b_i\}_{i\geq 0}$ be a sequence of points in $\mathbb{P}^N_{\mathbb{C}_p}(\mathbb{C}_p)$ satisfying $f(b_i) = b_{i-1}$ for all $i \geq 1$. Let V be a subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$. If there exists a subsequence $\{b_{n_i}\}_{i\geq 0}$ such that $|d(b_{n_i}, V)| \to 0$ when $n \to \infty$, then $b_{n_i} \in V$ for i large enough and there exists $r \geq 0$, such that $\{b_i\}_{i\geq 0} \subseteq \bigcup_{i=0}^{r} F^i(V)$.*

It implies the following Tate–Voloch type statement.

Corollary 1.9. *Let $F : \mathbb{P}^N_{\mathbb{C}_p} \to \mathbb{P}^N_{\mathbb{C}_p}$ be a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. Let $\{b_i\}_{i\geq 0}$ be a sequence of points in $\mathbb{P}^N_{\mathbb{C}_p}(\mathbb{C}_p)$ satisfying $f(b_i) = b_{i-1}$ for all $i \geq 1$. Let V be a subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$. Then there exists $c > 0$ such that for all $i \geq 0$, either $b_i \in V$ or $d(b_i, V) > c$.*

Overview of the proofs. Let us now see in more detail how our arguments work.

Denote by $K := \mathbb{C}_p$ and $K^\flat := \widehat{\overline{\mathbb{F}_p((t))}}$ the completion of the algebraic closure of $\mathbb{F}_p$. We denote by K° and $K^{\flat\circ}$ the valuation rings of K and $K^\flat$, respectively, and by $K^{\circ\circ}$ and $K^{\flat\circ\circ}$ the maximal ideal of K° and $K^{\flat\circ}$, respectively. Denote by $k := \overline{\mathbb{F}_p}$. We have $k = K^\circ/K^{\circ\circ} = K^{\flat\circ}/K^{\flat\circ\circ}$. Moreover, we have an embedding $k \hookrightarrow K^\flat$.

Let $F : \mathbb{P}^N_K \to \mathbb{P}^N_K$ be an endomorphism taking form

$$F : [x_0 : \cdots : x_N] \mapsto [x_0^q + p' P_0(x_0, \ldots, x_n) : \cdots : x_N^q + p' P_N(x_0, \ldots, x_N)],$$

where $p' \in K^{\circ\circ}$, q is a power of p, and $P_0, \ldots, P_N$ are homogeneous polynomials of degree q in $K^\circ[x_0, \ldots, x_N]$.

We associate to $\mathbb{P}^N_K$ and $\mathbb{P}^N_{K^\flat}$ nonarchimedean analytic spaces $\mathbb{P}^{N,\mathrm{ad}}_K$ and $\mathbb{P}^{N,\mathrm{ad}}_{K^\flat}$, respectively, with natural embeddings $\mathbb{P}^N_K(K) \subseteq \mathbb{P}^{N,\mathrm{ad}}_K$ and $\mathbb{P}^N_{K^\flat}(K^\flat) \subseteq \mathbb{P}^{N,\mathrm{ad}}_{K^\flat}$. The endomorphism F extends to an endomorphism F^{ad} on $\mathbb{P}^{N,\mathrm{ad}}_K$.

Denote by $\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}}$ the inverse limit of the $\mathbb{P}_K^{N,\mathrm{ad}}$ where the transition maps are F^{ad}. Then we may construct a perfectoid space $\mathbb{P}_K^{N,\mathrm{perf}}$ with an endomorphism F^{perf} for which the topological dynamical system $(\mathbb{P}_K^{N,\mathrm{perf}}, F^{\mathrm{perf}})$ is isomorphic to $(\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}}, T)$ where $T : (x_0, x_1, \ldots) \to (F^{\mathrm{ad}}(x_0), x_0, \ldots)$ is the shift map on $\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}}$. Moreover, we have a natural morphism $\pi : \mathbb{P}_K^{N,\mathrm{perf}} \to \mathbb{P}_K^{N,\mathrm{ad}}$ defined by projection to the first coordinate. This construction has been stated by Scholze [2014, §7].

Similarly, we construct a perfectoid space $\mathbb{P}_{K^\flat}^{N,\mathrm{perf}}$ which is isomorphic to the inverse limit $\varprojlim_{\Phi^s} \mathbb{P}_{K^\flat}^{N,\mathrm{ad}}$ where Φ is the Frobenius endomorphism on $\mathbb{P}_{K^\flat}^{N,\mathrm{ad}}$, and $q = p^s$. Denote by $\pi^\flat : \mathbb{P}_{K^\flat}^{N,\mathrm{perf}} \to \mathbb{P}_{K^\flat}^{N,\mathrm{ad}}$ the morphism defined by the projection to the first coordinate. Since Φ is a homeomorphism on the underlying topological space, $\pi^\flat$ induces an isomorphism from the topological dynamical system $(\mathbb{P}_{K^\flat}^{N,\mathrm{perf}}, \Phi^{s,\mathrm{perf}})$ to $(\mathbb{P}_{K^\flat}^{N,\mathrm{ad}}, \Phi^{s,\mathrm{ad}})$, where Φ^{perf} is the Frobenius on $\mathbb{P}_{K^\flat}^{N,\mathrm{perf}}$.

By the theory of perfectoid spaces, there is a natural homeomorphism of topological space

$$\rho : \mathbb{P}_K^{N,\mathrm{perf}} \to \mathbb{P}_{K^\flat}^{N,\mathrm{perf}}$$

satisfying $\Phi^{s,\mathrm{perf}} \circ \rho = \rho \circ F^{\mathrm{perf}}$.

As an example, we explain the proof of Theorem 1.1. Let V be any subvariety of $\mathbb{P}_{\mathbb{C}_p}^N$ such that $V \cap \mathrm{Per}$ is Zariski dense in V.

It is easy to see that the map $\pi \circ \rho^{-1} \circ (\pi^\flat)^{-1}$ induces a bijection from the set $\mathrm{Per}^\flat$ of periodic points of Φ^s in $\mathbb{P}_{K^\flat}^N(K^\flat)$ to the set Per of periodic points of F in $\mathbb{P}_K^N(K)$. We note that the set of periodic points of Φ^s in $\mathbb{P}_{K^\flat}^N(K^\flat)$ is exactly the set of points defined over k, i.e., the image of $\eta : \mathbb{P}_k^N(k) \hookrightarrow \mathbb{P}_{K^\flat}^N(K^\flat)$. We have a reduction map $\mathrm{red} : \mathbb{P}_K^N(K) \to \mathbb{P}_k^N(k)$. The map $\eta \circ \mathrm{red} : \mathrm{Per} \to \mathrm{Per}^\flat$ is bijective. Moreover, we have that $(\eta \circ \mathrm{red}) \circ (\pi \circ \rho^{-1} \circ (\pi^\flat)^{-1})$ is the identity on $\mathrm{Per}^\flat$.

Denote by $S^\flat$ the Zariski closure of $\eta \circ \mathrm{red}(V \cap \mathrm{Per})$. Since $S^\flat$ is defined over k, it is periodic under Φ^s. The main ingredient of our proof is to show that $S^\flat$ is a subset of $\pi^\flat(\rho(\pi^{-1}(V)))$. If $\pi^\flat(\rho(\pi^{-1}(V)))$ is algebraic, this is obvious. But, a priori, $\pi^\flat(\rho(\pi^{-1}(V)))$ is not algebraic, since the map ρ is very transcendental. Our strategy is to approximate $\pi^\flat(\rho(\pi^{-1}(V)))$ by algebraic subvarieties of $\mathbb{P}_{K^\flat}^N$. For simplicity, assume that V is an hypersurface of $\mathbb{P}_K^N$. Applying the approximation lemma of Scholze [2012, Corollary 6.7], for any $\epsilon > 0$, there exists an algebraic hypersurface H_ϵ of $\mathbb{P}_{K^\flat}^N$ which is ϵ-close to $\pi^\flat(\rho(\pi^{-1}(V)))$. Then $\eta \circ \mathrm{red}(V \cap \mathrm{Per})$ is ϵ-close to H_ϵ. Since $S^\flat$ is the Zariski closure of $\eta \circ \mathrm{red}(V \cap \mathrm{Per})$ in $\mathbb{P}_{K^\flat}^N$, we can show that it is ϵ-close to H_ϵ. Then we can show that $S^\flat$ is contained in $\pi^\flat(\rho(\pi^{-1}(V)))$ by letting ϵ tends to 0. Then we have $S := \pi(\rho^{-1}((\pi^\flat)^{-1}(S^\flat))) \subseteq V$. Since S is periodic and Zariski dense in V, it follows that V is periodic.

In this paper, we mainly consider the lifts of Frobenius on $\mathbb{P}_{\mathbb{C}_p}^N$ for simplicity, since the aim of this paper is to present a new method in dynamics. We suspect that our method can be applied to the more general case where F is a lift of Frobenius on any projective variety over $\mathbb{C}_p$. On the other hand, a lift of Frobenius on a projective variety X can often be extended to some lift of Frobenius on $\mathbb{P}_{\mathbb{C}_p}^N$ for some embedding $\tau : X \hookrightarrow \mathbb{P}_{\mathbb{C}_p}^N$. In the Appendix, we prove the existence of such embedding for *polarized* lifts of Frobenius on some projective varieties under some technical condition on the field of definition. Once this happens, many questions can be reduced to the special case where $X = \mathbb{P}_{\mathbb{C}_p}^N$.

The plan of the paper. The paper is organized as follows. In Section 2, we gather a number of results on the perfectoid spaces in Scholze's papers Scholze 2012; Scholze 2014. In Section 3, we construct the inverse limit and make it a perfectoid space with an automorphism. We also construct its tilt and give the isomorphism between these two topological dynamical systems. In Section 4, we study the periodic points of F. In particular, we prove Theorems 1.1 and 1.2. In Section 5, we study the coherent backward orbits of a point. In particular, we prove Theorems 1.7 and 1.8 and Corollary 1.9. In the Appendix, we study the polarized lift of Frobenius on projective varieties over $\mathbb{C}_p$.

2. Preliminary: perfectoid spaces

In this section, we introduce some necessary background in perfectoid spaces. All the results in this section can be found in Scholze's papers [2012; 2014]. The perfectoid spaces are some nonarchimedean analytic spaces. Following the technique of Scholze [2012], we work with Huber's adic spaces [1993; 1994; 1996].

Adic spaces. In this section, we denote by k a complete nonarchimedean field i.e., a complete topological field whose topology is induced by a nontrivial norm $|\cdot| : k \to [0, \infty)$. Denote by R a topological k-algebra. Moreover we suppose that R is a Tate k-algebra i.e., there exists a subring $R_0 \subseteq R$, such that aR_0, $a \in k^\times$, forms a basis of open neighborhoods of 0.

A subset $M \subseteq R$ is call bounded if $M \subseteq aR_0$, for some $a \in k^\times$. An element $x \in R$ is called power-bounded if $\{x^n \mid n \geq 0\} \subseteq R$ is bounded. Let $R^\circ \subseteq R$ be the subring of power-bounded elements.

Definition 2.1 [Scholze 2012]. An affinoid k-algebra is a pair (R, R^+), where R is a Tate k-algebra and R^+ is an open and integrally closed subring of R°.

A valuation on R is a map $|\cdot| : R \to \Gamma \cup \{0\}$, where Γ is a totally ordered abelian group, such that, $|0| = 0$, $|1| = 1$, $|xy| = |x||y|$ and $|x + y| \leq \max\{|x|, |y|\}$. We say that $|\cdot|$ is continuous, if for all $\gamma \in \Gamma$, the subset $\{x \in R : |x| < \gamma\} \subseteq R$ is open.

To a pair (R, R^+), Huber associates a space $\mathrm{Spa}(R, R^+)$ of equivalence classes of continuous valuations $|\cdot|$ on R such that $|R^+| \leq 1$, and calls it an *affinoid space*.

For a point $x \in \mathrm{Spa}(R, R^+)$, we denote by $f \to |f(x)|$ the associated valuation. It is a fact [Scholze 2012, Proposition 2.12.(iii)] that

$$R^+ = \{f \in R : |f(x)| \leq 1 \text{ for all } x \in \mathrm{Spa}(R, R^+)\}.$$

We equip $\mathrm{Spa}(R, R^+)$ with the topology generated by rational subsets:

$$U(f_1, \ldots, f_n; g) = \{x \in \mathrm{Spa}(R, R^+) : |f_i(x)| \leq |g(x)|\} \subseteq \mathrm{Spa}(R, R^+),$$

where $f_1, \ldots, f_n \in R$ generate R as an ideal and $g \in R$.

The completion $(\hat{R}, \hat{R}^+)$ of an affinoid algebra (R, R^+) is also an affinoid algebra. Then we recall [Huber 1993, Proposition 3.9].

Proposition 2.2. *We have* $\mathrm{Spa}(\hat{R}, \hat{R}^+) \simeq \mathrm{Spa}(R, R^+)$, *identifying rational subsets.*

We say a point $x \in \mathrm{Spa}(R, R^+)$ is a k-point, if the valuation x is induced by a morphism from R to k i.e., there exists a morphism $\phi : R \to k$ such that for any $f \in R$, $|f(x)| = |\phi(f)|$.

Roughly speaking, *adic spaces* over K are the objects obtained by gluing affinoid spaces. The morphisms between the adic spaces are the morphisms glued by the morphisms between affinoid spaces. Because in this paper we only consider some very concrete adic spaces, we give only a very brief definition of the adic spaces. One may find a detailed definition in [Huber 1994].

On an affinoid space $X = \mathrm{Spa}(R, R^+)$, one may define presheaves O_X and O_X^+ on X. Since we do not use these presheaves in this paper, we omit their definition. We do *not* know whether O_X is a sheaf in general. We note that once O_X is a sheaf, O_X^+ is a sheaf also. However, if (R, R^+) is of *topological finite type* then O_X is a sheaf.[5] Assume that O_X is a sheaf on X. For any $x \in X$, the valuation $f \mapsto |f(x)|$ extends to the stalk $O_{X,x}$, and we have $O_{X,x}^+ = \{f \in O_{X,x} : |f(x)| \leq 1\}$. The affinoid spaces X defines a triple $(X, O_X, |\cdot(x)| : x \in X)$.

An adic space over k is a triple $(Y, O_Y, |\cdot(x)| : x \in Y)$, consisting of a locally ringed topological space (Y, O_Y) where O_Y is a sheaf of complete topological k-algebras, and a continuous valuation $|\cdot(x)|$ on $O_{X,x}$ for every $x \in X$, which is locally on Y an affinoid adic space.

Let X be an affinoid space. We say a point $x \in X$ is a k-point if it is a k-point in any (and thus all) affinoid neighborhood of X.

Perfectoid fields. Denote by K a complete nonarchimedean field of residue characteristic $p > 0$ with norm $|\cdot| : K \to \mathbb{R}_{\geq 0}$. Denote by $K^\circ := \{x \in K : |x| \leq 1\}$ its valuation ring.

Definition 2.3. We say K is a *perfectoid field* if $|K| \subseteq \mathbb{R}_{\geq 0}$ is dense in $\mathbb{R}_{\geq 0}$ and the Frobenius map $\Phi : K^\circ/p \to K^\circ/p$ is surjective.

Observe that $\mathbb{C}_p$ and $\widehat{\mathbb{F}_p((t))}$ are perfectoid fields. Set

$$\mathbb{Q}_p(p^{1/p^\infty}) := \bigcup_{i \geq 0} \mathbb{Q}_p(p^{1/p^i}) \quad \text{and} \quad \mathbb{F}_p((t))(t^{1/p^\infty}) := \bigcup_{i \geq 0} \mathbb{F}_p((t))(t^{1/p^i}).$$

Then their completions $\mathbb{Q}_p\widehat{(p^{1/p^\infty}})$ and $\mathbb{F}_p\widehat{((t))(t^{1/p^\infty}})$ are perfectoid fields. Note that $\mathbb{Q}_p$ is not a perfectoid field, since $|\mathbb{Q}_p| = \{0\} \cup \{p^i \mid i \in \mathbb{Z}\} \subseteq \mathbb{R}_{\geq 0}$ is not dense.

For any perfectoid field K, we choose some element $\omega \in K^\times$ such that $|p| \leq |\omega| < 1$. We define

$$K^{\flat\circ} := \varprojlim_{x \mapsto \Phi(x)} K^\circ/\omega.$$

Recall [Scholze 2012, Lemma 3.2].

---

[5] An affinoid k-algebra (R, R^+) is said to be of topological finite type if R is a quotient of $k\{T_1, \dots, T_n\}$ for some n, and $R^+ = R^\circ$.

Lemma 2.4. (i) *There exists a multiplicative homeomorphism*

$$\varprojlim_{x\mapsto x^p} K^\circ \xrightarrow{\sim} \varprojlim_{x\mapsto \Phi(x)} K^\circ/\omega = K^{\flat\circ}$$

given by projection. Moreover, we have a map

$$K^{\flat\circ} = \varprojlim_{x\mapsto x^p} K^\circ = \{(x^{(0)}, x^{(1)}, \dots) \mid x^{(i)} \in K^\circ, (x^{(i+1)})^p = x^i\} \to K^\circ$$

defined by

$$x = (x^{(0)}, x^{(1)}, \dots) \to x^\# := x^{(0)}.$$

We may define a norm on $K^{\flat\circ}$ *by* $|x^\#| = |x|$ *for all* $x \in K^{\flat\circ}$.

(ii) *The addition on*

$$K^{\flat\circ} = \{x := (x^{(0)}, x^{(1)}, \dots) \mid x^{(i)} \in K^\circ, (x^{(i+1)})^p = x^i\}$$

is given by $(x+y)^i = \lim_{n\to\infty}(x^{(i+n)} + y^{(i+n)})^{p^n}$.

(iii) *There exists an element* $\omega^\flat \in \varprojlim_{x\mapsto x^p} K^\circ$, *satisfying* $(\omega^\flat)^\# = \omega$. *Define*

$$K^\flat := K^{\flat\circ}[(\omega^\flat)^{-1}].$$

Then norm $|\cdot|$ *on* $K^{\flat\circ}$ *extends to a norm on* $K^\flat$ *which makes* $K^{\flat\circ}$ *the valuation ring of* $K^\flat$.

(iv) *There exists a multiplicative homeomorphism*

$$K^\flat \xrightarrow{\sim} \varprojlim_{x\mapsto x^p} K.$$

Then $K^\flat$ *is a perfectoid field of characteristic* p. *We have* $|K^{b\times}| = |K^\times|$, $K^{\flat\circ}/\omega^\flat \simeq K^\circ/\omega$, *and* $K^{\flat\circ}/\mathfrak{m}^\flat \simeq K^\circ/\mathfrak{m}$, *where* $\mathfrak{m}$ *and* $\mathfrak{m}^\flat$ *are the maximal ideals of* K° *and* $K^{\flat\circ}$, *respectively.*

(v) *If* K *is of characteristic* p, *then* $K^\flat = K$.

We note that (i) and (ii) of Lemma 2.4 implies that the definition of $K^{\flat\circ}$ is independent of ω.

We call $K^\flat$ the *tilt* of K.

Example 2.5. The tilt of $\mathbb{C}_p$ is $\mathbb{C}_p^\flat = \widehat{\overline{\mathbb{F}_p((t))}}$.

Then we have the following theorem, which was known by the classical work of Fontaine and Wintenberger [1979]

Theorem 2.6. (i) *Let* L *be a finite extension of* K. *Then* L *with its natural topology induced by* K *is a perfectoid field.*

(ii) *The tilt functor* $L \mapsto L^\flat$ *induces an equivalence of categories between the category of finite extensions of* K *and the category of finite extensions of* $K^\flat$. *This equivalence preserves degrees.*

Almost mathematics. Let K be a perfectoid field and $\mathfrak{m}$ be the maximal ideal of K°.

A K°-module M is said to be *almost zero* if $\mathfrak{m}M = 0$. Define the category of almost K°-modules as

$$K^{\circ a}\text{–mod}\ K^{\circ a}-\ \text{mod} := K^\circ\text{–mod}\,(\mathfrak{m}\text{-torsion}).$$

We have a localization functor $M \mapsto M^a$ from K°-mod to $K^{\circ a}$-mod, whose kernel is the thick subcategory of almost zero modules.

For two $K^{\circ a}$-modules X and Y, we define $\mathrm{alHom}(X, Y) = \mathrm{Hom}(X, Y)^a$.

Proposition 2.7 [Gabber and Ramero 2003]. *The category* $K^{\circ a}\text{–mod}$ *is an abelian tensor category, where we define kernels, cokernels and tensor products in the unique way compatible with their definition in* $K^\circ\text{–mod}$, *that is*

$$M^a \otimes N^a = (M \otimes N)^a$$

for any two K°*-modules* M *and* N. *For any* $L, M, N \in K^{\circ a}\text{–mod}$ *there is a functorial isomorphism*

$$\mathrm{Hom}(L, \mathrm{alHom}(M, N)) = \mathrm{Hom}(L \otimes M, N).$$

This means that $K^{\circ a}$-mod has all properties of the category of modules over a ring and thus one can define the notion of $K^{\circ a}$-algebra. Any K°-algebra R defines a $K^{\circ a}$-algebra R^a as the tensor products are compatible. Moreover, localization also gives a functor from R-modules to R^a-modules.

Proposition 2.8 [Gabber and Ramero 2003]. *There exists a right adjoint functor*

$$K^{\circ a}\text{–mod} \to K^\circ\text{–mod} : M \mapsto M_* := \mathrm{Hom}_{K^{\circ a}}(K^{\circ a}, M)$$

to the localization functor $M \mapsto M^a$. *The adjunction morphism* $(M_*)^a \to M$ *is an isomorphism. If* M *is a* K°*-module, then* $(M^a)_* = \mathrm{Hom}(\mathfrak{m}, M)$.

If A is a $K^{\circ a}$-algebra, then A_* has a natural structure as K°-algebra and $(A^a)_* = A$. In particular, any $K^{\circ a}$-algebra comes via localization from a K°-algebra. Furthermore the functor $M \mapsto M_*$ induces a functor from A-modules to A_*-modules, and one can see also that all A-modules come via localization from A_*-modules. The category of A-modules is again an abelian tensor category, and all properties about the category of $K^{\circ a}$-modules stay true for the category of A-modules.

Let A be any $K^{\circ a}$-algebra. As in [Scholze 2012], an A-module M is said to be flat if the functor $X \mapsto M \otimes_A X$ on A-modules is exact.

Denote by ω an element in K° satisfying $|p| \le |\omega| < 1$. Let A be a K^a-algebra, we say A is ω-adically complete if $A \simeq \varprojlim A/\omega^n$.

Perfectoid algebras. Fix a perfectoid field K and an element $\omega \in K^\circ$ satisfying $|p| \le |\omega| < 1$.

Definition 2.9. (i) A *perfectoid* K*-algebra* is a Banach K-algebra R such that the subset $R^\circ \subseteq R$ of powerbounded elements is open and bounded, and the Frobenius morphism $\Phi : R^\circ/\omega \to R^\circ/\omega$ is surjective. Morphisms between perfectoid K-algebras are the continuous morphisms of K-algebras.

(ii) A *perfectoid $K^{\circ a}$-algebra* is a ω-adically complete flat $K^{\circ a}$-algebra A on which Frobenius induces an isomorphism

$$\Phi : A/\omega^{1/p} \simeq A/\omega.$$

Morphisms between perfectoid $K^{\circ a}$-algebras are the morphisms of $K^{\circ a}$-algebras.

(iii) A *perfectoid $K^{\circ a}/\omega$-algebra* is a flat $K^{\circ a}/\omega$-algebra $\bar{A}$ on which Frobenius induces an isomorphism

$$\Phi : \bar{A}/\omega^{1/p} \simeq \bar{A}.$$

Morphisms are the morphisms of $K^{\circ a}/\omega$-algebras.

Let K-Perf denote the category of perfectoid K-algebras and similarly for $K^{\circ a}$-Perf and $K^{\circ a}/\omega$-Perf. Let $K^\flat$ be the tilt of K and $\omega^\flat$ is an element in $K^\flat$ satisfying $(\omega^b)^\# = \omega$.

We recall [Scholze 2012, Theorem 5.2].

Theorem 2.10. *We have the following series of equivalences of categories*:

$$K\text{-Perf} \simeq K^{\circ a}\text{-Perf} \simeq (K^{\circ a}/\omega)\text{-Perf} = (K^{\flat a}/\omega^b)\text{-Perf} \simeq K^{\flat a}\text{-Perf} \simeq K^\flat\text{-Perf}.$$

In other words, a perfectoid K-algebra, which is an object over the generic fiber, has a canonical extension to the almost integral level as a perfectoid $K^{\circ a}$-algebra, and perfectoid $K^{\circ a}$-algebras are determined by their reduction modulo ω.

Let R be a perfectoid $K^{\circ a}$-algebra, with $A = R^{\circ a}$. Define

$$A^\flat := \varprojlim_{\Phi} A/\omega,$$

which we regard as a $K^{\flat\circ a}$-algebra via

$$K^{\flat\circ a} = (\varprojlim_{\Phi} K^\circ/\omega)^a = \varprojlim_{\Phi}(K^\circ/\omega)^a = \varprojlim_{\Phi} K^{\circ a}/\omega,$$

and set $R^\flat = A_*^b[(\omega^\flat)^{-1}]$.

Proposition 2.11. *This defines a perfectoid $K^\flat$-algebra $R^\flat$ with corresponding perfectoid $K^{\flat\circ a}$-algebra $A^\flat$, and $R^\flat$ is the tilt of R. Moreover,*

$$R^\flat = \varprojlim_{x\mapsto x^p} R, \quad A_*^\flat = \varprojlim_{x\mapsto x^p} A_*, \quad \textit{and} \quad A_*^\flat/\omega^\flat \simeq A_*/\omega.$$

In particular, we have a continuous multiplicative map $R^\flat = \varprojlim_{x\mapsto x^p} R \to R$,

$$x = (x^{(0)}, x^{(1)}, \dots) \mapsto x^\# := x^{(0)}.$$

Then the equivalence K-Perf $\to K^\flat$-Perf in Theorem 2.10 is given by $R \mapsto R^\flat$.

Proposition 2.12. *Let $R = K\langle T_1^{1/p^\infty}, \dots, T_n^{1/p^\infty}\rangle = K^\circ\widehat{[T_1^{1/p^\infty}, \dots, T_n^{1/p^\infty}]}[\omega^{-1}]$. Then R is a perfectoid K-algebra, and its tilt $R^\flat$ is given by $K^\flat\langle T_1^{1/p^\infty}, \dots, T_n^{1/p^\infty}\rangle$.*

Perfectoid spaces. Fix a perfectoid field K and an element $\omega \in K^\circ$ satisfying $|p| \le |\omega| < 1$. Let $K^\flat$ be the tilt of K and $\omega^\flat$ is an element in $K^\flat$ satisfying $(\omega^b)^\# = \omega$.

Definition 2.13. A *perfectoid affinoid K-algebra* is an affinoid K-algebra (R, R^+), where R is a perfectoid K-algebra, and $R^+ \subseteq R^\circ$ is an open and integrally closed subring.

Proposition 2.14. *The association $(R, R^+) \mapsto (R^b, R^{\flat +})$, where $R^{\flat +} = \varprojlim_{(x \to x^p)} R^+$. defines an equivalence between the category of perfectoid affinoid K-algebras and the category of perfectoid affinoid $K^\flat$-algebras.*

Theorem 2.15. *For any $x \in \operatorname{Spa}(R, R^+)$, one may define a point $x^\flat \in \operatorname{Spa}(R^\flat, R^{\flat +})$ by setting $|f(x^\flat)| := |f^\#(x)|$ for $f \in R^\flat$. This defines a homeomorphism $\rho : \operatorname{Spa}(R, R^+) \xrightarrow{\sim} \operatorname{Spa}(R^\flat, R^{\flat +})$ preserving rational subsets.*

Denote by $X := \operatorname{Spa}(R, R^+)$ and $X^\flat := \operatorname{Spa}(R^\flat, R^{\flat +})$. We note that in general the map $R^\flat \to R : f \to f^\#$ is not surjective. For any f in R, $\rho_* f := f \circ \rho^{-1}$ is a continuous function on $X^\flat$ but in general is not contained in $R^\flat$.

We have the following approximation lemma [Scholze 2012, Corollary 6.7].

Lemma 2.16. *For any $f \in R$ and any $c \ge 0$, $\epsilon > 0$, there exists $g_{c,\epsilon} \in R^\flat$ such that for all $x \in X$, we have*

$$|f(x) - g^\#_{c,\epsilon}(x)| \le |\omega|^{1-\epsilon} \max(|f(x)|, |\omega|^c).$$

Remark 2.17. Note that for $\epsilon < 1$, the given estimate says in particular that for all $x \in X$, we have

$$\max\{|f(x)|, |\omega|^c\} = \max\{|g^\#_{c,\epsilon}(x)|, |\omega|^c\}.$$

Remark 2.18. Let $R := K\langle x_1, \dots, x_N\rangle$ and $R^+ := R^\circ = K^\circ\langle x_1, \dots, x_N\rangle$. Then $R^\flat = K^\flat\langle x_1, \dots, x_N\rangle$ and $R^{\flat +} = K^{\flat\circ}\langle x_1, \dots, x_N\rangle$.

By Lemma 2.16, for any $c \in \mathbb{Z}^+$, there exists an element $g_c \in K^{\flat\circ}\langle x_1^{1/p^\infty}, \dots, x_N^{1/p^\infty}\rangle$ such that for all $x \in U_0^{\mathrm{perf}}$, we have

$$|H \circ \pi(x) - g_c^\#(x)| \le |p|^{1/2} \max(|H \circ \pi(x)|, |p|^c).$$

There exists $\ell \in \mathbb{N}$ and an element $G_c \in K^{\flat\circ}[x_1^{1/p^\ell}, \dots, x_N^{1/p^\ell}]$ such that $g_c - G_c \in t^{c+1} K^{\flat\circ}\langle x_1^{1/p^\infty}, \dots, x_N^{1/p^\infty}\rangle$. It follows that for all $x \in U_0^{\mathrm{perf}}$, we have

$$|H \circ \pi(x) - G_c^\#(x)| \le |p|^{1/2} \max(|H(x)|, |p|^c) = |p|^{1/2} \max(|G_c^\#(x)|, |p|^c),$$

and $G^{p^\ell} \in K^{\flat\circ}[x_1, \dots, x_N]$.

Moreover, when $K^\flat = \widehat{\overline{\mathbb{F}_p((t))}}$, we may arrange to have $G_c \subseteq E^\circ[x_1^{1/p^\ell}, \dots, x_N^{1/p^\ell}]$, where E is a finite extension of $\overline{\mathbb{F}_p}((t))$.

We next describe the structure sheaf O_X on $X := \operatorname{Spa}(R, R^+)$. Let $U = U(f_1, \dots, f_n; g) \subseteq X$ be a rational subset. Equip $R[g^{-1}]$ with the topology making the image of $R^+[f_1/g, \dots, f_n/g] \to R[g^{-1}]$

open and bounded. Let $R\langle f_1/g, \ldots, f_n/g\rangle$ be the completion of $R[g^{-1}]$ with respect to this topology. It is equipped with a subring

$$R\langle f_1/g, \ldots, f_n/g\rangle^+ \subseteq R\langle f_1/g, \ldots, f_n/g\rangle$$

which is the completion of the integral closure of $R^+[f_1/g, \ldots, f_n/g]$. By [Huber 1994, Proposition 1.3], the pair $(O_X(U), O_X^+(U)) := (R\langle f_1/g, \ldots, f_n/g\rangle, R\langle f_1/g, \ldots, f_n/g\rangle^+)$ depends only on the rational subset $U \subseteq X$ (and not on the choice of $f_1, \ldots, f_n, g \in R$). The map

$$\mathrm{Spa}(O_X(U), O_X^+(U)) \to \mathrm{Spa}(R, R^+)$$

is a homeomorphism onto U, preserving rational subsets. Moreover, $(O_X(U), O_X^+(U))$ is initial with respect to this property.

By [Scholze 2012, Theorem 6.3], we have the following:

Theorem 2.19. *For any rational subset $U \subseteq X$, let $U^\flat := \rho(U) \subseteq X^\flat$.*

(i) *The presheaves O_X and $O_{X^\flat}$ are sheaves.*

(ii) *For any rational subset $U \subseteq X$, the pair $(O_X(U), O_X^+(U))$ is a perfectoid affinoid K-algebra, which tilts to $(O_{X^\flat}(U^\flat), O_{X^\flat}^+(U^\flat))$.*

The resulting spaces $\mathrm{Spa}(R, R^+)$, equipped with the two structure sheaves of topological rings O_X and O_X^+, are called *affinoid perfectoid spaces* over K. The morphisms between the affinoid perfectoid spaces over K are the morphisms induced by the morphisms between affinoid perfectoid K-algebras.

One defines *perfectoid spaces* over K to be the objects obtained by gluing affinoid perfectoid spaces. The morphisms between the perfectoid spaces are the morphisms glued by the morphisms between affinoid perfectoid spaces.

We say that a perfectoid space $X^\flat$ over $K^\flat$ is the tilt of a perfectoid space X over K if there exists a functorial isomorphism $\mathrm{Hom}(\mathrm{Spa}(R^\flat, R^{\flat+}), X^\flat) = \mathrm{Hom}(\mathrm{Spa}(R, R^+), X)$ for all perfectoid affinoid K-algebras (R, R^+) with tilts $(R^\flat, R^{\flat+})$.

Theorem 2.20. *Any perfectoid space X over K admits a tilt $X^\flat$, unique up to isomorphism. This induces an equivalence between the category of perfectoid spaces over K and the category of perfectoid spaces over $K^\flat$. The underlying topological spaces of X and $X^\flat$ are naturally identified by ρ. A perfectoid space X is affinoid perfectoid if and only if its tilt $X^\flat$ is affinoid perfectoid. Finally, for any affinoid perfectoid subspace $U \subseteq X$, the pair $(O_X(U), O_X^+(U))$ is a perfectoid affinoid K-algebra with tilt $(O_{X^\flat}(U^\flat), O_{X^\flat}^+(U^\flat))$.*

For any morphism $F : X \to Y$ between perfectoid spaces over K, denote by $F^\flat : X^\flat \to Y^\flat$ the morphism between perfectoid spaces over $K^\flat$ induced by the equivalence of categories.

Points in perfectoid spaces. Fix a perfectoid field K and an element $\omega \in K^\circ$ satisfying $|p| \le |\omega| < 1$. Let X be a perfectoid space over K.

For any point $x \in X$, let $K(x)$ be the residue field of $O_{X,x}$ and $K(x)^+ \subseteq K(x)$ be the image of $O^+_{X,x}$. By [Scholze 2012, Proposition 2.25], the ω-adic completion of $O^+_{X,x}$ is equal to the ω-adic completion $\widehat{K(x)}^+$ of $\widehat{K(x)}^+$. By [Scholze 2012, Corollary 6.7], $\widehat{K(x)}$ is a perfectoid field.

Definition 2.21. An affinoid perfectoid field is a pair (K, K^+) consisting of a perfectoid field and an open valuation subring $K^+ \subseteq K$.

Then $(\widehat{K(x)}, \widehat{K(x)}^+)$ is an affinoid perfectoid field. Also note that affinoid perfectoid fields (L, L^+) for which $K \subseteq L$ are affinoid K-algebras.

Then we have the following description of points [Scholze 2012, Proposition 2.27].

Proposition 2.22. *The points of X are in bijection with maps $\iota : \mathrm{Spa}(L, L^+) \to X$ to affinoid fields (L, L^+) such that the quotient field of the image of $O_{X,x}$ in L is dense, where x is the image of* $\mathrm{Spa}(L, L^+)$ *in X.*

Any point $x \in X$ associates to a map $\iota : \mathrm{Spa}(\widehat{K(x)}, \widehat{K(x)}^+) \to X$. By the equivalence of categories, the point $x^\flat \in X^\flat$ associates to a map

$$\iota^\flat : \mathrm{Spa}(\widehat{K(x)}^\flat, \widehat{K(x)}^{\flat+}) \to X.$$

By [Scholze 2012, Lemma 5.21], $\mathrm{Spa}(\widehat{K(x)}^\flat, \widehat{K(x)}^{\flat+})$ is an affinoid perfectoid field. It follows that $\widehat{K^\flat(x^\flat)} = (\widehat{K(x)})^\flat$.

In particular, we have the following:

Lemma 2.23. *For any point $x \in X$, x is a K-point if and only if $x^\flat$ is a $K^\flat$-point in $X^\flat$.*

3. Inverse limit of lifts of Frobenius

In this section, fix a perfectoid field K. Denote by $p > 0$ the characteristic of the residue field $K^\circ/K^{\circ\circ}$ of K.

Let $F : \mathbb{P}^N_K \to \mathbb{P}^N_K$ be a lift of Frobenius i.e., an endomorphism taking the form

$$F : [x_0 : \cdots : x_N] \mapsto [x_0^q + p' P_0(x_0, \ldots, x_N) : \cdots : x_N^q + p' P_N(x_0, \ldots, x_N)],$$

where $p' \in K^{\circ\circ}$, $q = p^s$ is a power of p, and $P_0, \ldots, P_N$ are homogeneous polynomials of degree q in $K^\circ[x_0, \ldots, x_N]$. Let $\omega \in K^\circ$ be an element satisfying $\max\{|p'|, |p|\} \le |\omega| < 1$.

Adic projective spaces. At first, we define an adic space $\mathbb{P}^{N,\mathrm{ad}}_K$ which associates to the projective space $\mathbb{P}^N_K$. In fact, by [Scholze 2012, Theorem 2.22], for any projective variety X defined over K with an integral model $\mathfrak{X}$ over K°, we may associate an adic space X^{ad}. But in this paper, we don't need the general theory and we define $\mathbb{P}^{N,\mathrm{ad}}_K$ in the following explicit way:

For any $i \in \{0, \ldots, N\}$, denote by

$$U_i^{\mathrm{ad}} := \mathrm{Spa}\big(K\langle z_{i,0}, \ldots, z_{i,i-1}, z_{i,i+1}, \ldots, z_{i,N}\rangle, K\langle z_{i,0}, \ldots, z_{i,i-1}, z_{i,i+1}, \ldots, z_{i,N}\rangle^\circ\big)$$

the unit balls. Then we define $\mathbb{P}_K^{N,\mathrm{ad}}$ by gluing the unit balls U_i^{ad} together in the usual way: For any $i \neq j$, $U_i^{\mathrm{ad}} \cap U_j^{\mathrm{ad}} = U(1, z_{i,0}, \ldots, z_{i,i-1}, z_{i,i+1}, \ldots, z_{i,N}; z_{i,j}) \subseteq U_i^{\mathrm{ad}}$. On $U_i^{\mathrm{ad}} \cap U_j^{\mathrm{ad}}$, the transition map $\phi_{i,j}$ is defined by

$$\phi_{i,j}^*(z_{j,k}) = z_{i,k}/z_{i,j} \text{ for } k \neq i, j, \quad \text{and} \quad \phi_{i,j}^*(z_{j,i}) = 1/z_{i,j}.$$

Denote by $R(\mathbb{P}_K^{N,\mathrm{ad}})$ the set of K-points in $\mathbb{P}_K^{N,\mathrm{ad}}$.

Lemma 3.1. *There exists a natural embedding $\tau : \mathbb{P}_K^N(K) \hookrightarrow \mathbb{P}_K^{N,\mathrm{ad}}$. Moreover its image $\tau(\mathbb{P}_K^N(K)) = R(\mathbb{P}_K^{N,\mathrm{ad}})$.*

Proof of Lemma 3.1. For any point $q \in \mathbb{P}_K^N(K)$, there exists a finite extension L of K, and a point $q' = [x_0 : \cdots : x_N] \in \mathbb{P}_L^N(L)$, such that q is the image of that q' under the natural morphism $\pi_K^L : \mathbb{P}_L^N \to \mathbb{P}_K^N$ induced by the inclusion $K \hookrightarrow L$. Indeed, $(\pi_K^L)^{-1}$ is exactly the Galois orbit of q'. We may suppose that $\max\{|x_0|, \ldots, |x_N|\} = 1$ for all $j = 0, \ldots, N$. Denote by $I_q := \{i : |x_i| = 1\}$. Observe that I_q depends only on q. Pick $i \in I_q$, we define $\tau(q) \in U_i$ to be the point defined by $f \to |f(x_1/x_i, \ldots, x_n/x_i)|$, for all $f \in K\langle z_{i,0}, \ldots, z_{i,i-1}, z_{i,i+1}, \ldots, z_{i,N}\rangle$. Here $f(x_1/x_i, \ldots, x_n/x_i) \in L$ depends on the choice of q' in its Galois obit, but the value $|f(x_1/x_i, \ldots, x_n/x_i)|$ depends only on q. Moreover we may check that the definition of $\tau(q)$ does not depend on the choice of $i \in I_q$. Then τ is well defined. Moreover it is easy to check that τ is injective and $\tau(\mathbb{P}_K^N(K)) \subseteq R(\mathbb{P}_K^{N,\mathrm{ad}})$. By [Bosch et al. 1984, 6.1.2 Corollary 3], the map $\tau : \mathbb{P}_K^N(K) \to R(\mathbb{P}_K^{N,\mathrm{ad}})$ is surjective. ☐

***Lifts of Frobenius on* $\mathbb{P}_K^{N,\mathrm{ad}}$.** The endomorphism F induces a natural endomorphism F^{ad} on $\mathbb{P}_K^{N,\mathrm{ad}}$. We define F^{ad} in the following explicit way. For any $i = 0, \ldots, N$, $F^{\mathrm{ad}}|_{U_i^{\mathrm{ad}}} : U_i^{\mathrm{ad}} \to U_i^{\mathrm{ad}}$ is defined to be

$$F^*(z_{i,j}) = \frac{z_{i,j}^q + p' P_j(z_{i,0}, \ldots, z_{i,i-1}, 1, z_{i,i+1}, \ldots, z_{i,N})}{1 + p' P_i(z_{i,0}, \ldots, z_{i,i-1}, 1, z_{i,i+1}, \ldots, z_{i,N})}$$

for all $j \neq i$. We may write

$$\frac{z_{i,j}^q + p' P_j(z_{i,0}, \ldots, z_{i,i-1}, 1, z_{i,i+1}, \ldots, z_{i,N})}{1 + p' P_i(z_{i,0}, \ldots, z_{i,i-1}, 1, z_{i,i+1}, \ldots, z_{i,N})} = z_{i,j}^q + p' Q_{i,j}(z_{i,0}, \ldots, z_{i,i-1}, z_{i,i+1}, \ldots, z_{i,N})$$

where $Q_{i,j} \in K^\circ\langle z_{i,0}, \ldots, z_{i,i-1}, z_{i,i+1}, \ldots, z_{i,N}\rangle$. For any $i \neq j$, we may check that $F_i^{\mathrm{ad}}(U_i^{\mathrm{ad}} \cap U_j^{\mathrm{ad}}) \subseteq U_i^{\mathrm{ad}} \cap U_j^{\mathrm{ad}}$ and

$$F_i^{\mathrm{ad}}|_{U_i^{\mathrm{ad}} \cap U_j^{\mathrm{ad}}} = F_j^{\mathrm{ad}}|_{U_i^{\mathrm{ad}} \cap U_j^{\mathrm{ad}}}.$$

Then we may glue these F_i^{ad} to define $F^{\mathrm{ad}} : \mathbb{P}_K^{N,\mathrm{ad}} \to \mathbb{P}_K^{N,\mathrm{ad}}$. Observe that we have the following commutative diagram:

$$\begin{tikzcd}
\mathbb{P}_K^N(K) \arrow[r, "\tau"] \arrow[d, "F|_{\mathbb{P}_K^N(K)}"'] & \mathbb{P}_K^{N,\mathrm{ad}} \arrow[d, "F^{\mathrm{ad}}"] \\
\mathbb{P}_K^N(K) \arrow[r, "\tau^\flat"] & \mathbb{P}_K^{N,\mathrm{ad}}
\end{tikzcd}$$

Now we identify $\mathbb{P}^N_K(K)$ and $\mathbb{P}^N_{K^\flat}(K^\flat)$ with the image of τ and $\tau^\flat$ in $\mathbb{P}^{N,\mathrm{ad}}_K$ $\mathbb{P}^{N,\mathrm{ad}}_{K^\flat}$, respectively.

The inverse limit. The inverse limit $\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}^{N,\mathrm{ad}}_K$ is the topological space $\{(x_0, x_1, \dots) \in (\mathbb{P}^{N,\mathrm{ad}}_K)^{\mathbb{N}} \mid F^{\mathrm{ad}}(x_i) = x_{i-1} \text{ for all } i \geq 1\}$ with the product topology. There exists a natural automorphism T on $\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}^{N,\mathrm{ad}}_K$ defined by

$$T : (x_0, x_1, \dots) \to (F^{\mathrm{ad}}(x_0), x_0, x_1, \dots).$$

The aim of this section is to construct a perfectoid space $(\mathbb{P}^N_K)^{\mathrm{perf}}$ with an automorphism F^{perf} such that the topological dynamical system $((\mathbb{P}^N_K)^{\mathrm{perf}}, F^{\mathrm{perf}})$ is isomorphic to $(\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}^{N,\mathrm{ad}}_K, T)$.

Since $(F^{\mathrm{ad}})^{-1}(U_i^{\mathrm{ad}}) \subseteq U_i^{\mathrm{ad}}$ for all $i = 0, \dots, N$, we have

$$\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}^{N,\mathrm{ad}}_K = \bigcup_{i=0}^{N} (\varprojlim_{F^{\mathrm{ad}}} U_i^{\mathrm{ad}}).$$

Moreover we have $T(U_i^{\mathrm{ad}}) \subseteq U_i^{\mathrm{ad}}$. It follows that we only need to construct a perfectoid affinoid space U_i^{perf} with an automorphism F_i^{perf} such that the topological dynamical system $(U_i^{\mathrm{perf}}, F_i^{\mathrm{perf}})$ is isomorphic to $(\varprojlim_{F^{\mathrm{ad}}} U_i^{\mathrm{ad}}, T|_{U_i^{\mathrm{ad}}})$ and check that they can be glued together.

Denote $R_i^n := K\langle z_{i,0}^{(n)}, \dots, z_{i,i-1}^{(n)}, z_{i,i+1}^{(n)}, \dots, z_{i,N}^{(n)}\rangle$ for all $i = 0, \dots, N$ and $n \geq 0$. We identify $z_{i,j}^{(0)}$ and $z_{i,j}$ for $i \neq j$. For every $n \geq 0$, we have an embedding $R_i^n \hookrightarrow R_i^{n+1}$ defined by

$$z_{i,j}^{(n)} \mapsto (z_{i,j}^{(n+1)})^q + p' Q_{i,j}(z_{i,0}^{(n+1)}, \dots, z_{i,i-1}^{(n+1)}, z_{i,i+1}^{(n+1)}, \dots, z_{i,N}^{(n+1)})$$

where $Q_{i,j}$ is defined in Section 3. Then we denote by

$$R_i := K\langle z_{i,0}^{(\infty)}, \dots, z_{i,i-1}^{(\infty)}, z_{i,i+1}^{(\infty)}, \dots, z_{i,N}^{(\infty)}\rangle,$$

the completion of $\bigcup_{n=0}^{\infty} R_i^n$. Denote by $\|\cdot\|$ the norm on R_i induced by the norms on R_i^n, $n \geq 0$.

Lemma 3.2. *For every $i = 1, \dots, N$, R_i is a perfectoid K-algebra with*

$$R_i^\circ = K^\circ\langle z_{i,0}^{(\infty)}, \dots, z_{i,i-1}^{(\infty)}, z_{i,i+1}^{(\infty)}, \dots, z_{i,N}^{(\infty)}\rangle.$$

Its tilt is given by $R_i^\flat = K^\flat\langle z_{i,0}^{1/p^\infty}, \dots, z_{i,i-1}^{1/p^\infty}, z_{i,i+1}^{1/p^\infty}, \dots, z_{i,N}^{1/p^\infty}\rangle$.

Proof of Lemma 3.2. Observe that $K^\circ\langle z_{i,0}^{(\infty)}, \dots, z_{i,i-1}^{(\infty)}, z_{i,i+1}^{(\infty)}, \dots, z_{i,N}^{(\infty)}\rangle$ is the completion of $\bigcup_{n=0}^{\infty} (R_i^n)^\circ$. It is easy to check that

$$K^\circ\langle z_{i,0}^{(\infty)}, \dots, z_{i,i-1}^{(\infty)}, z_{i,i+1}^{(\infty)}, \dots, z_{i,N}^{(\infty)}\rangle \subseteq R_i^\circ.$$

For any $f \in R_i$, there exists a sequence $f_n \in R_i^n$ such that $f_n \to f$ as $n \to \infty$. There exists $M \geq 0$, such that for all $m, n \geq M$, $\|f_n - f_m\| \leq 1$. It follows that $f_n - f_M \in (R_i^n)^\circ$ for all $n \geq M$. Then $\widehat{\bigcup_{n=0}^{\infty} (R_i^n)^\circ}$. If $\|f_M\| \leq 1$, we have $f_M \in (R_i^M)^\circ$ and then $\widehat{\bigcup_{n=0}^{\infty} (R_i^n)^\circ}$. If $\|f_M\| > 1$, we have $\|f^n\| = \|f_M^n\| \to \infty$ as $n \to \infty$. Then f is not power bounded. It follows that

$$R_i^\circ \subseteq K^\circ\langle z_{i,0}^{(\infty)}, \dots, z_{i,i-1}^{(\infty)}, z_{i,i+1}^{(\infty)}, \dots, z_{i,N}^{(\infty)}\rangle.$$

It follows that R_i° is open and bounded.

We have $R_i^\circ/\omega = (K^\circ/\omega)\langle z_{i,0}^{(\infty)}, \ldots, z_{i,i-1}^{(\infty)}, z_{i,i+1}^{(\infty)}, \ldots, z_{i,N}^{(\infty)}\rangle$ is the completion of $\bigcup_{n=0}^\infty R_i^n/\omega$. The embedding $R_i^n/\omega \to R_i^n/\omega$ is given by

$$z_{i,j}^{(n)} \bmod \omega \mapsto (z_{i,j}^{(n+1)})^q + p' Q_{i,j}(z_{i,0}^{(n+1)}, \ldots, z_{i,i-1}^{(n+1)}, z_{i,i+1}^{(n+1)}, \ldots, z_{i,N}^{(n+1)}) \bmod \omega = (z_{i,j}^{(n+1)})^q \bmod \omega.$$

It follows that

$$R_i^\circ/\omega = (K^\circ/\omega)\langle z_{i,0}^{1/p^\infty}, \ldots, z_{i,i-1}^{1/p^\infty}, z_{i,i+1}^{1/p^\infty}, \ldots, z_{i,N}^{1/p^\infty}\rangle.$$

Then the Frobenius morphism $\Phi : R_i^\circ/\omega \to R_i^\circ/\omega$ is surjective. It follows that R_i is a perfectoid K-algebra.

By Proposition 2.12 and the categorical equivalence in Theorem 2.10, we have

$$R_i^\flat = K^\flat\langle z_{i,0}^{1/p^\infty}, \ldots, z_{i,i-1}^{1/p^\infty}, z_{i,i+1}^{1/p^\infty}, \ldots, z_{i,N}^{1/p^\infty}\rangle. \qquad \square$$

We define $U_i^{\mathrm{perf}} := \mathrm{Spa}(R_i, R_i^\circ)$ and $F_i^{\mathrm{perf}} : U_i^{\mathrm{perf}} \to U_i^{\mathrm{perf}}$ the map induced by the morphism $R_i \to R_i$ defined by

$$z_{i,j}^{(n)} \to z_{i,j}^{(n-1)} \text{ for all } n \geq 1 \quad \text{and} \quad z_{i,j}^{(0)} \to (z_{i,j}^{(0)})^q + p' Q_{i,j}(z_{i,0}^{(0)}, \ldots, z_{i,i-1}^{(0)}, z_{i,i+1}^{(0)}, \ldots, z_{i,N}^{(0)}).$$

Then we define $(\mathbb{P}_K^N)^{\mathrm{perf}}$ by gluing U_i^{perf} together in the usual way: For any $i \neq j$, $U_i^{\mathrm{perf}} \cap U_j^{\mathrm{perf}} = U(1, z_{i,0}^{(0)}, \ldots, z_{i,i-1}^{(0)}, z_{i,i+1}^{(0)}, \ldots, z_{i,N}^{(0)}; z_{i,j}^{(0)}) \subseteq U_i^{\mathrm{perf}}$. On $U_i^{\mathrm{perf}} \cap U_j^{\mathrm{perf}}$, the transition map $\phi_{i,j}$ is defined to be

$$(\phi_{i,j}^{\mathrm{perf}})^*(z_{j,k}^{(n)}) = z_{i,k}^{(n)}/z_{i,j}^{(n)} \text{ for } k \neq i, j \quad \text{and} \quad (\phi_{i,j}^{\mathrm{perf}})^*(z_{j,i}^{(n)}) = 1/z_{i,j}^{(n)}.$$

It is easy to check that for all $i \neq j$,

$$F_i^{\mathrm{perf}}(U_i^{\mathrm{perf}} \cap U_j^{\mathrm{perf}}) \subseteq U_i^{\mathrm{perf}} \cap U_j^{\mathrm{perf}}$$

and $F_i^{\mathrm{perf}} = F_j^{\mathrm{perf}}$ on $U_i^{\mathrm{perf}} \cap U_j^{\mathrm{perf}}$. Then we define F^{perf} by gluing F_i^{perf} for $i = 0, \ldots, N$.

Then we have the following:

Theorem 3.3. *There exists a natural homeomorphism* $\psi : (\mathbb{P}_K^N)^{\mathrm{perf}} \to \varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}}$ *which makes the following diagram commutative*:

$$\begin{tikzcd} \mathbb{P}_K^{N,\mathrm{perf}} \arrow[r, "\psi"] \arrow[d, "F^{\mathrm{perf}}"'] & \varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}} \arrow[d, "T"] \\ \mathbb{P}_K^{N,\mathrm{perf}} \arrow[r, "\psi"] & \varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}} \end{tikzcd}$$

In other words, the topological dynamical systems $(\mathbb{P}_K^{N,\mathrm{perf}}, F^{\mathrm{perf}})$ *and* $(\varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}}, T)$ *are isomorphic by* ψ.

Moreover a point $x \in \mathbb{P}_K^{N,\mathrm{perf}}$, *with image* $\psi(x) = (x_0, x_1, \ldots)$, *is a* K-*point if and only if* x_n *is a* K-*point for every* $n \geq 0$.

Proof of Theorem 3.3. Denote by $B_i := \bigcup_{n=0}^{\infty} R_i^n$. We have $B_i^\circ = \bigcup_{n=0}^{\infty} R_i^{n\circ}$. Then $\mathrm{Spa}(B, B^\circ)$ is an affinoid space and we have $R_i = \widehat{B_i}$ and $R_i^\circ = \widehat{B_i^\circ}$. By Proposition 2.2, the natural morphism $\mu_i : \mathrm{Spa}(R_i, R_i^\circ) \to \mathrm{Spa}(B_i, B_i^\circ)$ is a homeomorphism.

Denote by $\psi_i^n : U_i^{\mathrm{perf}} \to U_i^{\mathrm{ad}}$ the map induced by the morphism

$$K\langle z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}\rangle \to R_i^n \subseteq B_i \subseteq R_i$$

by sending $z_{i,j} \to z_{i,j}^{(n)}$. It is easy to check that ψ_i^n could be glued to a map $\psi^n : \mathbb{P}_K^{N,\mathrm{perf}} \to \mathbb{P}_K^{N,\mathrm{ad}}$.

Since $F^{\mathrm{ad}} \circ \psi^{n+1} = \psi^n$ for all $n \geq 0$, it induces a map

$$\psi := \varprojlim_{n} \psi^n : \mathbb{P}_K^{N,\mathrm{perf}} \to \varprojlim_{F^{\mathrm{ad}}} \mathbb{P}_K^{N,\mathrm{ad}}.$$

By checking in the affinoid spaces U_i^{perf}, it is easy to check that $T \circ \psi = \psi \circ F^{\mathrm{perf}}$.

So we only need to show that ψ is a homeomorphism. We only need to show it in U_i. Denote by

$$\psi_i := \psi|_{U_i^{\mathrm{perf}}} = \varprojlim_{n} \psi_i^n.$$

Now we define a morphism $\theta_i : \varprojlim_{F^{\mathrm{ad}}} U_i^{\mathrm{ad}} \to \mathrm{Spa}(B_i, B_i^\circ)$ as the following: Let $(x_0, x_1, \dots)$ be a point in $\varprojlim_{F^{\mathrm{ad}}} U_i^{\mathrm{ad}}$. For every $n \geq 0$, we identify $U_i \to \mathrm{Spa}(R_i^n, R_i^{n\circ})$ by $z_{i,j} \to z_{i,j}^{(n)}$. Then x_n defines a valuation on R_i^n with valuation group $\Gamma_n := \{|f(x_n)| : f \in R_i^n\}$. Moreover, for any $\ell \geq n$, and $f \in R_i^n$, we have $|f(x_l)| = |f(x_n)|$. Then we define $\theta_i((x_0, x_1, \dots))$ to be the natural valuation $B_i = \bigcup_{n=0}^{\infty} R_i^n \to \bigcup_{n=0}^{\infty} \Gamma_n$ by gluing all the valuations x_n on R_i^n. Since all the rational subset of $\mathrm{Spa}(B_i, B_i^\circ)$ are defined over some R_i^n, it is easy to check that θ is continuous. It is easy to check that $\psi_i \circ (\mu_i^{-1} \circ \theta_i) = \mathrm{id}$ and $(\mu_i^{-1} \circ \theta_i) \circ \psi_i = \mathrm{id}$. It follows that ψ_i is a homeomorphism.

Let x be a K-point in U_i^{perf} and $\psi_i(x) = (x_0, x_1, \dots)$. For any $n \geq 0$, we have

$$K \subseteq R_i^n / \{|f(x_n)| = 0 : f \in R_i^n\} \subseteq R_i / \{|f(x)| = 0 : f \in R_i\} = K.$$

It follows that x_n is a K-point.

Let x be a point in U_i^{perf} and set $\psi_i(x) = (x_0, x_1, \dots)$. We suppose that all x_n are K-points. Then $m_i^n := \{|f(x_n)| = 0 : f \in R_i^n\}$ is a maximal ideal in R_i^n and $R_i^n / m_i^n = K$. The valuation $R_i^n / m_i^n \to \mathbb{R}$ induced by x_i is the norm on K.

There exists a continuous morphism $\bigcup_{n=0}^{\infty} R_i^n \to K$ obtained by gluing the morphisms $R_i^n \to R_i^n / m_i^n \hookrightarrow K$. We can extend this morphism to a continuous morphism $g : R_i^n = \widehat{\bigcup_{n=0}^{\infty} R_i^n} \to K$. The valuation $f \to |g(f)|$ defines a point $y \in U_i^{\mathrm{perf}}$, which is a K-point. Observe that for all $f \in R_i^n$, $|f(y)| = |f(x_i)|$. Then we have $\psi(y) = (x_0, x_1, \dots) = \psi(x)$. Then $y = x$ and so, x is a K-point. $\square$

For every $i = 0, \dots, N$, the embedding $K\langle z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}\rangle \subseteq R_i$ induces a map $U_i^{\mathrm{perf}} \to U_i^{\mathrm{ad}}$. We define $\pi : \mathbb{P}_K^{N,\mathrm{perf}} \to \mathbb{P}_K^{N,\mathrm{ad}}$ by gluing these maps. It is easy to check that

$$F^{\mathrm{ad}} \circ \pi^\flat = \pi^\flat \circ F^{\mathrm{perf}}.$$

For any point $x \in \mathbb{P}_K^{N,\mathrm{perf}}$ with $\psi(x) = (x_0, x_1, \dots)$, we have $\pi(x) = x_0$.

Passing to the tilt. Denote by $U_i^{\flat,\mathrm{perf}} := \mathrm{Spa}(R_i^\flat, R_i^{\flat\circ})$ and $\Phi_i^{s,\mathrm{perf}} : U_i^{\flat,\mathrm{perf}} \to U_i^{\flat,\mathrm{perf}}$ the s-th power of the Frobenius i.e., the map induced by the morphism

$$R_i^\flat \to R_i^\flat : f \to f^q.$$

We define $(\mathbb{P}^N_{K^\flat})^{\mathrm{perf}}$ by gluing $U_i^{\flat,\mathrm{perf}}$ together in the usual way: For any $i \neq j$, $U_i^{\flat,\mathrm{perf}} \cap U_j^{\flat,\mathrm{perf}} = U(1, z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}; z_{i,j}) \subseteq U_i^{\flat,\mathrm{perf}}$.

On $U_i^{\flat,\mathrm{perf}} \cap U_j^{\flat,\mathrm{perf}}$, the transition map $\phi_{i,j}^\flat$ is defined to be

$$(\phi_{i,j}^{\flat,\mathrm{perf}})^*(z_{j,k}^{1/p^n}) = z_{i,k}^{1/p^n}/z_{i,j}^{1/p^n} \text{ for } k \neq i, j \quad \text{and} \quad (\phi_{i,j}^{\flat,\mathrm{perf}})^*(z_{j,i}^{1/p^n}) = 1/z_{i,j}^{1/p^n}.$$

By reducing modulo ω and the categorical equivalence in Theorem 2.10, we see that $\phi_{i,j}^{\flat,\mathrm{perf}} = (\phi_{i,j}^{\mathrm{perf}})^\flat$ and $\Phi_i^{s,\mathrm{perf}} = (F_i^{\mathrm{perf}})^\flat$. It follows that $(\mathbb{P}^N_{K^\flat})^{\mathrm{perf}} = ((\mathbb{P}^N_K)^{\mathrm{perf}})^\flat$ and we can define $\Phi^{s,\mathrm{perf}}$ by gluing $\Phi_i^{s,\mathrm{perf}}$ together. Moreover, we have $\Phi^{s,\mathrm{perf}} = (F^{\mathrm{perf}})^\flat$. Then we have the following:

Theorem 3.4. *The following diagram is commutative*:

$$\begin{tikzcd} \mathbb{P}_K^{N,\mathrm{perf}} \arrow[r, "\rho"] \arrow[d, "F^{\mathrm{perf}}"'] & \mathbb{P}_{K^\flat}^{N,\mathrm{perf}} \arrow[d, "\Phi^{s,\mathrm{perf}}"] \\ \mathbb{P}_K^{N,\mathrm{perf}} \arrow[r, "\rho"] & \mathbb{P}_{K^\flat}^{N,\mathrm{perf}} \end{tikzcd}$$

In other words, the topological dynamical systems $(\mathbb{P}_K^{N,\mathrm{perf}}, F^{\mathrm{perf}})$ *and* $(\mathbb{P}_{K^\flat}^{N,\mathrm{perf}}, \Phi^s)$ *are isomorphic by* ρ.

For any $i \in \{0, \dots, N\}$, denote by

$$U_i^{\flat,\mathrm{ad}} := \mathrm{Spa}\big(K^\flat\langle z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}\rangle, K^\flat\langle z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}\rangle^\circ\big).$$

As in Section 3, we define $\mathbb{P}_{K^\flat}^{N,\mathrm{ad}}$ by gluing $U_i^{\flat,\mathrm{ad}}$, $i = 0, \dots, N$. Denote by $\phi_i^{s,\mathrm{ad}}$ the s-th power of the Frobenius on U_i i.e., the map induced by the morphism $f \to f^q$ on $K^\flat\langle z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}\rangle$.

By Lemma 3.1, we have a natural embedding $\tau^\flat : \mathbb{P}^N_{K^\flat}(K^\flat) \hookrightarrow \mathbb{P}_{K^\flat}^{N,\mathrm{ad}}$. Then $\tau(\mathbb{P}^N_K(K)) = R(\mathbb{P}_K^{N,\mathrm{ad}})$ and we have the following commutative diagram:

$$\begin{tikzcd} \mathbb{P}^N_{K^\flat}(K^\flat) \arrow[r, "\tau"] \arrow[d, "\Phi^s|_{\mathbb{P}^N_K(K)}"'] & \mathbb{P}_{K^\flat}^{N,\mathrm{ad}} \arrow[d, "\Phi^{s,\mathrm{ad}}"] \\ \mathbb{P}^N_{K^\flat}(K^\flat) \arrow[r, "\tau"] & \mathbb{P}_{K^\flat}^{N,\mathrm{ad}} \end{tikzcd}$$

where Φ^s is the s-th power of the Frobenius on $P^N_{K^\flat}$.

For every $i = 0, \dots, N$, the embedding $K^\flat\langle z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}\rangle \subseteq R_i^\flat$ induces a map $U_i^{\flat,\mathrm{perf}} \to U_i^{\flat,\mathrm{ad}}$. We define $\pi^\flat : \mathbb{P}_{K^\flat}^{N,\mathrm{perf}} \to \mathbb{P}_{K^\flat}^{N,\mathrm{ad}}$ by gluing these maps. It is easy to check that

$$\Phi^{s,\mathrm{ad}} \circ \pi^\flat = \pi^\flat \circ \Phi^{s,\mathrm{perf}}. \tag{1}$$

By [Scholze 2012, Theorem 8.5] $\pi^\flat$ is a homeomorphism. Moreover, we have the following:

Lemma 3.5. *The map $\pi^\flat$ induces a bijection between $R(\mathbb{P}_{K^\flat}^{N,\mathrm{perf}})$ and $R(\mathbb{P}_{K^\flat}^{N,\mathrm{ad}})$.*

Proof of Lemma 3.5. It is clear that if x is a $K^\flat$-point then $\pi^\flat(x)$ is a $K^\flat$-point.

Now we suppose that $\pi^\flat(x)$ is a $K^\flat$-point. We suppose that x is contained in $U_i^{\flat,\mathrm{perf}}$ and then $x_0 := \pi^\flat(x) \in U_i^{\flat,\mathrm{ad}}$. Since x_0 is a $K^\flat$-point, it defines a morphism

$$g_0 : R_i^{\flat,0} := K^\flat\langle z_{i,0}, \dots, z_{i,i-1}, z_{i,i+1}, \dots, z_{i,N}\rangle \to K^\flat.$$

It follows that the Frobenius map $f \to f^p$ on $K^\flat$ is a field automorphism. For any $f \in R_i^{\flat,n} := K^\flat\langle z_{i,0}^{1/p^n}, \dots, z_{i,i-1}^{1/p^n}, z_{i,i+1}^{1/p^n}, \dots, z_{i,N}^{1/p^n}\rangle$, we have $f^{p^n} \in R_i^{\flat,0}$. Then the morphism g_0 extends to a morphism $g_n : R_i^{\flat,n} \to K^\flat$ by sending f to $(g_0(f^{p^n}))^{1/p^n}$. We glue g_n to define a continuous morphism $\bigcup_{n=0}^{\infty} R_i^{\flat,n} \to K^\flat$ and then extend it to a continuous morphism

$$g : R_i^\flat = \widehat{\left(\bigcup_{n=0}^{\infty} R_i^{\flat,n}\right)} \to K^\flat.$$

Then g induces a $K^\flat$-point $y \in U_i^{\flat,\mathrm{perf}}$. Since $\pi^\flat(y) = x_0 = \pi^\flat(x)$, we have $y = x$. Then x is a $K^\flat$-point. □

4. Periodic points

In this section, we denote by $K = \mathbb{C}_p$. Then K is a perfectoid field and $K^\flat$ is the completion of the algebraical closure of $\mathbb{F}_p((t))$. We may suppose that $|p| = |t| = p^{-1}$.

Let $F : \mathbb{P}_K^N \to \mathbb{P}_K^N$ be an endomorphism taking form

$$F : [x_0 : \cdots : x_N] \mapsto [x_0^q + p' P_0(x_0, \dots, x_N) : \cdots : x_N^q + p' P_N(x_0, \dots, x_N)],$$

where $p' \in K^{\circ\circ}$, q is a power of p, and $P_0, \dots, P_N$ are homogeneous polynomials of degree q in $K^\circ[x_0, \dots, x_N]$. The aim of this section is to study the periodic points of F. In particular, we prove Theorem 1.1 and 1.2.

Recall that Per is the set of periodic closed points in $\mathbb{P}_K^N$.

Let V be any irreducible subvariety of $\mathbb{P}_K^N$. Suppose that V is defined by the equations $H_j(x_0, \dots, x_N) = 0$, $j = 1, \dots, m$, where H_j are homogenous polynomials. We may suppose that $\|H_j\| = 1$ for all $j = 1, \dots, m$. For any $i = 0, \dots, N$, denote by

$$V_i^{\mathrm{ad}} := \{x \in U_i^{\mathrm{ad}} : |H_{i,j}(x)| = 0,\ j = 1, \dots, m\},$$

where $H_{i,j} := H(z_{i,0}, \dots, z_{i,i-1}, 1, z_{i,i+1}, \dots, z_{i,N})$. Observe that $\|H_{i,j}\| = 1$.

Set $R(V_i^{\mathrm{ad}}) := R(\mathbb{P}_K^{N,\mathrm{ad}}) \cap V_i^{\mathrm{ad}}$, $V^{\mathrm{ad}} := \bigcup_{i=0}^{N} V_i^{\mathrm{ad}}$ and $R(V^{\mathrm{ad}}) := R(\mathbb{P}_K^{N,\mathrm{ad}}) \cap V^{\mathrm{ad}}$. Then we have $\tau(R(V)) = R(V^{\mathrm{ad}})$.

Observe that for all points $x \in R(U_i^{\mathrm{ad}})$, we have $d(x, V) = \max\{|H_{i,j}(x)|\}$.

Passing to the reduction. Since K is algebraically closed, we have $\mathbb{P}^N_K(K) = \mathbb{P}^N_K(K)$. Denote by $k = \overline{\mathbb{F}_p}$, we have $k = K^\circ/K^{\circ\circ}$. At first, there exists a reduction map

$$\mathrm{red} : \mathbb{P}^N_K(K) \to \mathbb{P}^N_k(k)$$

defined by the following: For any point $x \in \mathbb{P}^N_K(K)$, we may write it as $x = [x_0 : \cdots : x_N]$ where $x_i \in K^\circ$, $i = 0, \ldots, N$ and $\max\{|x_i|, i = 0, \ldots, N\} = 1$. Then we define $\mathrm{red}(x) = [\overline{x_1} : \cdots : \overline{x_N}]$ where $\overline{x_i}$ is the image of x_i in $k = K^\circ/K^{\circ\circ}$. Observe that $\mathrm{red} \circ F = \overline{\Phi}^s \circ \mathrm{red}$, where $\overline{\phi}$ is the Frobenius on $\mathbb{P}^N_k$. For every point $y \in \mathbb{P}^N_k(k)$, there exists $m > 0$ such that $\overline{\Phi}^{sm}(y) = y$. Then we have $D_y := \mathrm{red}^{-1}(y) \simeq (K^{\circ\circ})^N$ is a polydisc fixed by F. Since $F^m|_{D_y}$ is attracting, $D_y \cap \mathrm{Per}$ has exactly one point. It follows that red induces a bijection between Per and $\mathbb{P}^N_k(k)$.

Similarly, we can define the reduction map $\mathrm{red}^\flat : \mathbb{P}^N_{K^\flat}(K^\flat) \to \mathbb{P}^N_k(k)$. This map induces a bijection between $\mathrm{Per}^\flat$ and $\mathbb{P}^N_k(k)$ where $\mathrm{Per}^\flat$ is the set of Φ^s-periodic closed points of $\mathbb{P}^N_{K^\flat}$.

Since k is a subfield of $K^\flat$, there exists an embedding $\eta : \mathbb{P}^N_k(k) \hookrightarrow \mathbb{P}^N_{K^\flat}(K^\flat)$. Observe that the image $\eta(\mathbb{P}^N_k(k))$ is exactly $\mathrm{Per}^\flat$. Moreover we have $\mathrm{red}^\flat \circ \eta = \mathrm{id}$. We may check that the map

$$\phi := \eta \circ \mathrm{red} : \mathrm{Per} \to \mathrm{Per}^\flat$$

is a bijection satisfying $\Phi^s \circ \phi = \phi \circ F$.

Passing to the tilt. Denote by $\mathrm{Per}^{\mathrm{ad}} = \tau(\mathrm{Per})$. It is exactly the set of periodic K-points in $\mathbb{P}^{N,\mathrm{ad}}_K$. For any point $x \in \mathrm{Per}^{\mathrm{ad}}$, denote by $n > 0$ a period of x under F^{ad}. We define a map $\chi : \mathrm{Per}^{\mathrm{ad}} \to \varprojlim_{F^{\mathrm{ad}}} \mathbb{P}^{N,\mathrm{ad}}_K$ by sending x to $(x_0, x_1, \ldots)$ where $x_i = (F^{\mathrm{an}})^{kn-i}(x)$ where $kn \geq i$. We note that $\chi(x)$ does not depend on the choice of n and k. Since $\pi \circ \chi = \mathrm{id}$, χ is injective. We have that $\chi(\mathrm{Per}^{\mathrm{ad}})$ is exactly the set of Per_T, where Per_T is the set of points $(x_0, x_1, \ldots) \in \varprojlim_{F^{\mathrm{ad}}} \mathbb{P}^{N,\mathrm{ad}}_K$ which is periodic under T such that every x_n is a K-point.

Denote by $\mathrm{Per}^{\flat,\mathrm{ad}}$ the set of $K^\flat$-points in $\mathbb{P}^{N,\mathrm{ad}}_{K^\flat}$ which are periodic under $\Phi^{s,\mathrm{ad}}$. By applying Lemma 3.1 over $K^\flat$, there exists a bijection $\tau^\flat : \mathbb{P}^N_{K^\flat}(K^\flat) \to R(\mathbb{P}^{N,\mathrm{ad}}_{K^\flat})$ and we have

$$\tau^\flat \circ \Phi^s = \Phi^{s,\mathrm{ad}} \circ \tau^\flat.$$

It follows that $\tau^\flat$ induces a bijection between $\mathrm{Per}^{\flat,\mathrm{ad}}$ and the set $\mathrm{Per}^\flat$ of Φ^s-periodic points in $R(\mathbb{P}^N_{K^\flat}) = \mathbb{P}^N_{K^\flat}(K^\flat)$.

By Theorems 3.3, 3.4, (1) and Lemma 3.5, the map

$$\iota := \pi^\flat \circ \rho \circ \psi^{-1} \circ \chi : \mathrm{Per}^{\mathrm{ad}} \to \mathrm{Per}^{\flat,\mathrm{ad}}$$

is bijective.

Denote by $\mathrm{Per}^{\mathrm{ad}}_i := \mathrm{Per}^{\mathrm{ad}} \cap U^{\mathrm{ad}}_i$ and $\mathrm{Per}^{\flat,\mathrm{ad}}_i := \mathrm{Per}^{\flat,\mathrm{ad}} \cap U^{\flat,\mathrm{ad}}_i$ for every $i = 0, \ldots, N$. Then we have $\iota(\mathrm{Per}^{\mathrm{ad}}_i) = \mathrm{Per}^{\flat,\mathrm{ad}}_i$.

Observe that for every point $x \in \mathrm{Per}$, we have $\mathrm{red}(x) = \mathrm{red}^\flat \circ \iota \circ \tau(x)$. Then on Per we have

$$\phi = \eta \circ \mathrm{red} = \eta \circ \mathrm{red}^\flat \circ \iota \circ \tau = \iota \circ \tau.$$

Proof of Theorem 1.2. We only need to show this theorem for the periodic points in U_i^{ad} for all $i = 0, \dots, N$. Without the loss of generality, we only need to show that there exists $\delta > 0$ such that for all $x \in \mathrm{Per} \cap U_0^{\mathrm{ad}}$, either $d(x, V) > \delta$ or $x \in V$.

At first, we prove our theorem for hypersurfaces.

Lemma 4.1. *Let $H \in K[x_1, \dots, x_N]$ be a polynomial. Then there exists $\varepsilon > 0$, such that for all $x \in \mathrm{Per} \cap U_0^{\mathrm{ad}}$, either $|H(x)| > \varepsilon$ or $H(x) = 0$.*

By this lemma, for any $H_{0,j}$, $j = 1, \dots, m$, we have $\varepsilon_j > 0$ such that for all $x \in \mathrm{Per} \cap U_0^{\mathrm{ad}}$, either $|H_{0,j}(x)| > \varepsilon_j$ or $H_{0,j}(x) = 0$. Set $\delta := \min_{1 \le j \le m}\{\varepsilon_j\}$. Let x be a point in $\mathrm{Per} \cap U_0^{\mathrm{ad}}$ satisfying $d(x, V) \le \delta$. Then for all $j = 1, \dots, m$, we have $H_{0,j}(x) = 0$. It follows that $x \in V$.

We only need to prove Lemma 4.1. To do this, we need the following lemma:

Lemma 4.2. *Let $E/\overline{\mathbb{F}_p}((t))$ be a finite extension. Then, for some $u \in E$ satisfying $|u| = |t|^{1/[E:\overline{\mathbb{F}_p}((t))]}$, $E = \overline{\mathbb{F}_p}((u))$.*

Proof of Lemma 4.2. Observe that E is a discrete valuation field.

Since $\overline{\mathbb{F}_p}$ is algebraically closed, the extension $E/\overline{\mathbb{F}_p}((t))$ is totally ramified. It follows that $E = \overline{\mathbb{F}_p}((t))(u)$ where the minimal polynomial of u over $\overline{\mathbb{F}_p}((t))$ is an Eisenstein polynomial. It follows that $|u| = |t|^{1/[E:\overline{\mathbb{F}_p}((t))]}$ and uE° is the maximal ideal of E°. For every $f \in E^{\circ}$, f can be written as $\sum_{i \ge 0} a_i u^i$ where $a_i \in \overline{\mathbb{F}_p}$ for all $i \ge 0$. This concludes our proof. $\square$

Lemma 4.3. *For any polynomial $G \in K^{\flat\circ}[x_1, \dots, x_N]$ and $\varepsilon > 0$, there exists a polynomial $G_\varepsilon \in K^{\flat\circ}[x_1, \dots, x_N]$ satisfying $\deg G_\varepsilon \le \deg G$, $\|G - G_\varepsilon\| < \varepsilon$ (resp. $\le \varepsilon$) and G_ε has the form $G_\varepsilon = \sum_{i \ge 0}^{m} u^i g_i$ where $g_i \in \overline{\mathbb{F}_p}[x_1, \dots, x_N]$, $u \in \overline{\mathbb{F}_p((t))}^{\circ}$ with norm $|u| = |t|^{1/[\overline{\mathbb{F}_p}((t))(u):\overline{\mathbb{F}_p}((t))]}$ and $|u|^m \ge \varepsilon$ (resp. $> \varepsilon$).*

Proof of Lemma 4.3. By Lemma 4.2, there exists $u \in \overline{\mathbb{F}_p((t))}^{\circ}$ with norm $|u| = |t|^{1/[\overline{\mathbb{F}_p}((t))(u):\overline{\mathbb{F}_p}((t))]}$ and $H \in \overline{\mathbb{F}_p}((t))(u)[x_1, \dots, x_N]$ such that $\deg H \le \deg G$, $\|G - H\| < \varepsilon$ and H takes form $H = \sum_{i \ge 0}^{\infty} u^i g_i$ where $g_i \in \overline{\mathbb{F}_p}[x_1, \dots, x_N]$. Let m be the largest integer such that $|u|^m \ge \varepsilon$ (resp. $> \varepsilon$). Set $G_\varepsilon = \sum_{i \ge 0}^{m} u^i g_i$ then we conclude our proof. $\square$

Proof of Lemma 4.1. We may suppose that $H \ne 0$ and $\|H\| = 1$.

By Remark 2.18, for any $c \in \mathbb{Z}^+$, there exists $\ell \in \mathbb{N}$ and an element $G_c \in K^{\flat\circ}[x_1^{1/p^\ell}, \dots, x_N^{1/p^\ell}]$ such that for all $x \in U_0^{\mathrm{perf}}$, we have

$$|H \circ \pi(x) - G_c^{\#}(x)| \le |p|^{1/2} \max(|H(x)|, |p|^c) = |p|^{1/2} \max(|G_c^{\#}(x)|, |p|^c),$$

and $G^{p^\ell} \in K^{\flat\circ}[x_1, \dots, x_N]$. By Lemma 4.3, we may suppose that $G_c^{p^\ell} = \sum_{i \ge 0}^{m} u^i g_{c,i}$ where $g_{c,j} \in \overline{\mathbb{F}_p}[x_1, \dots, x_N]$, $u \in \overline{\mathbb{F}_p((t))}^{\circ}$ with norm $|u| = |t|^{1/[\overline{\mathbb{F}_p}((t))(u):\overline{\mathbb{F}_p}((t))]}$ and $|u|^m > |t|^{(c+1/2)p^\ell}$.

Denote by I_c the ideal of $K^{\flat}[x_1, \dots, x_N]$ generated by all $g_{c,i}$.

If $x \in R(U_0^\flat)$ is a point such that for all $g \in I_c$, $g(x) = 0$, then we have

$$\begin{aligned}\left|H(\pi(\rho^{-1}((\pi^b)^{-1}(x))))\right| &= \left|H(\pi(\rho^{-1}((\pi^b)^{-1}(x)))) - G_c^{\#}(\rho^{-1}((\pi^\flat)^{-1}(x)))\right| \\ &\le \max\{|p|^{1/2}|G_c^{\#}(\rho^{-1}((\pi^\flat)^{-1}(x))|, |p|^{c+1/2}\} \\ &= |p|^{c+1/2}.\end{aligned}$$

On the other hand we have the following lemma.

Lemma 4.4. *Let x be a point in $\mathrm{Per} \cap U_0^{\mathrm{ad}}$ satisfying $|H(x)| \le |p|^{c+1/2}$. Then for all $g \in I_c$, we have $g(\phi(x)) = 0$.*

Proof of Lemma 4.4. Observe that $|(H - G_c^{\#})(\chi(x))| \le |p|^{1/2} \max(|H(x)|, |p|^c)$ and $|H(x)| \le |p|^{c+1/2}$. We have $|G_c(\phi(x))| = |G_c^{\#}(\chi(x))| \le |p|^{c+1/2}$. For all $j \ge 0$, we have $g_{c,j}(\phi(x)) \in k$. It follows that either $|g_{c,j}(\phi(x))| = 1$ or $g_{c,j}(\phi(x)) = 0$ for all $j \ge 0$. If $g_{c,j}(\phi(x)) = 0$ for all $j \ge 0$, then for all $g \in I_c$ we have $g(\phi(x)) = 0$. Otherwise, let j_0 be the smallest j satisfying $|g_{c,j}(\phi(x))| = 1$. It follows that $|G_c(\phi(x))| = |u|^{j_0/p^\ell}$. Since $|G_c(\phi(x))| = |G_c^{\#}(\chi(x))| \le |p|^{c+1/2}$, we get a contradiction. □

Set $I := \sum_{c \ge 1} I_c$. Since $K^\flat[x_1, \dots, x_N]$ is Noetherian, there exists $M \in \mathbb{Z}^+$, such that $I = \sum_{c=1}^M I_c$. Set $\varepsilon := |p|^{M+1/2}$. Let x be a point in $\mathrm{Per} \cap U_0^{\mathrm{ad}}$ satisfying $|H(x)| \le \varepsilon = |p|^{M+1/2}$. By Lemma 4.4, for all $g \in I = \sum_{c=1}^M I_c$, we have $|g(\phi(x))| = 0$. It follows that for all $c \ge 1$ and $g \in I_c$, we have $|g(\phi(x))| = 0$. Then we have $|H(x)| \le |p|^{c+1/2}$ for all $c \ge 0$. Let c tend to infinity, we have $H(x) = 0$. We conclude our proof of Lemma 4.1. □

Proof of Theorem 1.1. Suppose that $V \cap \mathrm{Per}$ is Zariski dense in V. We claim the following:

Lemma 4.5. *There exists a Zariski dense subset $S \subseteq V$ with the property that $F^\ell(S) = S$ for some positive integer ℓ.*

Since S is Zariski dense in V and $S = F^\ell(S)$ is Zariski dense in $F^\ell(V)$. It follows that $V = F^\ell(V)$. Then Lemma 4.5 implies Theorem 1.1.

Proof of Lemma 4.5. Since $\bigcup_{i=0}^N \tau^{-1}(\mathrm{Per}_i^{\mathrm{ad}} \cap V_i^{\mathrm{ad}}) = \mathrm{Per}$ is Zariski dense in V, there exists $i = 0, \dots, N$, such that $\tau^{-1}(\mathrm{Per}_i^{\mathrm{ad}} \cap V_i^{\mathrm{ad}})$ is Zariski dense in V. We may suppose that $i = 0$.

Let Z be the Zariski closure of $\phi(\tau^{-1}(\mathrm{Per}_0^{\mathrm{ad}} \cap V_0^{\mathrm{ad}})) \subseteq \mathbb{P}^N_{K^\flat}$. Since $\phi(\tau^{-1}(\mathrm{Per}_0^{\mathrm{ad}} \cap V_0^{\mathrm{ad}}))$ is defined over k and it is Zariski dense in Z, Z is defined over $k = \overline{\mathbb{F}_p}$. Then Z is defined over a finite extension of $\mathbb{F}_p$. It follows that there exists $\ell \ge 1$, such that $\Phi^{sl}(Z) = Z$.

Set $S^{\flat,\mathrm{ad}} := \tau^\flat(Z(K^\flat)) \cap U_0^\flat$. We have $\iota(\mathrm{Per}^{\mathrm{ad}} \cap V_0^{\mathrm{ad}}) \subseteq S^{\flat,\mathrm{ad}} \cap \pi^\flat(\rho(\pi^{-1}(V^{\mathrm{ad}})))$.

We claim the following:

Lemma 4.6. *We have $S^{\flat,\mathrm{ad}} \subseteq \pi^\flat(\rho(\pi^{-1}(V_0^{\mathrm{ad}})))$.*

Remark 4.7. We note that if $\pi^\flat(\rho(\pi^{-1}(V_0^{\mathrm{ad}})))$ is algebraic, our lemma is easy. Since $\phi(\tau^{-1}(\mathrm{Per}_0^{\mathrm{ad}} \cap V_0^{\mathrm{ad}}))$ is Zariski dense in Z, and $\pi^\flat(\rho(\pi^{-1}(V_0^{\mathrm{ad}})))$ is algebraic, we have $S^{\flat,\mathrm{ad}} \subseteq \pi^\flat(\rho(\pi^{-1}(V_0^{\mathrm{ad}})))$.

But in general $\pi^\flat(\rho(\pi^{-1}(V_0^{\mathrm{ad}})))$ is not algebraic since the map ρ is not algebraic. Our proof of Lemma 4.6 is based on Lemma 2.16, which allows us to approximate $\pi^\flat(\rho(\pi^{-1}(V_0^{\mathrm{ad}})))$ by algebraic subvarieties.

By assuming Lemma 4.6, we have $\pi(\rho^{-1}((\pi^\flat)^{-1}(S^{\flat,\mathrm{ad}}))) \subseteq V_0^{\mathrm{ad}}$. Set $S = \tau^{-1}(\pi(\rho^{-1}((\pi^\flat)^{-1}(S^{\flat,\mathrm{ad}}))))$. We have $S \subseteq V$ is a Zariski dense subset of V. Moreover, we have $F^\ell(S) = S$. This concludes the proof of Lemma 4.5. □

Now we only need to prove Lemma 4.6. First, we need the following:

Lemma 4.8. *Let $H \in K^b[z_{0,1}, \dots, z_{0,N}]$ be a polynomial with norm 1. Suppose that for every point $x \in \iota(\mathrm{Per}^{\mathrm{ad}} \cap V_0^{\mathrm{ad}})$, we have $|H(x)| \le 1/p^s$, where $s \in \mathbb{Z}^+$. Then for every point $y \in S^{\flat,\mathrm{ad}}$, we have $|H(y)| \le 1/p^s$.*

Proof of Lemma 4.8. Observe that we have a map

$$R(U_0^{\mathrm{ad}}) = (K^{\flat\circ})^N \to (K^{\flat\circ}/(t^s))^N = \mathbb{A}^N_{K^{\flat\circ}/(t^s)}(K^{\flat\circ}/(t^s))$$

defined by $(x_1, \dots, x_n) \mapsto (\overline{x_1}, \dots, \overline{x_N})$ where $\overline{x_i} = x_i \mod t^s$. Denote by

$$\overline{H} := H \mod t^s.$$

For every point $x \in \iota(\mathrm{Per}^{\mathrm{ad}} \cap V_0^{\mathrm{ad}})$, we have $\overline{H}(\bar{x}) = 0$. Observe that

$$\overline{\iota(\mathrm{Per}^{\mathrm{ad}} \cap V_0^{\mathrm{ad}})} = (\phi(\tau^{-1}(\mathrm{Per}_0^{\mathrm{ad}} \cap V_0^{\mathrm{ad}}))) \times_{\mathrm{Spec}\, k} \mathrm{Spec}(K^{\flat,\circ}/(t^s))$$

is Zariski dense in $Z \times_{\mathrm{Spec}\, k} \mathrm{Spec}(K^{\flat,\circ}/(t^s))$. It follows that

$$Z \times_{\mathrm{Spec}\, k} \mathrm{Spec}(K^{\flat,\circ}/(t^s)) \subseteq \{\overline{H} = 0\}.$$

Then we have

$$\overline{S^{\flat,\mathrm{ad}}} = Z(k) \times_{\mathrm{Spec}\, k} \mathrm{Spec}(K^{\flat,\circ}/(t^s)) \subseteq Z \times_{\mathrm{Spec}\, k} \mathrm{Spec}(K^{\flat,\circ}/(t^s)) \subseteq \{\overline{H} = 0\}.$$

It follows that for every $x \in S^{\flat,\mathrm{ad}}$, we have $H(x) = 0 \mod t^s$. Then we have $|H(x)| \le 1/p^s$, for all $x \in S^{\flat,\mathrm{ad}}$. □

Proof of Lemma 4.6. Now we apply Lemma 2.16 to $H_{0,j} \in K\langle z_{0,1}, \dots, z_{0,N}\rangle \subseteq R_0^{\mathrm{perf}}$ for every $j = 1, \dots, m$. For any $s \ge 2$ there exists $h_s \in R_0^{\flat,\mathrm{perf}}$ such that for all $x \in U_0^{\mathrm{perf}}$, we have

$$|H_{0,j}(x) - h_s^\#(x)| \le |t|^{1/2} \max(|H_{0,j}(x)|, |t|^s) = |t|^{1/2} \max\{|h_s^\#(x)|, |t|^s\} < 1. \tag{2}$$

It follows that $\|h_s\| = \|H_{0,j}\| = 1$.

For every point $x^\flat \in (\pi^\flat)^{-1}(\iota(\mathrm{Per}^{\mathrm{ad}} \cap V_0^{\mathrm{ad}}))$, we have

$$x := \rho^{-1}(x^\flat) \in \pi^{-1}(\mathrm{Per}^{\mathrm{ad}} \cap V_0^{\mathrm{ad}}).$$

Then we have $H_{0,j}(x) = 0$. By (2) we have

$$|h_s(x^\flat)| \le |t|^{s+1/2} = |t|^{1/2} \max\{|h_s(x^\flat)|, |t|^s\} = 1/p^{s+1/2}.$$

Since $h_s \in R_0^{\flat,\mathrm{perf}} = K^\flat\langle z_{0,1}^{1/p^\infty}, \dots, z_{0,N}^{1/p^\infty}\rangle$, there are $r \geq 0$ and a function

$$g_s \in K^\flat[z_{0,1}^{1/p^r}, \dots, z_{0,N}^{1/p^r}]$$

such that $\|h_s - g_s\| < 1/p^s$. It follows that $g_s^{p^r} \in K^\flat[z_{0,1}, \dots, z_{0,N}]$ and

$$\|h_s^{p^r} - g_s^{p^r}\| \leq |p|^{sp^r}.$$

Then for every point $x^\flat \in (\pi^\flat)^{-1}(\iota(\mathrm{Per}^{\mathrm{ad}} \cap V_0^{\mathrm{ad}}))$, we have

$$|g_s^{p^r}(\pi^\flat(x^\flat))| = |g_s^{p^r}(x^\flat)| = |h_s^{p^r}(x^\flat) + (g_s^{p^r}(x^\flat) - h_s^{p^r}(x^\flat))| \leq |p|^{sp^r}.$$

By Lemma 4.8, for all $y \in S^{\flat,\mathrm{ad}}$, we have $|g_s^{p^r}(y)| \leq |p|^{sp^r}$. Then we have $|h_s((\pi^\flat)^{-1}(y)| \leq 1/p^s$ and

$$|h_s^\#(\rho^{-1}((\pi^\flat)^{-1}(y))| = |h_s((\pi^\flat)^{-1}(y)| \leq 1/p^s$$

for all $y \in S^{\flat,\mathrm{ad}}$.

By (2), we have

$$|H_{0,j}(x) - h_s^\#(x)| \leq |t|^{1/2}\max\{|h_s^\#(x)|, |t|^s\} = 1/p^{s+1/2}$$

for all $x \in \rho^{-1}((\pi^\flat)^{-1}(S^{\flat,\mathrm{ad}}))$. It follows that for all $x \in \rho^{-1}((\pi^\flat)^{-1}(S^{\flat,\mathrm{ad}}))$, we have $|H_{0,j}(x)| \leq 1/p^s$. Let $s \to \infty$, we have $|H_{0,j}(x)| = 0$ for all $x \in \rho^{-1}((\pi^\flat)^{-1}(S^{\flat,\mathrm{ad}}))$. Since $|H_{0,j}(x)| = |H_{0,j}(\pi(x))|$, we have $|H_{0,j}(y)| = 0$ for all $j = 1, \dots, m$ and $y \in \pi(\rho^{-1}((\pi^\flat)^{-1}(S^{\flat,\mathrm{ad}})))$. It follows that $\pi(\rho^{-1}((\pi^\flat)^{-1}(S^{\flat,\mathrm{ad}}))) \subseteq V_0^{\mathrm{ad}}$. Then we have $S^{\flat,\mathrm{ad}} \subseteq \pi^\flat(\rho(\pi^{-1}(V_0^{\mathrm{ad}})))$. □

Scanlon's proof of Theorem 1.1. In this section, we discuss Scanlon's proof of Theorem 1.1. In this proof, we don't need the perfectoid spaces.

Let V be a subvariety of $\mathbb{P}^N$ such that $\mathrm{Per} \cap V$ is Zariski dense in V. We want to show that V is periodic.

We first treat the case where F is defined over $\overline{\mathbb{Q}_p}^\circ$. Since all points in Per are defined over $\overline{\mathbb{Q}_p}$ and $\mathrm{Per} \cap V$ is Zariski dense in V, V is defined over $\overline{\mathbb{Q}_p}$. There exists a finite extension K_p of $\mathbb{Q}_p$ such that F is defined over K_p i.e., F takes form

$$F : [x_0 : \dots : x_N] \mapsto [x_0^q + p' P_0(x_0, \dots, x_n) : \dots : x_N^q + p' P_N(x_0, \dots, x_N)],$$

where $p' \in K_p^{\circ\circ}$, q is a power of p, $P_0, \dots, P_N$ are homogeneous polynomials of degree $q = p^s$ in $K_p^\circ[x_0, \dots, x_N]$. After replacing F by a suitable iterate, we may assume that the residue field $\tilde{K} := K^\circ/K^{\circ\circ}$ is fixed by the q-power Frobenius.

By the structure of the absolute Galois group of K_p, there exists an element $\sigma \in \mathrm{Gal}(\overline{K_p}/K_p)$ which lifts the q-power Frobenius. Then we have the following lemma:

Lemma 4.9 [Medvedev and Scanlon 2014]. *We have* $\mathrm{Per} = \{x \in \mathbb{P}^N(\overline{\mathbb{Q}_p}) : F(x) = \sigma(x)\}$.

Proof of Lemma 4.9. Recall that the reduction map

$$\mathrm{red} : \mathbb{P}^N(\overline{\mathbb{Q}_p}) \to \mathbb{P}^N(\overline{\mathbb{F}_p})$$

gives a bijection between Per and $\mathbb{P}^N(\overline{\mathbb{F}_p})$.

Let x be any point in Per. We have that $F(x) \in \mathrm{Per}$ and $\mathrm{red}(F(x)) = \mathrm{red}(x)^q$. On the other hand, we have that $\sigma(x) \in \mathrm{Per}$ and $\mathrm{red}(\sigma(x)) = \mathrm{red}(x)^q$. Then we have $F(x) = \sigma(x)$.

Let x be any point in $\mathbb{P}^N(\overline{\mathbb{Q}_p})$ satisfying $F(x) = \sigma(x)$. Since x is defined over a finite extension of K_p, there exists $n \geq 1$ such that $\sigma^n(x) = x$. It follows that

$$F^n(x) = F^{n-1}(\sigma(x)) = \sigma(F^{n-1}(x)) = \cdots = \sigma^n(x) = x.$$

Then x is periodic. □

Observe that $\sigma(V)$ is a subvariety of $\mathbb{P}^N$. Then we have

$$\sigma(V \cap \mathrm{Per}) = F(V \cap \mathrm{Per}) \subseteq \sigma(V) \cap F(V).$$

Since $V \cap \mathrm{Per}$ is Zariski dense in V, we have $\sigma(V) = F(V)$. Since V is defined over a finite extension of $\mathbb{Q}_p$, there exists $n \geq 1$ such that $\sigma^n(V) = V$. It follows that

$$F^n(V) = F^{n-1}(\sigma(V)) = \sigma(F^{n-1}(V)) = \cdots = \sigma^n(V) = V.$$

Then V is periodic.

Now we treat the general case.

There exists a subring $R \subseteq \mathbb{C}_p^\circ$ which is finitely generated over $\mathbb{Z}$ such that F is defined over R. Let $m := R \cap \mathbb{C}_p^{\circ\circ}$ be a maximal ideal of R. By Lemma A.3, there exists $\sigma \in \mathrm{Gal}(\mathbb{C}_p/\mathbb{Q})$ such that $\sigma(R) \subseteq \overline{\mathbb{Q}_p}^\circ \subseteq \mathbb{C}_p^\circ$ and $\sigma(m) = \overline{\mathbb{Q}_p}^{\circ\circ} \cap R$.

Denote by F^σ the Galois conjugate of F by σ i.e., F^σ is obtained by changing every coefficient of F by its image under σ. Since $F^\sigma \mod \mathbb{C}_p^{\circ\circ} = F \mod \mathbb{C}_p^{\circ\circ}$, F^σ is a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. Moreover it is defined over $\overline{\mathbb{Q}_p}^\circ$.

Since $V \cap \mathrm{Per}$ is Zariski dense in V, $\sigma(V) \cap \sigma(\mathrm{Per})$ is Zariski dense in $\sigma(V)$. Moreover $\sigma(\mathrm{Per})$ is exactly the set of periodic points of F^σ. Then the previous argument shows that $\sigma(V)$ is periodic under F^σ. It follows that V is periodic under F.

5. Coherent backward orbits

In this section, we let $K = \mathbb{C}_p$. Then K is a perfectoid field and $K^\flat$ is the completion of the algebraic closure of $\mathbb{F}_p((t))$. We may suppose that $|p| = |t| = p^{-1}$. Let $k = \overline{\mathbb{F}_p}$ which is a subfield of $K^\flat$.

Let $F : \mathbb{P}^N_K \to \mathbb{P}^N_K$ be an endomorphism taking form

$$F : [x_0 : \cdots : x_N] \mapsto [x_0^q + p' P_0(x_0, \ldots, x_N) : \cdots : x_N^q + p' P_N(x_0, \ldots, x_N)],$$

where $p' \in K^{\circ\circ}$, q is a power of p, and $P_0, \ldots, P_N$ are homogeneous polynomials of degree q in $K^\circ[x_0, \ldots, x_N]$.

The aim of this section is to prove Theorems 1.7 and 1.8.

Without loss of generality, we may suppose that $b_0 \in R(U_0^{\mathrm{ad}})$. It follows that $b_i \in R(U_0^{\mathrm{ad}})$ for all $i \geq 0$. Set $w := \pi^\flat \circ \rho \circ \psi^{-1}((b_0, b_1, \ldots)) \in \mathbb{P}^N_{K^\flat}(K_b)$. Then $w \in R(U_0^{\flat,\mathrm{ad}}) := \{[1 : x_1 : \cdots : x_N] : |x_i| \leq 1\} \subseteq \mathbb{P}^N_{K^\flat}(K_b)$. It follows that $w^{1/q^n} \subseteq R(U_0^{\flat,\mathrm{ad}})$ for all $n \geq 0$.

If $\{b_i\}_{i\geq 0}$ is infinite, we may suppose that $b_1 \neq b_0$ and then the b_i, for $i \geq 0$, are all different. Let Z be the reduced subvariety of $U_0^\flat := \operatorname{Spec} K^\flat[x_1, \dots, x_N]$, whose support is the union of all positive dimensional irreducible components of the Zariski closure of $\{w^{1/q^{n_i}}\}_{i\geq 0}$.

There exists $A \geq 0$, such that Z is the Zariski closure of $\{w^{1/q^{n_i}}\}_{i\geq A}$ in $U_0^\flat$. Moreover, for all $n \geq A$, Z is the Zariski closure of $\{w^{1/q^{n_i}}\}_{i\geq n}$ in $U_0^\flat$.

Denote by $I(Z)$ the ideal in $K^\flat[x_1, \dots, x_N]$ which defines Z.

For every polynomial $f = \sum_I a_I x^I \in K^\flat[x_1, \dots .x_N]$ and $i \in \mathbb{Z}$, we denote by $f^{\sigma^i} := \sum_I a_I^{q^i} x^I$. Observe that $f(y^{1/q^i}) = (f^{\sigma^i}(y))^{1/q^i}$ for all $i \geq 0$ and $y \in R(U_0^{\flat,\mathrm{ad}})$.

Then we have the following lemma:

Lemma 5.1. *Let $f \in k[x_1, \dots, x_N]$ be a polynomial defined over k. If there exists $c \in (0, 1)$ and $B \geq A$, such that for all $i \geq B$, $|f(w^{1/q^{n_i}})| \leq c$, then $f \in I(Z)$.*

Proof of Lemma 5.1. There exists $L \geq 1$ such that f is defined over $\mathbb{F}_{q^L}$. Then we have $f^{\sigma^{nL}} = f$ for all $n \geq 0$. For $t = 0, \dots, L-1$, set $T_t := \{i \geq B \mid n_i = t \mod L\}$.

For all $t = 0, \dots, L-1$ satisfying $\#T_t = \infty$, we have

$$|f(w^{1/q^t})|^{1/q^{n_i-t}} = |f^{\sigma^{n_i-t}}(w^{1/q^t})|^{1/q^{n_i-t}} = |f(w^{1/q^{n_i}})| \leq c,$$

for all $i \in T_t$. It follows that $|f(w^{1/q^t})| \leq c^{q^{n_i-t}}$ for all $i \in T_t$. Since T_t is infinite, n_i can be arbitrary large. Then we have $|f(w^{1/q^t})| = 0$ for all $i \in T_t$. It follows that

$$|f(w^{1/q^{n_i}})| = |f^{\sigma^{n_i-t}}(w^{1/q^t})|^{1/q^{n_i-t}} = |f(w^{1/q^t})|^{1/q^{n_i-t}} = 0$$

for all $i \in T_t$. Set

$$T' := \bigsqcup_{\substack{0\leq t\leq L-1\\ \#T_t=\infty}} T_t.$$

It follows that $f(w^{1/q^{n_i}}) = 0$ for all $i \in T'$. Since $\{i \geq A\} \setminus T'$ is finite, $\{w^{1/q^{n_i}}\}_{i\in T'}$ is Zariski dense in Z. Then $f \in I(Z)$. □

Lemma 5.2. *We have that Z is defined over k. In particular, there exists $r \geq 1$ such that $\Phi^{sr}(Z) = Z$ and $\{w^{1/q^i}\}_{i\in\mathbb{Z}} \subseteq \bigcup_{i=0}^{r-1} \Phi^{si}(Z)$.*

Proof of Lemma 5.2. We only need to show that $I(Z)$ is generated by finitely many polynomials in $k[x_1, \dots, x_N] \subseteq K^\flat[x_1, \dots, x_N]$. In fact, if $I(Z) = (g_1, \dots, g_l)$ and $g_i \in k[x_1, \dots, x_N]$ for all $i = 1, \dots, \ell$, then there exists $r \geq 1$ such that all the coefficients of $g_i, i = 1, \dots, \ell$, are defined over $\mathbb{F}_{q^r}$. Then we have $\Phi^{sr}(Z) = Z$. Moreover, there exists $j \geq 0$, such that $w^{1/p^j} \in Z$. It follows that $\{w^{1/q^i}\}_{i\in\mathbb{Z}} \subseteq \bigcup_{i\in\mathbb{Z}} \Phi^{si}(Z) = \bigcup_{i=0}^{r-1} \Phi^{si}(Z)$.

Write $I(Z) = (f_1, \dots, f_m)$ where $m \geq 1$ and $f_i \in K^\flat[x_1, \dots, x_N]$ for all $i = 1, \dots, m$. Denote by $d := \max_{0\leq i\leq m}\{\deg(f_i)\}$.

By Lemma 4.3, for all $i = 1, \dots, m$, there exists a sequence of polynomial $\{f_{i,n}\}_{n\geq 1}$ such that $\|f_i - f_{i,n}\| \leq |t^n|$ and taking form $f_{i,n} = \sum_{j=0}^{m_{i,n}} u_{i,n}^j f_{i,n,j}$ where $f_{i,n,j} \in \overline{\mathbb{F}_p}[x_1, \dots, x_N]$ of degree at most d, $u_{i,n} \in \overline{\mathbb{F}_p((t))}^{\circ}$ with norm $|u_{i,n}| = |t|^{1/[\overline{\mathbb{F}_p}((t))(u_{i,n}):\overline{\mathbb{F}_p}((t))]}$ and $|u_{i,n}|^{m_{i,n}} > |t|^n$.

We claim that $f_{i,j,n} \in I(Z)$ for all $j = 0, \dots, m_{i,n}$.

We prove that claim by induction on j. For $j = 0$, we have

$$|f_{i,0,n}(w^{1/q^{n_l}})| = \left| f_{i,n}(w^{1/q^{n_l}}) - \sum_{j\geq 1} u_{i,n}^j f_{i,j,n}(w^{1/q^{n_l}}) \right| \leq \max\{|t^n|, |u_{i,n}|\} < 1$$

for all $\ell \geq A$. By Lemma 5.2, we have $f_{i,0} \in I(Z)$.

If $j \geq 1$ and $f_{i,0,n}, \dots, f_{i,j-1,n} \in I(Z)$, then

$$\begin{aligned} |f_{i,j,n}(w^{1/q^{n_l}})| &= \left| u_{i,n}^{-j}(f_{i,n}(w^{1/q^{n_l}}) - \sum_{0\leq t'\leq j-1} u_{i,n}^{t'} f_{i,t',n}(w^{1/q^{n_l}})) - \sum_{t'\geq j+1} u_{i,n}^{t'-j} f_{i,j,n}(w^{1/q^{n_l}}) \right| \\ &\leq \max\left\{ \frac{|t|^n}{|u_{i,n}|^j}, \left| \sum_{t'\geq j+1} u_{i,n}^{t'-j} f_{i,j,n}(w^{1/q^{n_l}}) \right| \right\} \\ &\leq \max\left\{ \frac{|t|^n}{|u_{i,n}|^j}, |u_{i,n}| \right\} \\ &< 1 \end{aligned}$$

for all $\ell \geq A$. By Lemma 5.2, we have $f_{i,j,n} \in I(Z)$. This concludes the proof of the claim. It follows that $f_{i,n} \in I$.

Set $I_d := \{f \in I \mid \deg(f) \leq d\}$. Then I_d is a finite-dimensional $K^\flat$-vector space. For all $n \geq 0$ and $j = 0, \dots, m_{i,n}$, denote by $I_{i,j,n}$ the K_b-vector space spanned by $f_{i,0,0} \dots, f_{i,j,0}, \dots, f_{i,0,n} \dots, f_{i,j,n}$. Then $\bigcup_{n\geq 0, j=0,\dots,m_{i,n}} I_{i,j,n}$ is a subspace of I_d. Since $\dim I_d$ is finite, $\bigcup_{n\geq 0, j=0,\dots,m_{i,n}} I_{i,j,n}$ is closed. Observe that f_i is contained in the closure of $\bigcup_{n\geq 0, j=0,\dots,m_{i,n}} I_{i,j,n}$, we have $f_i \in \bigcup_{n\geq 0, j=0,\dots,m_{i,n}} I_{i,j,n}$. There exists $l_i \geq 0$, such that $f_i \in I_{i,m_{i,l_i},l_i}$. It follows that $I = (f_1, \dots, f_m) \subseteq \sum_{1\leq i\leq m}(I_{i,m_{i,l_i},l_i}) \subseteq I$. Then we have $I = (f_{i,j,n})_{1\leq i\leq m, 0\leq n\leq l_i, 0\leq j\leq m_{l_i}}$ and $f_{i,j} \in k[x_1, \dots, x_N]$ for all i, j. □

Proof of Theorem 1.8. Let V be a subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$ such that there exists a subsequence $\{b_{n_i}\}_{i\geq 0}$ such that $|d(b_{n_i}, V)| \to 0$ when $i \to \infty$. We need to show that $b_{n_i} \in V$ for i large enough and there exists $r \geq 0$, such that $\{b_i\}_{i\geq 0} \subseteq \bigcup_{i=0}^{r-1} F^i(V)$.

If $\{b_i\}_{i\geq 0}$ is finite, Theorem 1.8 is trivial. So we suppose that $\{b_i\}_{i\geq 0}$ is infinite.

Let $I(V)$ denote the ideal in $K[x_1, \dots, x_N]$ which defines $V \cap U_1$. Then for any point in $R(U_0^{\mathrm{ad}})$, we have $d(y, V) = \max\{|H(y)| : H \in I(V) \text{ and } \|H\| = 1\}$.

Let Z denote the union of all positive dimensional irreducible components of the Zariski closure of $\{w^{1/q^{n_i}}\}_{i\geq 0}$. There exists an $A \geq 0$, such that Z is that Zariski closure of $\{w^{1/q^{n_i}}\}_{i\geq A}$ in $U_0^\flat :=$ Spec $K^\flat[x_1, \dots, x_N]$. Moreover, for all $n \geq A$, Z is the Zariski closure of $\{w^{1/q^{n_i}}\}_{i\geq n}$ in $U_0^\flat$.

Let $I(Z)$ denote the ideal in $K^\flat[x_1, \dots, x_N]$ which defines Z. Let H be a polynomial in $I(V)$.

Lemma 5.3. *For any point $x \in Z \cap R(U_0^{\flat,\mathrm{ad}})$, we have $H(\pi(\rho^{-1}((\pi^\flat)^{-1}(x)))) = 0$.*

Proof of Lemma 5.3. By Remark 2.18, for any $c \in \mathbb{Z}^+$, there exists $\ell \in \mathbb{N}$ and an element $G_c \in K^{\flat\circ}[x_1^{1/p^\ell}, \ldots, x_N^{1/p^\ell}]$ such that for all $x \in U_0^{\mathrm{perf}}$, we have

$$|H \circ \pi(x) - G_c^\#(x)| \le |p|^{1/2} \max(|H(x)|, |p|^c) = |p|^{1/2} \max(|G_c^\#(x)|, |p|^c),$$

and $G_c^{p^\ell} \in K^{\flat\circ}[x_1, \ldots, x_N]$. By Lemma 4.3, we may suppose that $G_c^{p^\ell} = \sum_{i \ge 0}^m u^i g_i$ where $g_i \in \overline{\mathbb{F}_p}[x_1, \ldots, x_N]$, $u \in \overline{\mathbb{F}_p((t))}^\circ$ with norm $|u| = |t|^{1/[\overline{\mathbb{F}_p}((t))(u):\overline{\mathbb{F}_p}((t))]}$ and $|u|^m > |t|^{(c+1/2)p^\ell}$.

There exists $A_1 \ge 0$, such that $|H(b_{n_i})| \le |p|^{c+1}$, for all $i \ge A_1$.

For all $i \ge A_1$, we have

$$|G_c(w^{1/q^{n_i}})| \le \max\{|H(b_{n_i})|, |H(b_{n_i}) - G_c^\#(\rho^{-1}(w^{1/q^{n_i}}))|\} \le |p|^{c+1/2}.$$

Then we have

$$|G_c(w^{1/q^{n_i}})^{p^\ell}| \le |t|^{p^\ell(c+1/2)}.$$

We claim that for all $j = 0, \ldots, m$, we have $g_j \in I(Z)$.

We prove this claim by induction on j. Suppose that for all $0 \le t' < j \le m$, we have $g_{t'} \in I(Z)$. For all $i \ge \max\{A, A_1\}$, we have

$$\left| u^j g_j(w^{1/q^{n_i}}) + \sum_{t' \ge j+1} u^{t'} g_{t'}(w^{1/q^{n_i}}) \right| = |G_c(w^{1/q^{n_i}})^{p^\ell}| \le |t|^{p^\ell(c+1/2)}.$$

It follows that $|g_j(w^{1/q^{n_i}})| \le \max\{|t|^{p^\ell(c+1/2)}/|u|^j, |u|\} < 1$ for all $i \ge \max\{A, A_1\}$. Then Lemma 5.1 implies that $g_j \in I(Z)$ for $j = 0, \ldots, m$. This proves the claim.

Then for any $x \in Z \cap R(U_0^{\flat,\mathrm{ad}})$, we have

$$\begin{aligned} |H(\pi(\rho^{-1}((\pi^b)^{-1}(x))))| &= |H(\pi(\rho^{-1}((\pi^b)^{-1}(x)))) - G_c^\#(\rho^{-1}((\pi^\flat)^{-1}(x)))| \\ &\le \max\{|p|^{1/2}|G_c^\#(\rho^{-1}((\pi^\flat)^{-1}(x))|, |p|^{c+1/2}\} \\ &= |p|^{c+1/2}. \end{aligned}$$

Let c tend to infinity, then we have $|H(\pi(\rho^{-1}((\pi^b)^{-1}(x))))| = 0$. We complete the proof of our lemma. □

This lemma shows that $S := \pi(\rho^{-1}((\pi^b)^{-1}(Z \cap R(U_0^{\flat,\mathrm{ad}})))) \subseteq V$. Then $b_{n_i} \in V$ for $i \ge A$. By Lemma 5.2, there exists $r \ge 1$ such that $\Phi^{sr}(Z) = Z$ and $\{w^{1/p^i}\}_{i \in \mathbb{Z}} \subseteq \bigcup_{i=0}^r \Phi^{si}(Z)$. It follows that $\{b_i\}_{i \ge 0} \subseteq \bigcup_{i=0}^{r-1} F^i(S) \subseteq \bigcup_{i=0}^{r-1} F^i(V)$.

Proof of Corollary 1.9. Let V be a subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$ of positive dimension. We need to show there exists $c > 0$ such that for all $i \ge 0$ either $b_i \in V$ or $d(b_i, V) > c$.

Otherwise, there exists a subsequence $\{b_{n_i}\}_{i \ge 0} \subseteq \{b_i\}_{i \ge 0} \setminus V$ such that $d(b_{n_i}, V)$ tends to 0. By Theorem 1.8, we have $b_{n_i} \in V$ for sufficiently large i, which is a contradiction. □

Proof of Theorem 1.7. Let V be a positive subvariety of $\mathbb{P}^N_{\mathbb{C}_p}$ such that $\{b_i\}_{i\geq 0}\cap V$ is Zariski dense in V. Let $\{n_1 < n_2 < \cdots\}$ be the set of $n\geq 0$ such that $b_n\in V$. We need to show that V is periodic under F.

If $\{b_i\}_{i\geq 0}$ is finite, then all points in $\{b_i\}_{i\geq 0}$ are periodic. Moreover V is a union of finitely many periodic points. So V is periodic.

Now we may suppose that $\{b_i\}_{i\geq 0}$ is infinite. Denote by $I(V)$ the ideal in $K[x_1,\dots,x_N]$ which defines $V\cap U_1$. Let H be a polynomial $I(V)$. By Lemma 5.3, for any point $x\in Z\cap R(U_0^{\flat,\mathrm{ad}})$, we have $H(\pi(\rho^{-1}((\pi^b)^{-1}(x))))=0$.

It follows that $S:=\pi(\rho^{-1}((\pi^b)^{-1}(Z\cap R(U_0^{\flat,\mathrm{ad}}))))\subseteq V$. Since $b_{n_i}\in S$ for all $i\geq A$, S is Zariski dense in V. Since $\Phi^{rs}(Z\cap R(U_0^{\flat,\mathrm{ad}}))=Z\cap R(U_0^{\flat,\mathrm{ad}})$, we have $F^r(S)=S$. It follows that $F^r(V)=V$. This concludes the proof. □

Appendix

Let X be any projective variety over $\mathbb{C}_p$ and $F: X\to X$ be an endomorphism. Let $\mathfrak{X}\to\operatorname{Spec}\mathbb{C}_p^\circ$ be a finitely presented projective scheme which is flat over $\operatorname{Spec}\mathbb{C}_p^\circ$ whose generic fiber is X and $\mathcal{L}$ an ample line bundle on $\mathfrak{X}$. If there exists an endomorphism $\tilde{F}$ of $\mathfrak{X}$ over $\mathbb{C}_p^\circ$ such that $\tilde{F}^*\mathcal{L}=\mathcal{L}^{\otimes q}$ where $q=p^s$, $s\geq 1$, the restriction of $\tilde{F}$ on the generic fiber is F and the restriction $\overline{F}$ of $\tilde{F}$ on the special fiber $\overline{X}$ is a power of the Frobenius, then we say that F *is a polarized lift of Frobenius on* X *with respect to* $(\mathfrak{X},\tilde{F},\mathcal{L})$. In particular, a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$ in the previous sections is a lift of Frobenius on X with respect to a pair $(\mathbb{P}^N_{\mathbb{C}_p^\circ},\tilde{F},O_{\mathbb{P}^N_{\mathbb{C}_p^\circ}}(1))$.

Now assume that F is a polarized lift of Frobenius on X with respect to the pair $(\mathfrak{X},\tilde{F},\mathcal{L})$ and we identify X with the generic fiber of $\mathfrak{X}$.

In this appendix, we show that under a technical condition, the dynamical system (X,F) can be embedded in a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$ (with respect to some $(\mathbb{P}^N_{\mathbb{C}_p^\circ},\tilde{F},O_{\mathbb{P}^N_{\mathbb{C}_p^\circ}}(1))$).

Theorem A.1. *Assume that $\mathfrak{X}$ and $\tilde{F}$ are defined over $\overline{\mathbb{Q}_p}^\circ\subseteq\mathbb{C}_p^\circ$. Then there exists $N\geq 1$, a lift of Frobenius G on $\mathbb{P}^N_{\mathbb{C}_p}$ and an embedding $\tau: X\hookrightarrow\mathbb{P}^N_{\mathbb{C}_p}$ such that $\tau\circ F^l=G\circ\tau$ for some $l\geq 1$.*

This theorem can be viewed as a version of [Fakhruddin 2003, Proposition 2.1] for the lifts of Frobenius.

As an application, it implies the dynamical Manin–Mumford Conjecture and Conjecture 1.5, for any polarized lift of Frobenius on X with respect to some $(\mathfrak{X},\tilde{F},\mathcal{L})$.

Corollary A.2. *Let V be any positive dimensional irreducible subvariety of X. Denote by Per_F the set of periodic closed points in X. Let $\{b_i\}_{i\geq 0}$ be a sequence of closed points in X satisfying $f(b_i)=b_{i-1}$ for all $i\geq 1$. Then we have that*:

(i) *If $V\cap\mathrm{Per}_F$ is Zariski dense in V, then V is periodic.*

(ii) *If the $\{b_i\}_{i\geq 0}\cap V$ is Zariski dense in V, then V is periodic under F.*

Proof of Corollary A.2. There exists subring $R \subseteq \mathbb{C}_p^\circ$ which is finitely generated over $\mathbb{Z}$ such that $\mathfrak{X}$, $\tilde{F}$ and $\mathcal{L}$ are defined over R, i.e., there exists a projective scheme $\mathfrak{X}_R$ over Spec R with an endomorphism $\tilde{F}_R$ and an ample line bundle $\mathcal{L}_R$ such that $\mathfrak{X} = \mathfrak{X}_R \otimes_R \mathbb{C}_p^\circ$, $\mathcal{L} = \mathcal{L}_R \otimes_R \mathbb{C}_p^\circ$, $\tilde{F} = \tilde{F}_R \otimes_R \mathbb{C}_p^\circ$ and $\tilde{F}_R^*\mathcal{L}_R = \mathcal{L}_R^{\otimes q}$.

If $R \subseteq \overline{\mathbb{Q}_p}^\circ$, Theorem A.1 reduces it to the case where $X = \mathbb{P}^N_{\mathbb{C}_p}$ and F is a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. We conclude the proof by applying Theorems 1.1 and 1.8.

Now we assume that $R \not\subseteq \overline{\mathbb{Q}_p}^\circ$. Set $m := R \cap \mathbb{C}_p^{\circ\circ}$. It is a maximal ideal of R.

Lemma A.3. *Let R be a subring of $\mathbb{C}_p^\circ$ which is finitely generated over $\mathbb{Z}$. Let $m := R \cap \mathbb{C}_p^{\circ\circ}$ be a maximal ideal of R. Then there exists $\sigma \in \mathrm{Gal}(\mathbb{C}_p/\mathbb{Q})$ such that $\sigma(R) \subseteq \overline{\mathbb{Q}_p}^\circ \subseteq \mathbb{C}_p^\circ$ and $\sigma(m) = \overline{\mathbb{Q}_p}^{\circ\circ} \cap R$.*

Now consider $\mathfrak{X}^\sigma := \mathfrak{X}_R \otimes_R^\sigma \mathbb{C}_p^\circ$, $\mathcal{L}^\sigma := \mathcal{L}_R \otimes_R^\sigma \mathbb{C}_p^\circ$, $\tilde{F}^\sigma := \tilde{F} \otimes_R^\sigma \mathbb{C}_p^\circ$, $X^\sigma := X_R \otimes_R^\sigma \mathbb{C}_p$ and $F^\sigma := X_R \otimes_R^\sigma \mathbb{C}_p$. In the tensor product $\bullet \otimes_R^\sigma \mathbb{C}_p$, we use the embedding $\sigma|_R$. We note that if we view $\mathbb{C}_p$ as an abstract field, (X, F) and (X^σ, F^σ) are Galois conjugate. Since the statements of (i) and (ii) are purely algebraic, we only need to show it for (X^σ, F^σ). Observe that the special fiber $\overline{X}^\sigma$ of $\mathfrak{X}^\sigma$ is

$$\overline{X}^\sigma = \mathfrak{X}_R \otimes_R^\sigma (\mathbb{C}_p^\circ/\mathbb{C}_p^{\circ\circ}) = \mathfrak{X}_R \otimes_R (R/m) \otimes_{R/m}^{\bar{\sigma}} (\mathbb{C}_p^\circ/\mathbb{C}_p^{\circ\circ}) = \mathfrak{X}_R \otimes_R^{\bar{\sigma}} (\mathbb{C}_p^\circ/\mathbb{C}_p^{\circ\circ}) \simeq \overline{X}.$$

Moreover the restriction of F^σ on $\overline{X}^\sigma$ is exactly $\overline{F}$ under this identification. So F^σ is some power of Frobenius and F is a lift of the Frobenius with respect to $(\mathfrak{X}^\sigma, \tilde{F}^\sigma, \mathcal{L}^\sigma)$. Since $(\mathfrak{X}^\sigma, \tilde{F}^\sigma, \mathcal{L}^\sigma)$ is defined over $\sigma(R) \subseteq \overline{\mathbb{Q}_p}^\circ$, Theorem A.1 reduces it to the case where $X = \mathbb{P}^N_{\mathbb{C}_p}$ and F is a lift of Frobenius on $\mathbb{P}^N_{\mathbb{C}_p}$. We conclude the proof by applying Theorems 1.1 and 1.8. $\square$

Proof of Lemma A.3. Since $\mathbb{C}_p$ is algebraically closed, any embedding $R \hookrightarrow \mathbb{C}_p$ extends to an automorphism in $\mathrm{Gal}(\mathbb{C}_p/\mathbb{Q})$. We only need to find an embedding $\sigma : R \hookrightarrow \mathbb{C}_p^\circ \subseteq \mathbb{C}_p$ satisfying $\sigma(m) \subseteq \mathbb{C}_p^{\circ\circ}$. Indeed since $\sigma^{-1}(\mathbb{C}_p^{\circ\circ} \cap \sigma(R))$ is a maximal ideal of R which contains m, we have $\sigma(m) = \sigma(R) \cap \mathbb{C}_p^{\circ\circ}$.

Let $t_1, \ldots, t_l \in R$ be a set of generators of R over $\mathbb{Z}$. Let $u_1, \ldots, u_s$ be a set of generators of m. Set $Y :=$ Spec R and $Y_{\mathbb{C}_p} :=$ Spec $R \otimes_{\mathbb{Z}} \mathbb{C}_p$. We endow $Y_{\mathbb{C}_p}(\mathbb{C}_p)$ with the p-adic topology induced by the topology on $\mathbb{C}_p$. An element $f \in R$ can be viewed as an analytic function on $Y_{\mathbb{C}_p}(\mathbb{C}_p)$.

Denote by $i : R \hookrightarrow \mathbb{C}_p$ the inclusion. It defines a point $o \in Y_{\mathbb{C}_p}(\mathbb{C}_p)$. Set $U := \{x \in Y_{\mathbb{C}_p}(\mathbb{C}_p) : |t_i| \leq 1, i = 1, \ldots, l, \text{ and } |u_i| < 1, i = 1, \ldots, s\}$. Then U is an open neighborhood of o.

For any nonzero element P of R, denote by V_P the subscheme of $Y_{\mathbb{C}_p}$ defined by $\{P = 0\}$. Since the set of nonzero prime ideals is countable, and $Y_{\mathbb{C}_p}(\mathbb{C}_p)$ has a complete metric, $Y_{\mathbb{C}_p}(\mathbb{C}_p) \setminus \left(\bigcup_{R\setminus\{0\}} V_P\right)$ is dense in $Y_{\mathbb{C}_p}(\mathbb{C}_p)$. Then there exists a point $y \in U \setminus \left(\bigcup_{R\setminus\{0\}} V_P\right)$. It defines a morphism $\sigma : R \to \mathbb{C}_p$ by $f \mapsto f(y)$. Because $y \in U$, we have $\sigma(t_i) \in \mathbb{C}^\circ$, $i = 1, \ldots, l$, and $\sigma(u_i) \in \mathbb{C}_p^{\circ\circ}$, $i = 1, \ldots, s$. It follows that $\sigma(R) \subseteq \mathbb{C}_p^\circ$ and $\sigma(m) \subseteq \mathbb{C}_p^{\circ\circ}$. Since $y \notin \left(\bigcup_{R\setminus\{0\}} V_P\right)$, $\sigma : R \to \mathbb{C}_p$ is an embedding. This concludes the proof. $\square$

Proof of Theorem A.1. In this section, we assume that $\mathfrak{X}$, $\tilde{F}$ and $\mathcal{L}$ are defined over $\overline{\mathbb{Q}_p}^\circ \subseteq \mathbb{C}_p^\circ$. Since $\mathfrak{X}$ is finitely presented, there exists a finite extension K of $\mathbb{Q}_p$ such that $\mathfrak{X}$, $\tilde{F}$ and $\mathcal{L}$ are defined over K°. We note that $R := K^\circ$ is a discrete valuation ring. Set $m := K^{\circ\circ}$ the maximal ideal of R and π a generator of m.

There exists a flat and geometrically irreducible projective scheme $\mathfrak{X}_R$ over Spec R an ample line bundle $\mathcal{L}_R$ and an endomorphism $\tilde{F}_R$ such that $\mathfrak{X} = \mathfrak{X}_R \otimes_R \mathbb{C}_p^\circ$ $\mathcal{L} = \mathcal{L}_R \otimes_R \mathbb{C}_p^\circ$ and $\tilde{F} = \tilde{F}_R \otimes_R \mathbb{C}_p^\circ$. We may assume that $\tilde{F}_R^*\mathcal{L}_R = \mathcal{L}_R^{\otimes q}$. We denote by X_s the special fiber of $\mathfrak{X}_R$ and F_s the restriction of $\tilde{F}_R$ on X_s. Let X_K be the generic fiber of $\mathfrak{X}_R$ and F_K the restriction of $\tilde{F}_R$ on X_K. Write $L_K := \mathcal{L}_R|_{X_K}$ and $L_s := \mathcal{L}_R|_{X_s}$. Since $\mathcal{L}_R$ is ample, after replacing $\mathcal{L}_R$ by a suitable power, we may assume that $\mathcal{L}$ is very ample and the morphisms

$$\Psi_R := H^0(X_R, \mathcal{L}_R)^{\otimes q} \to H^0(X_R, \mathcal{L}_R^{\otimes q}) \quad \text{and} \quad \Psi_s := H^0(X_s, \mathcal{L}_s)^{\otimes q} \to H^0(X_s, \mathcal{L}_R^{\otimes q})$$

are surjective. Moreover, we may assume that

$$H^i(X_R, \mathcal{L}_R) = H^i(X_R, \mathcal{L}_R^{\otimes q}) = 0$$

for all $i \geq 1$. It follows that the natural morphisms

$$r_1 : H^0(X_R, \mathcal{L}_R) \otimes_R R/m \to H^0(X_s, \mathcal{L}_s) \quad \text{and} \quad r_q : H^0(X_R, \mathcal{L}_R^{\otimes q}) \otimes_R R/m \to H^0(X_s, \mathcal{L}_s^{\otimes q})$$

are isomorphisms.

Since $\tilde{F}_R^*\mathcal{L}_R = \mathcal{L}_R^{\otimes q}$, it induces morphisms

$$\tilde{F}_R^* : H^0(X_R, \mathcal{L}_R) \to H^0(X_R, \mathcal{L}_R^{\otimes q}) \quad \text{and} \quad \tilde{F}_s^* : H^0(X_s, \mathcal{L}_s) \to H^0(X_s, \mathcal{L}_s^{\otimes q}).$$

We have $r_q \circ \tilde{F}_R^* = F_s^* \circ r_1$.

Since R is a discrete valuation ring, and $H^0(X_R, \mathcal{L}_R)$ has no torsions, $H^0(X_R, \mathcal{L}_R)$ is a free R-module. Let $s_0, \dots s_N$ a basis of $H^0(X_R, \mathcal{L}_R)$. We note that

$$r_q(\tilde{F}^*(s_i)) = r_1(s_i)^q$$

for $i = 0, \dots, N$. It follows that

$$\tilde{F}^*(s_i) - \Psi_R(s_i^q) \in m H^0(X_R, \mathcal{L}_R^{\otimes q}) = \pi H^0(X_R, \mathcal{L}_R^{\otimes q})$$

for $i = 0, \dots, N$. In other words, there exists $g_i \in H^0(X_R, \mathcal{L}_R^{\otimes q})$ such that

$$\tilde{F}^*(s_i) = s_i^q + \pi g_i, i = 1, \dots, N.$$

Since Ψ_R is surjective, there exists $G_i \in R[x_0, \dots, x_N]$ homogenous of degree q such that $g_i = G_i(s_0, \dots, s_N), i = 1, \dots, N$. It follows that

$$\tilde{F}^*(s_i) = s_i^q + \pi G_i(s_0, \dots, s_N), \quad i = 1, \dots, N.$$

Let $G_K : \mathbb{P}_K^N \to \mathbb{P}_K^N$ be the morphism

$$[x_0, \dots, x_N] \mapsto [x_0^q + \pi G_0(x_0, \dots, x_N) : \dots : x_N^q + \pi G_N(x_0, \dots, x_N)].$$

Set $G := G_K \otimes_K \mathbb{C}_p : \mathbb{P}_{\mathbb{C}_p}^N \to \mathbb{P}_{\mathbb{C}_p}^N$. It is a lift of Frobenius on $\mathbb{P}_{\mathbb{C}_p}^N$. Let $\tau_K : X \to \mathbb{P}_K^N$ be the morphism

$$x \mapsto [s_0(x) : \dots : s_N(x)].$$

Since L_K is very ample, τ is an embedding. We may check that $G_K \circ \tau_K = \tau_K \circ F_K$. We conclude the proof by setting $\tau := \tau_K \otimes_K \mathbb{C}_p$. $\square$

Acknowledgement

I would like to thank Charles Favre, Serge Cantat, Stéphane Lamy, Jean Gillibert, Dragos Ghioca, Thomas Tucker and Peter Scholze for useful discussions. I thank Fabien Mehdi Pazuki and Umberto Zannier for their comments on the first version of this paper. I thank Thomas Scanlon, who told me about his new proof of Theorem 1.1 and let me know of the Tate–Voloch conjecture. We thank the referees for numerous insightful remarks. I especially thank Shou-Wu Zhang, who introduced me to the theory of perfectoid spaces and posed Question 1.3.

References

[Baker and Hsia 2005] M. H. Baker and L.-C. Hsia, "Canonical heights, transfinite diameters, and polynomial dynamics", *J. Reine Angew. Math.* **585** (2005), 61–92. MR Zbl

[Bell et al. 2010] J. P. Bell, D. Ghioca, and T. J. Tucker, "The dynamical Mordell–Lang problem for étale maps", *Amer. J. Math.* **132**:6 (2010), 1655–1675. MR Zbl

[Bell et al. 2016] J. P. Bell, D. Ghioca, and T. J. Tucker, *The dynamical Mordell–Lang conjecture*, Mathematical Surveys and Monographs **210**, American Mathematical Society, Providence, RI, 2016. MR Zbl

[Bosch et al. 1984] S. Bosch, U. Güntzer, and R. Remmert, *Non-Archimedean analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **261**, Springer, 1984. MR Zbl

[Buium 1996a] A. Buium, "An approximation property for Teichmüller points", *Math. Res. Lett.* **3**:4 (1996), 453–457. MR Zbl

[Buium 1996b] A. Buium, "Geometry of p-jets", *Duke Math. J.* **82**:2 (1996), 349–367. MR Zbl

[Dujardin and Favre 2017] R. Dujardin and C. Favre, "The dynamical Manin–Mumford problem for plane polynomial automorphisms", *J. Eur. Math. Soc.* (*JEMS*) **19**:11 (2017), 3421–3465. MR Zbl

[Fakhruddin 2003] N. Fakhruddin, "Questions on self maps of algebraic varieties", *J. Ramanujan Math. Soc.* **18**:2 (2003), 109–122. MR Zbl

[Fakhruddin 2014] N. Fakhruddin, "The algebraic dynamics of generic endomorphisms of $\mathbb{P}^n$", *Algebra Number Theory* **8**:3 (2014), 587–608. MR Zbl

[Faltings 1994] G. Faltings, "The general case of S. Lang's conjecture", pp. 175–182 in *Barsotti Symposium in Algebraic Geometry* (Abano Terme, 1991), edited by V. Cristante and W. Messing, Perspect. Math. **15**, Academic Press, San Diego, CA, 1994. MR Zbl

[Fontaine and Wintenberger 1979] J.-M. Fontaine and J.-P. Wintenberger, "Extensions algébrique et corps des normes des extensions APF des corps locaux", *C. R. Acad. Sci. Paris Sér. A-B* **288**:8 (1979), A441–A444. MR Zbl

[Gabber and Ramero 2003] O. Gabber and L. Ramero, *Almost ring theory*, Lecture Notes in Mathematics **1800**, Springer, 2003. MR Zbl

[Ghioca and Tucker 2009] D. Ghioca and T. J. Tucker, "Periodic points, linearizing maps, and the dynamical Mordell–Lang problem", *J. Number Theory* **129**:6 (2009), 1392–1403. MR Zbl

[Ghioca and Tucker 2010] D. Ghioca and T. J. Tucker, "Proof of a dynamical Bogomolov conjecture for lines under polynomial actions", *Proc. Amer. Math. Soc.* **138**:3 (2010), 937–942. MR Zbl

[Ghioca et al. 2011] D. Ghioca, T. J. Tucker, and S. Zhang, "Towards a dynamical Manin–Mumford conjecture", *Int. Math. Res. Not.* **2011**:22 (2011), 5109–5122. MR Zbl

[Ghioca et al. 2015] D. Ghioca, K. D. Nguyen, and H. Ye, "The Dynamical Manin–Mumford Conjecture and the Dynamical Bogomolov Conjecture for split rational maps", 2015. arXiv

[Ghioca et al. 2018] D. Ghioca, K. D. Nguyen, and H. Ye, "The dynamical Manin–Mumford conjecture and the dynamical Bogomolov conjecture for endomorphisms of $(\mathbb{P}^1)^n$", *Compos. Math.* **154**:7 (2018), 1441–1472. MR

[Hrushovski 2001] E. Hrushovski, "The Manin–Mumford conjecture and the model theory of difference fields", *Ann. Pure Appl. Logic* **112**:1 (2001), 43–115. MR Zbl

[Huber 1993] R. Huber, "Continuous valuations", *Math. Z.* **212**:3 (1993), 455–477. MR Zbl

[Huber 1994] R. Huber, "A generalization of formal schemes and rigid analytic varieties", *Math. Z.* **217**:4 (1994), 513–551. MR Zbl

[Huber 1996] R. Huber, *Étale cohomology of rigid analytic varieties and adic spaces*, Friedr. Vieweg & Sohn, Braunschweig, 1996. MR Zbl

[Medvedev and Scanlon 2014] A. Medvedev and T. Scanlon, "Invariant varieties for polynomial dynamical systems", *Ann. of Math.* (2) **179**:1 (2014), 81–177. MR Zbl

[Pazuki 2010] F. Pazuki, "Zhang's conjecture and squares of abelian surfaces", *C. R. Math. Acad. Sci. Paris* **348**:9-10 (2010), 483–486. MR Zbl

[Pazuki 2013] F. Pazuki, "Polarized morphisms between abelian varieties", *Int. J. Number Theory* **9**:2 (2013), 405–411. MR Zbl

[Pink and Roessler 2002] R. Pink and D. Roessler, "On Hrushovski's proof of the Manin–Mumford conjecture", pp. 539–546 in *Proceedings of the International Congress of Mathematicians, Vol. I* (Beijing, 2002), edited by T. Li, Higher Ed. Press, Beijing, 2002. MR Zbl

[Raynaud 1983a] M. Raynaud, "Courbes sur une variété abélienne et points de torsion", *Invent. Math.* **71**:1 (1983), 207–233. MR Zbl

[Raynaud 1983b] M. Raynaud, "Sous-variétés d'une variété abélienne et points de torsion", pp. 327–352 in *Arithmetic and geometry, Vol. I*, edited by M. Artin and J. Tate, Progr. Math. **35**, Birkhäuser, Boston, 1983. MR Zbl

[Scanlon 1999] T. Scanlon, "The conjecture of Tate and Voloch on p-adic proximity to torsion", *Internat. Math. Res. Notices* 17 (1999), 909–914. MR Zbl

[Scholze 2012] P. Scholze, "Perfectoid spaces", *Publ. Math. Inst. Hautes Études Sci.* **116** (2012), 245–313. MR Zbl

[Scholze 2014] P. Scholze, "Perfectoid spaces and their applications", pp. 461–486 in *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. II*, edited by S. Y. Jang et al., Kyung Moon Sa, Seoul, 2014. MR Zbl

[Tate and Voloch 1996] J. Tate and J. F. Voloch, "Linear forms in p-adic roots of unity", *Internat. Math. Res. Notices* 12 (1996), 589–601. MR Zbl

[Ullmo 1998] E. Ullmo, "Positivité et discrétion des points algébriques des courbes", *Ann. of Math.* (2) **147**:1 (1998), 167–179. MR Zbl

[Vojta 1996] P. Vojta, "Integral points on subvarieties of semiabelian varieties. I", *Invent. Math.* **126**:1 (1996), 133–181. MR Zbl

[Xie 2017] J. Xie, "The dynamical Mordell–Lang conjecture for polynomial endomorphisms of the affine plane", pp. vi+110 in *Journées de Géométrie Algébrique d'Orsay*, Astérisque **394**, Société Mathématique de France, Paris, 2017. MR Zbl

[Yuan and Zhang 2017] X. Yuan and S.-W. Zhang, "The arithmetic Hodge index theorem for adelic line bundles", *Math. Ann.* **367**:3-4 (2017), 1123–1171. MR Zbl

[Zhang 1995] S. Zhang, "Small points and adelic metrics", *J. Algebraic Geom.* **4**:2 (1995), 281–300. MR Zbl

[Zhang 1998] S.-W. Zhang, "Equidistribution of small points on abelian varieties", *Ann. of Math.* (2) **147**:1 (1998), 159–165. MR Zbl

Communicated by Shou-Wu Zhang
Received 2017-10-02 Revised 2018-06-15 Accepted 2018-07-17

junyi.xie@univ-rennes1.fr *Institut de Recherche Mathématique de Rennes, CNRS - Université de Rennes 1, Bâtiment 22-23 du campus de Beaulieu, Rennes, France*

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1749

A dynamical variant of the Pink–Zilber conjecture

Dragos Ghioca and Khoa Dang Nguyen

Let $f_1, \dots, f_n \in \overline{\mathbb{Q}}[x]$ be polynomials of degree $d > 1$ such that no f_i is conjugate to x^d or to $\pm C_d(x)$, where $C_d(x)$ is the Chebyshev polynomial of degree d. We let φ be their coordinatewise action on $\mathbb{A}^n$, i.e., $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ is given by $(x_1, \dots, x_n) \mapsto (f_1(x_1), \dots, f_n(x_n))$. We prove a dynamical version of the Pink–Zilber conjecture for subvarieties V of $\mathbb{A}^n$ with respect to the dynamical system $(\mathbb{A}^n, \varphi)$, if $\min\{\dim(V), \operatorname{codim}(V) - 1\} \le 1$.

1. Introduction

1A. *Notation.* As always in dynamics, we write φ^m for the m-th compositional power of the self-map φ for any $m \in \mathbb{N}_0$ (where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$); also, φ^0 is the identity map. The orbit of some point α under φ is denoted by $\mathcal{O}_\varphi(\alpha)$ and it consists of all $\varphi^m(\alpha)$ for $m \in \mathbb{N}_0$. For a subvariety $Y \subset \mathbb{A}^n$ under the action of an endomorphism φ, we say that Y is periodic if there exists a positive integer m such that $Y = \varphi^m(Y)$; similarly, we say that Y is preperiodic under the action of φ if there exists $m \in \mathbb{N}_0$ such that $\varphi^m(Y)$ is periodic.

For every $d \ge 2$, the Chebyshev polynomial of degree d, denoted $C_d(x)$, is the polynomial of degree d satisfying the functional equation $C_d\left(x + \frac{1}{x}\right) = x^d + \frac{1}{x^d}$. Following [Medvedev and Scanlon 2014], a *disintegrated polynomial* is a polynomial of degree $d \ge 2$ that is not linearly conjugate to x^d or $\pm C_d(x)$.

1B. *Our results.* In [Ghioca and Nguyen 2016], a dynamical version of the bounded height conjecture (see [Bombieri et al. 2007] for the formulation of this classical conjecture in the context of algebraic tori) was proven for endomorphisms of $\mathbb{A}^n$ given by coordinatewise action of disintegrated polynomials. The results of [Ghioca and Nguyen 2016] suggest the following variant of the Pink–Zilber conjecture in a dynamical setting; see [Bombieri et al. 1999; Zilber 2002; Pink $\ge$ 2018] for the statement of this conjecture in the classical setting of algebraic tori, or more generally, of semiabelian schemes.

Conjecture 1.1. *Let $f_1, \dots, f_n \in \overline{\mathbb{Q}}[x]$ be disintegrated polynomials of degree $d \ge 2$. We let φ be their coordinatewise action on $\mathbb{A}^n$, i.e., $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ is given by $(x_1, \dots, x_n) \mapsto (f_1(x_1), \dots, f_n(x_n))$. For each positive integer $s \le n$, we let $\mathrm{Per}^{[s]}$ be the union of all irreducible periodic subvarieties of $\mathbb{A}^n$ of codimension s; similarly, we let $\mathrm{Prep}^{[s]}$ be the union of all irreducible preperiodic subvarieties of $\mathbb{A}^n$ of codimension s. Let $X \subset \mathbb{A}^n$ be an irreducible subvariety of dimension m.*

The research of the first author was partially supported by an NSERC Discovery grant.
MSC2010: primary 11G50; secondary 11G35, 14G25.
Keywords: dynamical Pink–Zilber conjecture, heights.

(1) *If $X \cap \mathrm{Per}^{[m+1]}$ is Zariski dense in X, then X is contained in a proper, irreducible subvariety of $\mathbb{A}^n$, which is periodic under the action of φ.*

(2) *If $X \cap \mathrm{Prep}^{[m+1]}$ is Zariski dense in X, then X is contained in a proper, irreducible subvariety of $\mathbb{A}^n$, which is preperiodic under the action of φ.*

We prove the following result in support of Conjecture 1.1.

Theorem 1.2. *Let $f_1, \dots, f_n \in \overline{\mathbb{Q}}[x]$ be disintegrated polynomials of degree $d > 1$ and let φ be their coordinatewise action on $\mathbb{A}^n$, i.e., $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ is given by $\varphi(x_1, \dots, x_n) = (f_1(x_1), \dots, f_n(x_n))$. Let $X \subset \mathbb{A}^n$ be an irreducible subvariety defined over $\overline{\mathbb{Q}}$ such that $\min\{\dim(X), \mathrm{codim}(X) - 1\} \le 1$. If $X \cap \mathrm{Per}^{[\dim(X)+1]}$ is Zariski dense, then X is contained in a proper, irreducible, periodic subvariety of $\mathbb{A}^n$.*

Therefore Theorem 1.2 provides a proof for Conjecture 1.1(1) in the following 3 *nontrivial* cases:

(I) X is a hypersurface (see Theorem 3.1, which proves the more general result that any irreducible subvariety of $\mathbb{A}^n$ containing a Zariski dense set of periodic points must be itself periodic).

(II) X is a curve (see Theorem 4.1).

(III) $X \subset \mathbb{A}^n$ has codimension 2 (see Theorem 5.1 which proves a generalization of this statement by showing that for *any* irreducible subvariety $X \subset \mathbb{A}^n$ of codimension at least equal to 2, we have that if $X \cap \mathrm{Per}^{[n-1]}$ is Zariski dense in X, then X must be contained in a proper, periodic, irreducible subvariety of $\mathbb{A}^n$).

Clearly, if X is a point (i.e., $\dim(X) = 0$), or if $X = \mathbb{A}^n$ (i.e., $\mathrm{codim}(X) = 0$, in which case $\mathrm{Per}^{[n+1]}$ is void since there is no periodic subvariety of codimension larger than n), Conjecture 1.1 holds.

Remark 1.3. In particular, we observe that Theorem 1.2 proves completely Conjecture 1.1(1) for all subvarieties of $\mathbb{A}^n$ if $n \le 4$.

1C. *The dynamical Pink–Zilber conjecture.*

We discuss next some subtleties involved in Conjecture 1.1.

Remark 1.4. It is natural to wonder whether Conjecture 1.1 could be formulated alternatively by asking if $X \cap \left(\bigcup_{i>\dim(X)} \mathrm{Per}^{[i]}\right)$ or, respectively, $X \cap \left(\bigcup_{i>\dim(X)} \mathrm{Prep}^{[i]}\right)$ is Zariski dense in X. However, since each periodic subvariety of codimension $m+1$ is contained in a periodic subvariety of codimension m (see Section 2), this alternative formulation would reduce to our Conjecture 1.1.

It makes sense to restrict Conjecture 1.1 to polynomials which are not conjugate to monomials or Chebyshev polynomials since otherwise we would encounter the classical Pink–Zilber conjecture (see [Zannier 2012] for a comprehensive discussion). Also, we note that if X is contained in a proper, irreducible (pre)periodic subvariety Y of $\mathbb{A}^n$, then (simply, by geometric considerations of counting the dimensions) X intersects nontrivially each (pre)periodic subvariety of relative codimension in Y equal to $\dim(X)$, and thus, X has a Zariski dense intersection with $\mathrm{Per}^{[\dim(X)+1]}$ and $\mathrm{Prep}^{[\dim(X)+1]}$, respectively; this scenario is identical to the classical case when a subvariety $X \subset \mathbb{G}_m^n$ contained in a proper algebraic subtorus would have a Zariski dense intersection with the union of all subtori in $\mathbb{G}_m^n$ of codimension equal to $\dim(X)+1$.

We also note that the two parts of Conjecture 1.1 are *independent*, neither one implying the other one. Furthermore, it is likely that the methods one would need to employ in proving the above two conjectures might differ slightly. For example, we would expect that some of the p-adic techniques developed for attacking the dynamical Mordell–Lang conjecture (for more details, see [Bell et al. 2016, Chapter 4]) could prove useful in treating Conjecture 1.1(1) in full generality. On the other hand, in attacking Conjecture 1.1(2), one might need to develop generalizations of the arguments employed in [Ghioca et al. 2018]. Also, Conjecture 1.1(2) is particularly challenging since one lacks a corresponding dynamical bounded height conjecture for preperiodic subvarieties, in the spirit of the one proven in [Ghioca and Nguyen 2016] (which is valid only for periodic subvarieties). Attempting to prove a variant of the bounded height conjecture for preperiodic subvarieties of $\mathbb{A}^n$ leads to subtle diophantine questions similar to the ones encountered in [DeMarco et al. 2017].

It is important to observe that if we did not impose the condition that the polynomials have the same degree, then there would be simple counterexamples, similar to those of a naive formulation of the dynamical Manin–Mumford conjecture (see Section 1D) which does not require the polarizability of the given endomorphism. Indeed, if $f \in \overline{\mathbb{Q}}[x]$ has degree $d \geq 2$, then its graph $y = f(x)$ is a (rational) plane curve containing infinitely many points which are periodic under the coordinatewise action of

$$(x, y) \mapsto (f(x), f^2(y));$$

however, this graph is *not periodic* under the action of (f, f^2).

It is *difficult* to extend any of our results to dynamical systems given by the coordinatewise action of rational functions due to the absence of Medvedev and Scanlon's [2014] classification of periodic subvarieties in that case (see also [Ghioca and Nguyen 2016]). Also, it is difficult to extend Theorem 1.2 to subvarieties $X \subset \mathbb{A}^n$ of dimension either larger than 1, or codimension larger than 2; see the following Example, which can be generalized to any subvariety of $\mathbb{A}^n$ of dimension in the range $\{2, \dots, n-3\}$.

Example 1.5. Let $f \in \overline{\mathbb{Q}}[x]$ be a polynomial of degree $d \geq 2$ and let φ be its coordinatewise action on $\mathbb{A}^6$. Let $X \subset \mathbb{A}^6$ be a surface which projects to a nonpreperiodic point on each of the first 3 coordinates, i.e., $X = \zeta \times X_1$, where $\zeta \in \mathbb{A}^3(\overline{\mathbb{Q}})$ and $X_1 \subset \mathbb{A}^3$ is a surface defined over $\overline{\mathbb{Q}}$. We also assume X_1 is not a periodic surface, while ζ is not contained in a proper periodic subvariety of $\mathbb{A}^3$; this last assumption can be achieved (see Section 2) by assuming the coordinates of $\zeta := (\zeta_1, \zeta_2, \zeta_3)$ belong to different orbits under f, i.e., there are no $i, j \in \{1, 2, 3\}$ and no $m, n \in \mathbb{N}$ such that $f^m(\zeta_i) = f^n(\zeta_j)$. Then X is not contained in a proper periodic subvariety of $\mathbb{A}^6$ and therefore, Conjecture 1.1 predicts that $X \cap \mathrm{Per}^{[3]}$ is not Zariski dense in X. In particular, this would yield that

$$X_1 \cap \big(\mathcal{O}_f(\zeta_1) \times \mathcal{O}_f(\zeta_2) \times \mathcal{O}_f(\zeta_3)\big) \tag{1.6}$$

is not Zariski dense in X_1. However, understanding the intersection from (1.6) is equivalent to solving a stronger form of the dynamical Mordell–Lang conjecture for hypersurfaces in $\mathbb{A}^3$ and at the present moment, this problem seems very difficult; for a comprehensive discussion about the dynamical Mordell–Lang conjecture, see [Bell et al. 2016].

As shown by Bombieri, Masser and Zannier [Bombieri et al. 1999; 2006], even the classical Pink–Zilber conjecture in the context of algebraic tori is very difficult and initially only the case of curves was established; for more details, see the beautiful book [Zannier 2012]. In the dynamical context, the fact that we do not even know the validity of the dynamical Mordell–Lang conjecture makes Conjecture 1.1 particularly challenging.

It is also natural to formulate Conjecture 1.1 for polynomials with complex coefficients. The difficulty in extending our present results to this more general setting lies in a couple of points. First, there is no easy specialization argument which would yield a similar result to the one from Theorem 1.2 for polynomials with complex coefficients by simply using the conclusion of Theorem 1.2. Secondly, it is essential for our strategy of proof (for more details, see Section 1E) to use the dynamical Bogomolov conjecture (proven in [Ghioca et al. 2018] for subvarieties of $(\mathbb{P}^1)^n$) and that result was proven when the maps are defined over $\overline{\mathbb{Q}}$.

1D. ***The dynamical Manin–Mumford and the dynamical Bogomolov conjectures.*** Our Conjecture 1.1 is related to (and, in fact, motivated by) the dynamical Manin–Mumford conjecture and the dynamical Bogomolov conjecture, proposed in [Zhang 2006]. We state next a special case of the dynamical Manin–Mumford conjecture and of the dynamical Bogomolov conjecture for split endomorphisms of $\mathbb{A}^n$.

Theorem 1.7 [Ghioca et al. 2018]. *Let $f_1, \dots, f_n \in \overline{\mathbb{Q}}[x]$ be disintegrated polynomials of degree $d > 1$ and we let φ be their coordinatewise action on $\mathbb{A}^n$, i.e., $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ is given by $(x_1, \dots, x_n) \mapsto (f_1(x_1), \dots, f_n(x_r))$. For any irreducible $\overline{\mathbb{Q}}$-subvariety $X \subset \mathbb{A}^n$, if X contains a Zariski dense set of preperiodic points, then X is preperiodic. Furthermore, if for each $\epsilon > 0$, the set of points $(a_1, \dots, a_n) \in X(\overline{\mathbb{Q}})$ such that*

$$\hat{h}_{f_1}(a_1) + \cdots + \hat{h}_{f_n}(a_n) < \epsilon$$

is Zariski dense in X, then X is a preperiodic subvariety.

In Theorem 1.7, given a polynomial $f \in \overline{\mathbb{Q}}[x]$ of degree larger than 1, $\hat{h}_f(\cdot)$ is the canonical height defined as $\hat{h}_f(a) := \lim_{n\to\infty} h(f^n(a))/\deg(f)^n$ for any $a \in \overline{\mathbb{Q}}$, where $h(\cdot)$ is the usual Weil height. For more details regarding heights, see [Bombieri and Gubler 2006].

Actually, in [Ghioca et al. 2018, Theorem 1.1], the conclusion of Theorem 1.7 was established for all polarizable endomorphisms of $(\mathbb{P}^1)^n$, i.e., maps of the form $(x_1, \dots, x_n) \mapsto (f_1(x_1), \dots, f_n(x_n))$ where each $f_i \in \overline{\mathbb{Q}}(x)$ is a rational function of degree $d \geq 2$, which is not conjugate to a monomial, a $\pm$Chebyshev polynomial, or a Lattés map. We will prove in Theorem 3.1 a slightly more precise version of Theorem 1.7 for any subvariety of $\mathbb{A}^n$ which contains a Zariski dense set of periodic points.

In Theorem 1.7, if each polynomial f_i is conjugated with either a monomial or a $\pm$Chebyshev polynomial, then we recover the classical conjectures of Manin–Mumford and Bogomolov for algebraic tori. Actually, those conjectures (including in their version for abelian varieties) motivated Zhang to formulate in the early 1990s a far-reaching dynamical conjecture for polarizable algebraic dynamical systems generalizing both these classical diophantine problems and Theorem 1.7 (see also [Zhang 2006]).

In Theorem 1.7, since the coordinatewise action of φ on $\mathbb{A}^n$ is given by polynomials, one does not encounter the counterexamples (see [Ghioca et al. 2011]) to the original formulation of the dynamical Manin–Mumford conjecture (and of the dynamical Bogomolov conjecture), and hence one is not expected to require the stronger hypothesis for the reformulation from [Ghioca et al. 2011, Conjecture 1.4] of the dynamical Manin–Mumford conjecture. We note that Theorem 1.7 was initially proven when $X \subset \mathbb{A}^n$ is a curve in [Ghioca et al. 2015].

1E. ***Strategy for our proof.*** We prove Theorem 1.2 by splitting it into its 3 nontrivial cases (I)-(III), i.e., X is a hypersurface (Theorem 3.1), X is a curve (Theorem 4.1) and finally, X has codimension 2 (Theorem 5.1). The common ingredients for proving these results are the classification of periodic subvarieties of $\mathbb{A}^n$ under the coordinatewise action of n one-variable polynomials (as obtained in [Medvedev and Scanlon 2014], along with some further refinements obtained in [Ghioca and Nguyen 2016]) and also the proof of the dynamical Manin–Mumford and of the dynamical Bogomolov conjectures for endomorphisms of $(\mathbb{P}^1)^n$ (see Theorem 1.7 and [Ghioca et al. 2015; 2018]). In the case of curves $X \subset \mathbb{A}^n$, we also need to employ the recent result of [Xie 2017], who proved the dynamical Mordell–Lang conjecture for plane curves.

We discuss next a bit more about the actual strategy of proof for our results. First, we note that the case of hypersurfaces in Theorem 1.2 (see also its extension from Theorem 3.1) is significantly easier than both the case of curves and also the case of subvarieties of codimension 2 from Theorem 1.2. Next, we sketch a proof for a special case of both Theorems 4.1 and 5.1.

Assume $X \subset \mathbb{A}^3$ is a curve which contains an infinite set of points in common with the union of all periodic curves of $\mathbb{A}^3$. We assume $f_1 = f_2 = f_3 =: f$ is a polynomial which commutes only with iterates of itself; this is actually the generic case for a polynomial mapping. With this assumption, the result of [Medvedev and Scanlon 2014] yields that each periodic curve of $\mathbb{A}^3$ (which projects dominantly on each coordinate axis) is of the form

$$C_{k,\ell} := \{(x, f^k(x), f^{k+\ell}(x)) : x \in \mathbb{A}^1_{\overline{\mathbb{Q}}}\},$$

for some integers $k, \ell \geq 0$, after a suitable reordering of the coordinate axes. We show that we can reduce to the case $X \cap \bigcup_{k,\ell} C_{k,\ell}$ is infinite. Now, if there exists some integer j such that either $X \cap \bigcup_k C_{k,j}$ or $X \cap \bigcup_\ell C_{j,\ell}$ is infinite, we derive that X is contained in a periodic surface of $\mathbb{A}^3$. So, then we are left with the case that there exists an infinite set of pairs (k_n, ℓ_n) such that

$$X \cap C_{k_n,\ell_n} \neq \varnothing \quad \text{and} \quad \lim_{n\to\infty} k_n = \lim_{n\to\infty} \ell_n = \infty.$$

Then letting $(a_n, b_n, c_n) \in (X \cap C_{k_n,\ell_n})(\overline{\mathbb{Q}})$, using the fact that for each point on X, the height of any given coordinate is bounded uniformly (depending only on X, but independently of the given point) in terms of the heights of the other two coordinates of the point, while $\hat{h}_f(c_n) \gg \hat{h}_f(b_n) \gg \hat{h}_f(a_n)$, we can show that

$$\lim_{n\to\infty} \hat{h}_f(a_n) = \lim_{n\to\infty} \hat{h}_f(b_n) = 0.$$

This allows us to apply Theorem 1.7 to derive that the projection of X on the first two coordinate axes must be a periodic curve and therefore, X must be contained in a periodic surface.

1F. *Plan for our paper.* In Section 2, using [Medvedev and Scanlon 2014] (along with its refinements from [Nguyen 2015; Ghioca and Nguyen 2016]) we introduce the structure of periodic subvarieties of $\mathbb{A}^n$ under the coordinatewise action of n one-variable polynomials. In Section 3 we prove Theorem 1.2 for hypersurfaces $X \subset \mathbb{A}^n$ (see Theorem 3.1, which actually proves that *any* subvariety of $\mathbb{A}^n$ containing a Zariski dense set of periodic points must be periodic itself). Then we continue by proving Theorem 1.2 when X is a curve (see Theorem 4.1) in Section 4. We conclude our paper by proving Theorem 1.2 when $\operatorname{codim}(X) = 2$ in Section 5 (see Theorem 5.1, which proves that if any irreducible subvariety $X \subset \mathbb{A}^n$ of codimension at least equal to 2 intersects $\mathrm{Per}^{[n-1]}$ in a Zariski dense subset, then X is contained in a periodic hypersurface).

2. Structure of preperiodic subvarieties

Most of this section is taken from [Ghioca and Nguyen 2016; 2017] which, in turn, follows from [Medvedev and Scanlon 2014; Nguyen 2015]. Throughout this section, let $n \geq 2$, and let $f_1, \ldots, f_n$ be disintegrated polynomials in $\mathbb{C}[x]$. For $m \geq 2$, an irreducible curve (or more generally, a higher dimensional subvariety) in $\mathbb{A}^m$ is said to be *fibered* if its projection to one of the coordinate axes is constant, otherwise the curve (or the subvariety) is called *nonfibered*. For any two disintegrated polynomials $f(x)$ and $g(x)$, write $f \approx g$ if the self-map $(x, y) \mapsto (f(x), g(y))$ of $\mathbb{A}^2$ admits an irreducible nonfibered periodic curve. The relation $\approx$ is an equivalence relation in the set of disintegrated polynomials (see [Ghioca and Nguyen 2016, Section 7]).

Let $\varphi = f_1 \times \cdots \times f_n$ be the self-map of $\mathbb{A}^n$ given by $\varphi(x_1, \ldots, x_n) = (f_1(x_1), \ldots, f_n(x_n))$. Let s denote the number of equivalence classes arising from $f_1, \ldots, f_n$ (under $\approx$). Let $n_1, \ldots, n_s$ denote the sizes of these classes (hence $n_1+\cdots+n_s = n$). We relabel the polynomials $f_1, \ldots, f_n$ as $f_{i,j}$ for $1 \leq i \leq s$ and $1 \leq j \leq n_i$ according to the equivalence classes. After rearranging the polynomials $f_1, \ldots, f_n$ so that equivalence polynomials stay in blocks, we have $\varphi = \varphi_1 \times \cdots \times \varphi_s$, where φ_i is the self-map $f_{i,1} \times \cdots \times f_{i,n_i}$ of $\mathbb{A}^{n_i}$. There exist a positive integer N, nonconstant $p_{i,j} \in \mathbb{C}[x]$ for $1 \leq i \leq s$ and $1 \leq j \leq n_i$ and disintegrated $w_1, \ldots, w_s \in \mathbb{C}[x]$ in s different equivalence classes such that the following holds. For $1 \leq i \leq s$, let ψ_i be the self-map $w_i \times \cdots \times w_i$ on $\mathbb{A}^{n_i}$, and let $\psi = \psi_1 \times \cdots \times \psi_s$. Let η_i be the self-map $p_{i,1} \times \cdots \times p_{i,n_i}$ of $\mathbb{A}^{n_i}$ and let $\eta = \eta_1 \times \cdots \times \eta_s$. We have the commutative diagram:

$$\begin{array}{ccc} \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_s} & \xrightarrow{\psi} & \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_s} \\ \downarrow{\scriptstyle \eta} & & \downarrow{\scriptstyle \eta} \\ \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_s} & \xrightarrow{\psi^N} & \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_s} \end{array} \tag{2.1}$$

We have the following simple observations:

Lemma 2.2. *Let V be an irreducible φ-preperiodic subvariety of dimension r.*

(a) *Every irreducible component of $\eta^{-1}(V)$ is ψ-preperiodic and has dimension r.*

(b) *If V is φ-periodic then some irreducible component of $\eta^{-1}(V)$ is ψ-periodic.*

(c) *Let X be an irreducible subvariety in $\mathbb{A}^n$ and let $\mathrm{Per}_{\varphi}^{[r]}$ (respectively $\mathrm{Per}_{\psi}^{[r]}$) be the union of φ-periodic (respectively ψ-periodic) subvarieties of codimension r. If $X \cap \mathrm{Per}_{\varphi}^{[r]}$ is Zariski dense in X then there is an irreducible component X' of $\eta^{-1}(X)$ such that $X' \cap \mathrm{Per}_{\psi}^{[r]}$ is Zariski dense in X'.*

Proof. Part (a) follows from the commutative diagram (2.1) and the fact that η is finite. For part (b), if $\varphi^{M_0}(V) = V$ then ψ^{M_0} maps the set of irreducible components of $\eta^{-1}(V)$ to itself; hence at least one element in this set is a ψ-periodic subvariety.

For part (c), we have a collection of points $\{P_i : i \in \mathcal{S}\}$ that is Zariski dense in X and satisfies the property that for each $i \in \mathcal{S}$, there is an irreducible φ-subvariety V_i of codimension r such that $P_i \in X \cap V_i$. For each $i \in \mathcal{S}$, there is an irreducible component W_i of $\eta^{-1}(V_i)$ that is ψ-periodic and there is a point $Q_i \in W_i$ such that $\eta(Q_i) = P_i$. Let $X_1, \ldots, X_M$ denote all the irreducible components of $\eta^{-1}(X)$. We partition $\mathcal{S}$ into $\mathcal{S}_1, \ldots, \mathcal{S}_M$ such that $i \in \mathcal{S}_j$ implies $Q_i \in X_j$ for every $1 \le j \le M$. We claim that there exists some $j \in \{1, \ldots, M\}$ such that $\{Q_i : i \in \mathcal{S}_j\}$ is Zariski dense in X_j; consequently $X_j \cap \mathrm{Per}_{\psi}^{[r]}$ is Zariski dense in X_j. To prove this claim, assume that the Zariski closure of $\{Q_i : i \in \mathcal{S}_j\}$ is strictly smaller than X_j for every $j \in \{1, \ldots, M\}$. Then the image under η of the union of these M Zariski closures contains $\{P_i : i \in \mathcal{S}\}$ and is strictly smaller than X, a contradiction. □

Remark 2.3. We will also use the following simple observation which can be proved by arguments which are similar to the ones employed in the proof of part (c) above. If X is an irreducible subvariety of $\mathbb{A}^n$ and $\{V_i : i \in \mathcal{S}\}$ is a collection of irreducible subvarieties of $\mathbb{A}^n$ such that $X \cap \bigcup_{i \in \mathcal{S}} V_i$ is Zariski dense in X and $\mathcal{S}_1, \ldots, \mathcal{S}_M$ is a partition of $\mathcal{S}$ then there exists j such that $X \cap \bigcup_{i \in \mathcal{S}_j} V_i$ is Zariski dense in X.

Each irreducible φ-preperiodic subvariety V of $\mathbb{A}^n$ has the form $V_1 \times \cdots \times V_s$ where each V_i is an irreducible φ_i-preperiodic subvariety of $\mathbb{A}^{n_i}$. Let W be an arbitrary irreducible component of $\eta^{-1}(V)$. Then W is ψ-preperiodic and has the form $W_1 \times \cdots \times W_s$ where each W_i is an irreducible component of $\psi_i^{-1}(V_i)$ and it is ψ_i-preperiodic. Note that ψ_i is the coordinate-wise self-map of $\mathbb{A}^{n_i}$ induced by the *common* polynomial w_i.

Let f be a disintegrated polynomial and let $\Phi = f \times \cdots \times f$ be the corresponding self-map of $\mathbb{A}^n$. We recall the structure of Φ-periodic subvarieties of $\mathbb{A}^n$ given in [Ghioca and Nguyen 2016, Section 2]. Write $I_n = \{1, \ldots, n\}$. For each *ordered* subset J of I_n, we define:

$$\mathbb{A}^J := \mathbb{A}^{|J|}$$

equipped with the canonical projection $\pi_J : \mathbb{A}^n \to \mathbb{A}^J$. In this paper, we will consider ordered subsets of I_n whose orders need not be induced from the usual order of the set of integers. If $J_1, \ldots, J_m$ are ordered subsets of I_n which partition I_n, then we have the canonical isomorphism

$$(\pi_{J_1}, \ldots, \pi_{J_m}) : \mathbb{A}^n = \mathbb{A}^{J_1} \times \cdots \times \mathbb{A}^{J_m}.$$

For each irreducible subvariety V of $\mathbb{A}^n$, let J_V denote the set of all $j \in I_n$ such that the projection from V to the j-th coordinate axis is constant. If $J_V \neq \varnothing$, we equip J_V with the natural order of the set of integers, and we let $a_V \in \mathbb{A}^{J_V}(\mathbb{C})$ denote $\pi_{J_V}(V)$. Even when $J_V = \varnothing$, we will *vacuously* define $(\mathbb{A}^1)^{J_V}$ as the variety consisting of one point and define a_V to be that point. We have the following:

Proposition 2.4. (a) *Let V be an irreducible Φ-periodic subvariety of $\mathbb{A}^n$ of dimension r. Then there exists a partition of $I_n - J_V$ into r nonempty subsets $J_1, \dots, J_r$ such that the following hold. We fix an order on each $J_1, \dots, J_r$, and identify*

$$\mathbb{A}^n = \mathbb{A}^{J_V} \times \mathbb{A}^{J_1} \times \cdots \times \mathbb{A}^{J_r}.$$

For $1 \leq i \leq r$, let Φ_i denote the coordinatewise self-map of $\mathbb{A}^{J_i}$ induced by f. For $1 \leq i \leq r$, there exists an irreducible Φ_i-periodic curve C_i in $\mathbb{A}^{J_i}$ such that

$$V = \{a_V\} \times C_1 \times \cdots \times C_r.$$

(b) *Let C be an irreducible Φ-periodic curve in $\mathbb{A}^n$ and denote $m := |I_n - J_C| \geq 1$. Then there exist a permutation $(i_1, \dots, i_m)$ of $I_n - J_C$ and nonconstant polynomials $g_2, \dots, g_m \in \overline{\mathbb{Q}}[x]$ such that C is given by the equations $x_{i_2} = g_2(x_{i_1}), \dots, x_{i_m} = g_m(x_{i_{m-1}})$. Furthermore, the polynomials $g_2, \dots, g_m$ commute with an iterate of f.*

Remark 2.5. Let C be a nonfibered irreducible preperiodic curve in $\mathbb{A}^2$ under the map $\Phi(x, y) = (f(x), f(y))$. Then $\Phi^r(C)$ is periodic for some r. So we know that C satisfies an equation of the form $f^r(x_2) = g(f^r(x_1))$ or $f^r(x_1) = g(f^r(x_2))$ where g commutes with an iterate of f. We can express both cases by an equation of the form $g(x_1) = G(x_2)$ where both g and G commute with an iterate of f.

Remark 2.6. The above discussion gives a very precise description of irreducible φ-preperiodic subvarieties of $\mathbb{A}^n$ (recall that $\varphi = f_1 \times \cdots \times f_n$). Occasionally, the following simpler observation is sufficient for our purpose. Let $V \subsetneq \mathbb{A}^n$ be an irreducible φ-periodic subvariety. Then there exist $1 \leq i < j \leq n$ and an irreducible curve C in $\mathbb{A}^2$ which is periodic under $(x, y) \mapsto (f_i(x), f_j(y))$ such that $V \subseteq \pi^{-1}(C)$ where π is the projection from $\mathbb{A}^n$ to the i-th and j-th coordinates $\mathbb{A}^2$.

Remark 2.7. The permutation $(i_1, \dots, i_m)$ mentioned in part (b) of Proposition 2.4 induces the order $i_1 \prec \cdots \prec i_m$ on $I_n - J_C$. Such a permutation and its induced order are not uniquely determined by V. For example, let L be a *linear* polynomial commuting with an iterate of f. Let C be the periodic curve in $\mathbb{A}^2$ defined by the equation $x_2 = L(x_1)$. Then $I - J_C = \{1, 2\}$, and $1 \prec 2$ is an order satisfying the conclusion of part (b). However, we can also express C as $x_1 = L^{-1}(x_2)$. Then the order $2 \prec 1$ also satisfies part (b). Therefore, in part (a), the choice of an order on each J_i is not unique. Nevertheless, the partition of $I_n - J_V$ into the subsets $J_1, \dots, J_r$ is unique (see [Nguyen 2015, Section 2]).

Next we describe all polynomials g commuting with an iterate of f.

Proposition 2.8. *Let $f \in \mathbb{C}[x]$ be a disintegrated polynomial of degree greater than 1. We have:*

(a) *If* $g \in \mathbb{C}[x]$ *has degree at least* 2 *such that* g *commutes with an iterate of* f *then* g *and* f *have a common iterate.*

(b) *Let* $M(f^\infty)$ *denote the collection of all linear polynomials commuting with an iterate of* f. *Then* $M(f^\infty)$ *is a finite cyclic group under composition.*

(c) *Let* $\tilde{f} \in \mathbb{C}[x]$ *be a polynomial of minimum degree* $\tilde{d} \geq 2$ *such that* $\tilde{f}$ *commutes with an iterate of* f. *Then there exists* $D = D_f > 0$ *relatively prime to the order of* $M(f^\infty)$ *such that* $\tilde{f} \circ L = L^D \circ \tilde{f}$ *for every* $L \in M(f^\infty)$.

(d) $\{\tilde{f}^m \circ L : m \geq 0, L \in M(f^\infty)\} = \{L \circ \tilde{f}^m : m \geq 0, L \in M(f^\infty)\}$, *and these sets describe exactly all polynomials* g *commuting with an iterate of* f. *As a consequence, there are only finitely many polynomials of bounded degree commuting with an iterate of* f.

Remark 2.9. In the diagram (2.1), if $f_1, \ldots, f_n$ are in $\overline{\mathbb{Q}}[x]$ then the polynomials w_i and $p_{i,j}$ can be chosen to be in $\overline{\mathbb{Q}}[x]$. In Proposition 2.8, if $f(x) \in \overline{\mathbb{Q}}[x]$ then $\tilde{f} \in \overline{\mathbb{Q}}[x]$ and elements of $M(f^\infty)$ are in $\overline{\mathbb{Q}}[x]$.

We will use the following immediate corollary to recognize when a point is f-periodic.

Corollary 2.10. *Let* $f \in \mathbb{C}[x]$ *be a disintegrated polynomial of degree greater than* 1.

(a) *Let* $g(x) \in \mathbb{C}[x]$ *such that* $\deg(g) \geq 2$ *and* g *commutes with an iterate of* f. *Then* $\alpha \in \mathbb{C}$ *is* g*-periodic if and only if it is* f*-periodic.*

(b) *Let* $p(x) \in \mathbb{C}[x]$ *such that* $\deg(p) \geq 1$ *and* p *commutes with an iterate of* f. *Let* $\alpha \in \mathbb{C}$ *be* f*-periodic. Then* $p(\alpha)$ *is also* f*-periodic.*

(c) *If* α *is* f*-preperiodic then for any polynomial* g *that commutes with an iterate of* f *and* $\deg(g)$ *is sufficiently large,* $g(\alpha)$ *is* f*-periodic.*

(d) *If* α *is* f*-preperiodic then the set*

$$\{g(\alpha) : g \text{ commutes with an iterate of } f\}$$

is finite.

Proof. Part (a) is obvious since g and f have a common iterate. For part (b), choose m such that f^m commutes with p and $\alpha = f^m(\alpha)$. Then $f^m(p(\alpha)) = p(f^m(\alpha)) = p(\alpha)$. For part (c), let $r \geq 0$ such that $f^r(\alpha)$ is f-periodic, then if $\deg(g) \geq \deg(f)^r$, we can write $g = g_1 \circ f^r$ where g_1 commutes with an iterate of f by Proposition 2.8(d). Now $g(\alpha) = g_1(f^r(\alpha))$ is f-periodic by part (b). For part (d), let $\tilde{f}$ be as in Proposition 2.8, we can write g as $L \circ \tilde{f}^m$ for some $m \geq 0$ and $L \in M(f^\infty)$. Since α is $\tilde{f}$-preperiodic and $M(f^\infty)$ is finite, there are only finitely many possibilities for $g(\alpha)$. □

We now consider the more general self-map $\varphi = f_1 \times \cdots \times f_n$ as in the beginning of this section. Let V be an irreducible φ-preperiodic subvariety of $\mathbb{A}^n$ with $r := \dim(V)$. As before, J_V denotes the set of $i \in I_n$ such that the projection from V to the i-th coordinate $\mathbb{A}^1$ is constant and $a_V \in \mathbb{A}^{J_V}(\mathbb{C})$ is the image $\pi^{J_V}(V)$. By Proposition 2.4 and the diagram (2.1), we can partition the set $I_n \setminus J_V$ into r nonempty subsets $J_1, \dots, J_r$ such that after identifying

$$\mathbb{A}^n = \mathbb{A}^{J_V} \times \mathbb{A}^{J_1} \times \cdots \times \mathbb{A}^{J_r},$$

we have:

$$V = \{a_V\} \times C_1 \times \cdots \times C_r$$

where each C_j is a preperiodic curve in $\mathbb{A}^{J_j}$ with respect to the coordinatewise self-map induced by the polynomials f_i's for $i \in J_j$. Moreover, if V is periodic then a_V and each C_j are periodic. Since each C_i is necessarily nonfibered thanks to the definition of J_V, we have that $f \approx g$ for $f, g \in J_j$ for $1 \le j \le r$. We have the following:

Definition 2.11. The weak signature of V is the collection consisting of the set J_V and the partition of $I_n \setminus J_V$ into the sets $J_1, \dots, J_r$.

3. Proof of Theorem 1.2 for hypersurfaces

The case of hypersurfaces $X \subset \mathbb{A}^n$ in Theorem 1.2 is a consequence of the following more general result.

Theorem 3.1. *Let $f_1, \dots, f_n$, d, and φ be as in Theorem 1.2. Let X be an irreducible subvariety of $\mathbb{A}^n$ such that X contains a Zariski dense set of φ-periodic points, then X is periodic. Consequently, Theorem 1.2 holds when* $\operatorname{codim}(X) = 1$.

We thank the referee for suggesting the following proof for Theorem 3.1, which is simpler than our original proof.

Proof. By Theorem 1.7 X is preperiodic; so there exist positive integers m and r such that $\varphi^{m+r}(X) = \varphi^m(X)$. We define a function

$$\Psi : \mathbb{N} \to \{1, 2, \dots, m+r-1\}$$

given by

$$\Psi(n) = \begin{cases} n & \text{if } 1 \le n \le m-1, \\ \rho & \text{if } n \ge m, \end{cases}$$

where ρ is the unique integer in the set $\{m, m+1, \dots, m+r-1\}$ satisfying the property that $\rho \equiv n \pmod r$. In particular, using the fact that $\varphi^m(X) = \varphi^{m+r}(X)$, we get that $\varphi^n(X) = \varphi^{\Psi(n)}(X)$ for each $n \in \mathbb{N}$.

Let S be the set of periodic points in X. For each point $x \in S$, we denote by $r_x \ge 1$ its period (under the action of φ). Then for each $i = 1, \dots, m+r-1$, we let

$$S_i := \{x \in S : \Psi(r_x) = i\}.$$

Since S is Zariski dense in X (and X is irreducible), there exists some $i \in \{1, \dots, r+m-1\}$ such that S_i is Zariski dense in X. Now, for each $x \in S_i$, we have that

$$x = \varphi^{r_x}(x) \in \varphi^{r_x}(X) = \varphi^i(X).$$

It follows that $X \subseteq \varphi^i(X)$ and since φ is finite, we conclude that $X = \varphi^i(X)$; therefore, X is periodic, as claimed. $\square$

4. Proof of Theorem 1.2 for curves

In this section we prove the following result

Theorem 4.1. *Theorem 1.2 holds when $X \subset \mathbb{A}^n$ is a curve.*

Proof. The case $n = 2$ follows from Theorem 3.1. We will prove next the result for $n \in \{3, 4\}$ and proceed by induction for $n \geq 5$. We recall the notation and terminology from Section 2. We are given that the curve X has a Zariski dense (i.e., infinite) set of points each of which is contained in a periodic subvariety V of codimension 2. Since there are only finitely many possibilities for the weak signature, by Remark 2.3, we may assume that all of the above periodic subvarieties have a common weak signature consisting of a (possibly empty) subset $\mathcal{J} = J_V$ of I_n and a partition of $I_n \setminus \mathcal{J}$ into $n-2$ nonempty subsets $J_1, \ldots, J_{n-2}$. Let h denote the absolute logarithmic Weil height on $\mathbb{P}^1(\overline{\mathbb{Q}})$. We also let h denote the height on $\mathbb{A}^n(\overline{\mathbb{Q}}) \subset (\mathbb{P}^1)^n(\overline{\mathbb{Q}})$ given by

$$h(x_1, \ldots, x_n) = h(x_1) + \cdots + h(x_n).$$

For each f_i, let $\hat{h}_{f_i}$ denote the canonical height on $\mathbb{P}^1(\overline{\mathbb{Q}})$ associated to f_i, and let $\hat{h}$ denote the function on $\mathbb{A}^n \subset (\mathbb{P}^1)^n(\overline{\mathbb{Q}})$ given by

$$\hat{h}(x_1, \ldots, x_n) = \hat{h}_{f_1}(x_1) + \cdots + \hat{h}_{f_n}(x_n).$$

Note that $\hat{h}$ is the canonical height associated to φ (which is the coordinatewise action of the polynomials f_i on $\mathbb{A}^n$). We refer the readers to [Bombieri and Gubler 2006; Silverman 2007, Chapter 3] for more details on height and canonical height functions.

4A. ***The case when the ambient space has dimension 3.*** Without loss of generality, we have the following possibilities for the weak signature $(\mathcal{J}, J_1)$:

Case A: $\mathcal{J} = \varnothing$ and $J_1 = \{1, 2, 3\}$. By part (c) of Lemma 2.2, we may assume that $f_1 = f_2 = f_3 =: f$. By Proposition 2.4 and Remark 2.3, we may assume that there are infinitely many points $\{P_i\}_{i=1}^{\infty}$ such that for each i, there is a periodic curve V_i defined by the equations $x_2 = g_{i,2}(x_1)$ and $x_3 = g_{i,3}(x_2)$ such that $P_i \in X \cap V_i$ where $g_{i,2}$ and $g_{i,3}$ are polynomials commuting with an iterate of f. If $\{\deg(g_{i,2})\}_{i\geq 1}$ has a bounded subsequence then Proposition 2.8(d) yields that there exists a polynomial g such that $g_{i,2} = g$ for infinitely many i. Hence X is contained in the periodic surface defined by $x_2 = g(x_1)$ because it is a curve containing infinitely many points from this surface. The case when $\{\deg(g_{i,3})\}_{i\geq 1}$ has a bounded subsequence is treated similarly. We now assume that

$$\lim_{i\to\infty} \deg(g_{i,2}) = \lim_{i\to\infty} \deg(g_{i,3}) = \infty.$$

Write $P_i = (a_i, b_i, c_i)$. Let $\pi_{1,2}$ denote the projection from $\mathbb{A}^3$ to the first two coordinates $\mathbb{A}^2$ and let Y be the Zariski closure of $\pi_{1,2}(X)$.

We consider the case when $\pi_{1,2}$ is nonconstant on X, in other words Y is a curve in $\mathbb{A}^2$. Then there exist positive constants C_1 and C_2 depending only on the curve X such that for every point $(a, b, c) \in X(\overline{\mathbb{Q}})$, we have

$$h(c) \le C_1 \max\{h(a), h(b)\} + C_2. \tag{4.2}$$

Inequality (4.2) is a special case of [Ghioca and Nguyen 2016, Lemma 3.2(b)] (see also Corollary 3.4 of that paper). Essentially, inequality (4.2) says that the height of each coordinate of a point on a curve is bounded in terms of the heights of the other coordinates, as long as the curve is not fibered. Since $|h - \hat{h}_f| = O(1)$, there exist positive constants C_3 and C_4 depending on X and f such that

$$\hat{h}_f(c) \le C_3 \max\{\hat{h}_f(a), \hat{h}_f(b)\} + C_4$$

for every $(a, b, c) \in X(\overline{\mathbb{Q}})$ (see also [Ghioca and Nguyen 2016, Corollary 3.4]). In particular, this inequality holds for the points $P_i = (a_i, b_i, c_i)$. On the other hand, we have

$$\hat{h}_f(c_i) = \deg(g_{i,3})\hat{h}_f(b_i) \quad \text{and} \quad \hat{h}_f(b_i) = \deg(g_{i,2})\hat{h}_f(a_i).$$

Overall, we have

$$\deg(g_{i,3}) \max\{\hat{h}_f(a_i), \hat{h}_f(b_i)\} \le \hat{h}_f(c_i) \le C_3 \max\{\hat{h}_f(a_i), \hat{h}_f(b_i)\} + C_4.$$

Since $\lim \deg(g_{i,3}) = \infty$, we get $\lim_{i\to\infty}(\max\{\hat{h}_f(a_i), \hat{h}_f(b_i)\}) = 0$ and so, Theorem 1.7 yields that the curve Y is preperiodic.

A more careful analysis shows that X is contained in a *periodic* surface, as follows.

First, consider the case when the projection from X to the first or second coordinate $\mathbb{A}^1$ is constant, then this constant, denoted γ, is necessarily preperiodic since Y is preperiodic. From $c_i = g_{i,3}(b_i) = g_{i,3}(g_{i,2}(a_i))$ and Corollary 2.10, we have that c_i is periodic for all sufficiently large i and the sequence $\{c_i\}_{i\ge1}$ consists of only finitely many points. Hence there is a periodic point ζ such that $c_i = \zeta$ for infinitely many i. We conclude that X is contained in the periodic surface $\mathbb{A}^2 \times \{\zeta\}$.

When the projection from X to neither the first nor second $\mathbb{A}^1$ is constant, by Proposition 2.4 and Remark 2.5, the preperiodic curve Y satisfies an equation of the form $g(x_1) = G(x_2)$ where g and G commute with an iterate of f. Therefore the point (a_i, b_i) satisfies both $g(a_i) = G(b_i)$ and $b_i = g_{i,2}(a_i)$.

The following observation will be used repeatedly throughout our proof.

Lemma 4.3. *With the above notation, for all i sufficiently large, we have that b_i is periodic.*

Proof of Lemma 4.3.. When i is sufficiently large so that $\deg(g_{i,2}) \ge \deg(g)$, from Proposition 2.8(d), we can write $g_{i,2} = u_i \circ g$ where u_i is a polynomial commuting with an iterate of f. Therefore

$$b_i = u_i(g(a_i)) = u_i(G(b_i))$$

and Corollary 2.10(a) implies that b_i is f-periodic. □

Using that b_i is periodic along with the fact that $c_i = g_{i,3}(b_i)$, we obtain that c_i is also f-periodic (by Corollary 2.10(b)). Let Y' be the Zariski closure of the projection from X to the second and third coordinates $\mathbb{A}^2$. Since (b_i, c_i) is periodic for all sufficiently large i, we have that Y' is periodic (according to Theorem 3.1). Hence X is contained in the periodic subvariety $\mathbb{A}^1 \times Y'$.

The case when $\pi_{1,2}$ is constant on X is obvious. Indeed, $X = \{(a, b)\} \times \mathbb{A}^1$ and since $X \cap V_1 \neq \varnothing$, we have $b = g_{1,2}(a)$ and $g_{1,2}$ commutes with an iterate of f. Hence X is contained in the periodic surface defined by $x_2 = g_{1,2}(x_1)$.

Case B: $\mathcal{J} = \{1\}$ and $J_1 = \{2, 3\}$. As in Case A, we may assume that $f_2 = f_3 =: f$ and there are infinitely many points $\{P_i = (a_i, b_i, c_i)\}_{i\geq 1}$ such that for each i, there is a periodic curve V_i defined by $x_1 = \zeta_i$ and $x_3 = g_i(x_2)$ such that $P_i \in X \cap V_i$ where ζ_i is f_1-preperiodic and g_i commutes with an iterate of f. By arguments similar to Case A, we may assume $\lim_{i\to\infty} \deg(g_i) = \infty$.

When $\pi_{1,2}$ is nonconstant on X, we can use similar arguments as in Case A. This time, we have an inequality of the form

$$\hat{h}_f(c) \leq C_5 \max\{\hat{h}_{f_1}(a), \hat{h}_f(b)\} + C_6 \tag{4.4}$$

for every $(a, b, c) \in X(\overline{\mathbb{Q}})$ where C_5 and C_6 are constants depending only on X, f_1, and f. So we can conclude that $\lim_{i\to\infty} \hat{h}_f(b_i) = 0$. Since Y contains the Zariski dense set $\{(a_i = \zeta_i, b_i)\}_i$, we have that Y is preperiodic, by Theorem 1.7. If the projection π_1 from X (and Y) to the first $\mathbb{A}^1$ is constant then we have $a_i = \zeta_1$ for every i and X is contained in the periodic surface $\{\zeta_1\} \times \mathbb{A}^2$. If the projection π_2 from X (and Y) to the second $\mathbb{A}^1$ is constant, then inequality (4.4) combined with the fact that $a_i = \zeta_i$ is periodic and the fact that $\lim_{i\to\infty} \deg(g_i) = \infty$ yields that b_i must be preperiodic. But then, because b_i is constant as we vary i and $\deg(g_i) \to \infty$, Corollary 2.10(c) yields that c_i must be constant *and* periodic, thus providing the desired conclusion in Theorem 4.1. If π_1 and π_2 are nonconstant then Y satisfies an equation $g(x_1) = G(x_2)$, where g and G commute with an iterate of f. In particular $g(\zeta_i) = g(a_i) = G(b_i)$; so, by Corollary 2.10, $G(b_i)$ is f-periodic (note that ζ_i is periodic). When $\deg(g_i) \geq \deg(G)$, by (the proof of) part (c) of Corollary 2.10, we have that $c_i = g_i(b_i)$ is also periodic. Now the Zariski closure of the projection from X to the first and third coordinates $\mathbb{A}^2$ contains the Zariski dense set $\{(a_i, c_i) : i \text{ is large}\}$ of periodic points, it must be periodic thanks to Theorem 3.1. Hence X is contained in a periodic surface.

The case $\pi_{1,2}$ is constant on X is also obvious. Indeed, $X = \{(a, b)\} \times \mathbb{A}^1$ and since $X \cap V_1 \neq \varnothing$, we have that $a = \zeta_1$. Hence X is contained in the periodic surface $\{\zeta_1\} \times \mathbb{A}^2$.

Case C: $\mathcal{J} = \{1, 2\}$ and $J_1 = \{3\}$. This time, each periodic curve V_i has the form $\{(\alpha_i, \beta_i)\} \times \mathbb{A}^1$ where α_i is f_1-periodic and β_i is f_2-periodic. If $\pi_{1,2}$ is nonconstant on X then Theorem 3.1 implies that Y is a periodic curve in $\mathbb{A}^2$, hence X is contained in the periodic surface $Y \times \mathbb{A}^1$. If $\pi_{1,2}$ is constant on X, since $X \cap V_1 \neq \varnothing$, we have $X = V_1$ is periodic.

4B. ***The case when the ambient space has dimension 4.*** We will need the following result:

Proposition 4.5. *Let* $f(x), g(x) \in \overline{\mathbb{Q}}[x]$ *with* $\deg(f) = \deg(g) =: d \geq 2$. *Let* $C \subset \mathbb{A}^2$ *be an irreducible* $\overline{\mathbb{Q}}$*-curve with the following properties*:

- *C is nonfibered.*
- *There exist $\alpha, \beta \in \overline{\mathbb{Q}}$ such that $C \cap (\mathcal{O}_f(\alpha) \times \mathcal{O}_g(\beta))$ is infinite.*

Then C is periodic under the action $(x_1, x_2) \mapsto (f(x_1), g(x_2))$.

Before proceeding to its proof, we explain the necessity of Proposition 4.5 for our proof of Theorem 4.1 when $n = 4$. In this case, we have a curve $X \subset \mathbb{A}^4$ which intersects the union of all periodic surfaces in an infinite set. For example, if $f_1 = f_2 = f_3 = f_4 =: f$ we could deal with the special case that X projects to a point (a, b) on the first two coordinate axes, where both a and b are not preperiodic under the action of f; we let Y be the projection of X on the last two coordinate axes of $\mathbb{A}^4$. Each surface $S_{k,\ell} \subset \mathbb{A}^4$ given by the equations $x_3 = f^k(x_1)$ and $x_4 = f^\ell(x_2)$ is periodic. Next, assume

$$X \cap \Bigl(\bigcup_{k,\ell} S_{k,\ell}\Bigr) \text{ is infinite.}$$

So, we are left to prove that if $Y \cap (\mathcal{O}_f(a) \times \mathcal{O}_f(b))$ is infinite, then Y is periodic under the induced action of f on the last two coordinate axes of $\mathbb{A}^4$, which is precisely the conclusion from Proposition 4.5.

Proof of Proposition 4.5. As in Cases A and B (see also [Ghioca and Nguyen 2016, Corollary 3.4]), since C is nonfibered there exist positive constants C_7 and C_8 depending on C, f, and g such that for each $(a_1, a_2) \in C(\overline{\mathbb{Q}})$, we have

$$\max\{\hat{h}_f(a_1), \hat{h}_g(a_2)\} \le C_7 \min\{\hat{h}_f(a_1), \hat{h}_g(a_2)\} + C_8. \tag{4.6}$$

Now, since $C \cap (\mathcal{O}_f(\alpha) \times \mathcal{O}_g(\beta))$ is infinite and C projects dominantly to both coordinates, we get that α and β are not f-preperiodic and g-preperiodic, respectively. Hence $\hat{h}_f(\alpha) > 0$ and $\hat{h}_g(\beta) > 0$. From this observation, inequality (4.6) for each point $(f^m(\alpha), g^n(\beta)) \in C(\overline{\mathbb{Q}})$, and the fact that $\hat{h}_f(f^m(\alpha)) = d^m \hat{h}_f(a)$ and $\hat{h}_g(g^n(\beta)) = d^n \hat{h}_g(\beta)$, we conclude that $|m - n|$ is uniformly bounded as we vary among all points $(f^m(\alpha), g^n(\beta)) \in C(\overline{\mathbb{Q}})$. Therefore, there exists an integer ℓ such that there exist infinitely many $(m, n) \in \mathbb{N}_0 \times \mathbb{N}_0$ with the property that $(f^m(\alpha), g^n(\beta)) \in C(\overline{\mathbb{Q}})$ and also $m - n = \ell$. Without loss of generality, we assume that $\ell \ge 0$, and therefore get that C contains infinitely many points from the orbit of $(f^\ell(\alpha), \beta)$ under the action of $(x_1, x_2) \mapsto (f(x_1), g(x_2))$. Since the dynamical Mordell–Lang conjecture (see [Bell et al. 2016, Chapter 3]) is known in the case of endomorphisms of $\mathbb{A}^2$ (as proven in [Xie 2017]), we conclude that C is periodic under the action of $(x_1, x_2) \mapsto (f(x_1), g(x_2))$, as desired. $\square$

We now return to the proof of Theorem 4.1. We have the following cases for the weak signature $(\mathcal{J}, J_1, J_2)$:

Case D: $|J_1| = 1$ or $|J_2| = 1$. Without loss of generality, assume $|J_2| = 1$, more specifically $J_2 = \{4\}$. Now there are infinitely many points $\{P_i = (a_i, b_i, c_i, d_i)\}_{i \ge 1}$ such that for each i, there is a periodic surface V_i such that $P_i \in X \cap V_i$. Moreover, we have that $V_i = W_i \times \mathbb{A}^1$ where W_i is a periodic curve under the self-map $f_1 \times f_2 \times f_3$ of $\mathbb{A}^3$.

Let $\pi_{1,2,3}$ denote the projection from $\mathbb{A}^4$ to the first three coordinates $\mathbb{A}^3$. If $\pi_{1,2,3}$ is nonconstant on X, then the Zariski closure Y of $\pi_{1,2,3}(X)$ in $\mathbb{A}^3$ is a curve and we can apply Theorem 4.1 to the data $(n = 3, f_1, f_2, f_3, Y)$ to conclude that Y is contained in a periodic surface S in $\mathbb{A}^3$. Hence X is contained in the periodic hypersurface $S \times \mathbb{A}^1$. The case $\pi_{1,2,3}$ is constant on X is obvious. We have that $X = \{(a, b, c)\} \times \mathbb{A}^1$. Since $P_i = (a_i, b_i, c_i, d_i) = (a, b, c, d_i)$ lies in $V_i = W_i \times \mathbb{A}^1$, we have that X itself is contained in the periodic subvariety V_i (for every i).

Case E: $|J_1| = |J_2| = 2$. Without loss of generality, assume $J_1 = \{1, 2\}$ and $J_2 = \{3, 4\}$. As in Case A, we may assume $f_1 = f_2 =: f$ and $f_3 = f_4 =: g$. By Proposition 2.4 and without loss of generality, we may assume that there are infinitely many points $\{P_i = (a_i, b_i, c_i, d_i)\}_{i \geq 1}$ such that for each i, there is a periodic surface V_i defined by $x_2 = U_i(x_1)$ and $x_4 = T_i(x_3)$ such that $P_i \in X \cap V_i$ and $U_i(x)$ and $T_i(x)$ commute with an iterate of $f(x)$ and $g(x)$, respectively. For such polynomials $U_i(x)$ and $T_i(x)$, and for any $a \in \overline{\mathbb{Q}}$, we have (see [Nguyen 2015, Lemma 3.3])

$$\hat{h}_f(U_i(a)) = \deg(U_i)\hat{h}_f(a) \quad \text{and} \quad \hat{h}_g(T_i(a)) = \deg(T_i)\hat{h}_g(a). \tag{4.7}$$

As in Case A, we may assume that $\lim_{i\to\infty} \deg(U_i) = \lim_{i\to\infty} \deg(T_i) = \infty$. Let $\pi_{1,3}$ denote the projection from $\mathbb{A}^4$ to the first and third coordinates $\mathbb{A}^2$ and let Y denote the Zariski closure of $\pi_{1,3}(X)$.

We consider first the case when $\pi_{1,3}$ is nonconstant on X, in other words Y is a curve in $\mathbb{A}^2$.

As in Case A, there are positive constants C_9 and C_{10} depending only on X and f such that for every point $(a, b, c, d) \in X(\overline{\mathbb{Q}})$, we have

$$\hat{h}_f(b) + \hat{h}_g(d) \leq C_9(\hat{h}_f(a) + \hat{h}_g(c)) + C_{10}.$$

Combining with (4.7) and the fact that $P_i = (a_i, b_i, c_i, d_i) \in X \cap V_i$, we have

$$(\deg(U_i) - C_9)\hat{h}_f(a_i) + (\deg(T_i) - C_9)\hat{h}_g(b_i) \leq C_{10}.$$

Since $\lim_{i\to\infty} \deg(U_i) = \lim_{i\to\infty} \deg(T_i) = \infty$, we get that $\lim_{i\to\infty} \hat{h}_f(a_i) = \lim_{i\to\infty} \hat{h}_g(c_i) = 0$. By Theorem 1.7, the curve Y is preperiodic under the map $(x_1, x_3) \mapsto (f(x_1), g(x_3))$. A more careful analysis shows that X is contained in a periodic subvariety as follows.

When the projection from X to the first (or respectively the third) coordinate is constant, then this constant is necessarily preperiodic since Y is preperiodic. Since $b_i = U_i(a_i)$ (respectively $d_i = T_i(c_i)$), we can argue as in Case A to conclude that there is an f-periodic point (respectively g-periodic point) ζ such that $b_i = \zeta$ (respectively $d_i = \zeta$) for infinitely many i. Hence X is contained in the periodic surface $\mathbb{A}^1 \times \{\zeta\} \times \mathbb{A}^2$ (respectively $\mathbb{A}^3 \times \{\zeta\}$).

Now consider the case when the projection from X to both the first and third coordinates is nonconstant, or equivalently Y is a nonfibered curve in $\mathbb{A}^2$. This implies $f \approx g$. By Lemma 2.2, we may assume that $f = g$ (i.e., $f_1 = f_2 = f_3 = f_4 = f$). Remark 2.5 gives that Y satisfies an equation of the form $g(x_1) = G(x_3)$ where g and G commute with an iterate of f. In particular $b_i = U_i(a_i)$, $d_i = T_i(c_i)$, and $g(a_i) = G(c_i)$. When i is sufficiently large so that $\deg(U_i) \geq \deg(g)$ and $\deg(T_i) \geq \deg(G)$, we can write

$$U_i = U_i^* \circ g \quad \text{and} \quad T_i = T_i^* \circ G$$

where U_i^* and T_i^* commute with an iterate of f. Obviously, either $\deg(U_i^*) \geq \deg(T_i^*)$ or $\deg(T_i^*) \geq \deg(U_i^*)$. By restricting to an infinite subsequence of $\{P_i\}$ and without loss of generality, we may assume that $\deg(T_i^*) \geq \deg(U_i^*)$ for every i. From Proposition 2.8, we can write $T_i^* = S_i \circ U_i^*$ where S_i commutes with an iterate of f. We have

$$d_i = T_i(c_i) = T_i^*(G(c_i)) = T_i^*(g(a_i)) = S_i(U_i^*(g(a_i))) = S_i(U_i(a_i)) = S_i(b_i).$$

If $\{\deg(S_i)\}_i$ has a bounded subsequence then by similar arguments as before, X would be contained in a periodic surface of the form $x_4 = S(x_2)$ and we are done. Now assume $\lim_{i\to\infty} \deg(S_i) = \infty$. Since the projection from X to the first 3 coordinates is nonconstant, there exist C_{11} and C_{12} such that:

$$\hat{h}_f(d_i) \leq C_{11} \max\{\hat{h}_f(a_i), \hat{h}_f(b_i), \hat{h}_f(c_i)\} + C_{12}.$$

On the other hand

$$\hat{h}_f(d_i) = \deg(T_i)\hat{h}_f(c_i), \quad \hat{h}_f(d_i) = \deg(S_i)\hat{h}_f(b_i) = \deg(S_i)\deg(U_i)\hat{h}_f(a_i)$$

and $\{\deg(S_i)\}_i$, $\{\deg(T_i)\}_i$, and $\{\deg(U_i)\}_i$ become arbitrarily large; so, we conclude that

$$\lim_{i\to\infty} \hat{h}_f(a_i) = \lim_{i\to\infty} \hat{h}_f(b_i) = \lim_{i\to\infty} \hat{h}_f(c_i) = 0.$$

By Theorem 1.7, the Zariski closure Z of the projection from X to the first 2 coordinates $\mathbb{A}^2$ is preperiodic. We are assuming that the projection from X to the first coordinate is nonconstant. If the projection to the second coordinate is constant then it must be preperiodic (since Z is preperiodic), denoted γ. Now $d_i = S_i(b_i) = S_i(\gamma)$ and we can argue as in Case A to conclude that X is contained in a periodic hypersurface of the form $\mathbb{A}^3 \times \{\zeta\}$. It remains to treat the case when the projection to the second coordinate is nonconstant. Then Z satisfies an equation $g^*(x_1) = G^*(x_2)$ where g^* and G^* commute with an iterate of f. By similar arguments as in Case A (see Lemma 4.3), we conclude that b_i is f-periodic when i is sufficiently large, and so, $d_i = S_i(b_i)$ is also f-periodic. Then Theorem 3.1 implies that the projection from X to the second and fourth coordinates axes is a periodic curve and we are done since we obtain that X is contained in the periodic (irreducible) hypersurface in $\mathbb{A}^4$, which is the pullback of the aforementioned periodic plane curve under the projection map $(x_1, x_2, x_3, x_4) \mapsto (x_2, x_4)$.

Finally, we treat the case when $\pi_{1,3}$ is constant on X.

Write $\{(\alpha, \gamma)\} = \pi_{1,3}(X)$, hence $(a_i, c_i) = (\alpha, \gamma)$ for every i. If α is f-preperiodic then for all i sufficiently large, we get that $b_i = U_i(a_i) = U_i(\gamma)$ must be some given periodic point β and thus, X is contained in the periodic hypersurface $\mathbb{A}^1 \times \{\beta\} \times \mathbb{A}^2$ and hence, we are done. Therefore we may assume that α and γ are not f-preperiodic and g-preperiodic, respectively. Hence $\hat{h}_f(\alpha) > 0$ and $\hat{h}_g(\gamma) > 0$. From (4.7) and the fact that

$$\lim_{i\to\infty} \deg(U_i) = \lim_{i\to\infty} \deg(T_i) = \infty,$$

we conclude that $\lim_{i\to\infty} \hat{h}_f(b_i) = \lim_{i\to\infty} \hat{h}_g(d_i) = \infty$. Consequently, X projects dominantly to both the second and fourth coordinates of $\mathbb{A}^4$. Let X' be the curve in $\mathbb{A}^2$ which is the Zariski closure of the image of X under the projection to the second and fourth coordinates.

From Proposition 2.8, we can write

$$U_i = f^{m_i} \circ u_i,\ T_i = g^{n_i} \circ t_i$$

where $m_i, n_i \in \mathbb{N}_0$, u_i and t_i commute with an iterate of f and g, respectively, and $\max\{\deg(u_i), \deg(t_i)\} \le \deg(f) = \deg(g)$. From Proposition 2.8 again, there are only finitely many possibilities for the pair (u_i, t_i). Hence there exist polynomials u and t such that $(u_i, t_i) = (u, t)$ for infinitely many i. Overall, the curve X' in $\mathbb{A}^2$ satisfies the following properties:

- X' is nonfibered.
- $X' \cap (\mathcal{O}_f(u(\alpha)) \times \mathcal{O}_g(t(\beta)))$ is infinite.

By Proposition 4.5, X' is periodic under the map $(x_2, x_4) \mapsto (f(x_2), g(x_4))$. Therefore X is contained in the periodic hypersurface

$$\{(x_1, x_2, x_3, x_4) : (x_2, x_4) \in X'\}$$

and we finish the proof of this case.

4C. ***The case when the ambient space has dimension larger than 4.*** Let $N \ge 5$, assume Theorem 4.1 holds for $n \le N-1$. We now consider $n = N$. Note that the common weak signature $(\mathcal{J}, J_1, \dots, J_{n-2})$ of the V_i's is a partition of $\{1, \dots, n\}$ for which $\mathcal{J}$ could possibly be empty while each J_j is nonempty. Since $2(n-2) > n$, there must be some j such that $|J_j| = 1$. Without loss of generality, assume $J_{n-2} = \{n\}$. We can now proceed as in Case D: if the projection from X to the first $(n-1)$ coordinates is nonconstant then we reduce to $n = N-1$ and apply the induction hypothesis, otherwise we can easily conclude that X is contained in V_i for every i. This finishes the proof of Theorem 4.1. □

5. Proof of Theorem 1.2 for subvarieties of codimension 2

Theorem 1.2 is proven once we deal with the last case of it, which is covered by the following more general result:

Theorem 5.1. *Let $X \subset \mathbb{A}^n$ be an irreducible subvariety of codimension at least equal to* 2. *If $X \cap \mathrm{Per}^{[n-1]}$ is Zariski dense in X, then X must be contained in a proper, periodic, irreducible subvariety of $\mathbb{A}^n$.*

The reason why we can obtain the stronger Theorem 5.1 for the intersection of any subvariety $X \subset \mathbb{A}^n$ of codimension at least equal to 2 with $\mathrm{Per}^{[n-1]}$ is that in this case we intersect X with periodic curves C and this gives a firmer control on the magnitude of the canonical heights for the points from the intersection $X \cap C$. Indeed, we sketch below our approach for the proof of Theorem 5.1. So, assume (for simplicity) that $f_1 = \cdots = f_n =: f$; then for each nonzero integers $k_1, \dots, k_{n-1}$, we let $C_{k_1,\dots,k_{n-1}} \subset \mathbb{A}^n$

be the curve given by the equations

$$x_2 = f^{k_1}(x_1), x_3 = f^{k_2}(x_2), \cdots, x_n = f^{k_{n-1}}(x_{n-1}).$$

Also, assume that X intersects the union of all curves $C_{k_1,\dots,k_{n-1}}$ in a Zariski dense subset. Then, arguing as in the proof of Theorem 4.1, we can assume that the integers k_i are arbitrarily large. This yields that the projection Y of X on the first $(n-1)$ coordinate axes contains a Zariski dense set of points of canonical height tending to 0. Then Theorem 1.7 yields that Y must be preperiodic; also, note that Y is a proper subvariety of $\mathbb{A}^{n-1}$ since the codimension of $X \subset \mathbb{A}^n$ is at least equal to 2. So, using the results of [Medvedev and Scanlon 2014], Y itself must be contained in some hypersurface of $\mathbb{A}^{n-1}$ of the form $C \times \mathbb{A}^{n-3}$, for some preperiodic plane curve C. Then arguing as in the proof of Lemma 4.3, we obtain that C must be periodic and so, X is contained in a proper, periodic, irreducible subvariety of $\mathbb{A}^n$. However, there are extra complications appearing in the proof of Theorem 5.1 compared to the proof of Theorem 4.1 since we cannot reduce our arguments to the case n is small (note that the case $n \geq 5$ reduces to the cases $n = 3, 4$ in the proof of Theorem 4.1); this leads to significant difficulties in showing that the aforementioned curve C is actually *periodic*.

Proof of Theorem 5.1.. Here we are assuming that the intersection between X and the union of all periodic curves is Zariski dense in X and we need to prove that X is contained in a periodic hypersurface of $\mathbb{A}^n$. We argue by induction on n; the case $n = 2$ is trivial while the case $n = 3$ was proven in Theorem 3.1. We assume $n \geq 4$ from now on.

By using Remark 2.3 as in the proof of Theorem 4.1, we can assume that all of the above periodic curves have a common weak signature J_1 which is assumed to be $\{1, \dots, s\}$ where $1 \leq s \leq n$. By Lemma 2.2, Remark 2.3, and Proposition 2.4, we may assume that $f_1 = \cdots = f_s =: f$ and there are periodic curves $\{V_m\}_{m\geq 1}$ (in $\mathbb{A}^n$) such that the following hold:

(a) There is a Zariski dense set of points $\{P_m\}_{m\geq 1}$ in X such that $P_m \in X \cap V_m$ for every m.

(b) Each V_m is defined by equations $x_2 = g_{m,1}(x_1), \dots, x_s = g_{m,s-1}(x_{s-1})$ where the $g_{m,i}$ are polynomials commuting with an iterate of f, along with equations $x_{s+1} = a_{m,s+1}, \dots, x_n = a_{m,n}$ where each $a_{m,i}$ is f_i-periodic for $s+1 \leq i \leq n$.

Write

$$P_m = (b_{m,1}, \dots, b_{m,n}),$$

with $b_{m,j+1} = g_{m,j}(b_{m,j})$ for $1 \leq j \leq s-1$ and $b_{m,j} = a_{m,j}$ for $s+1 \leq j \leq n$.

By restricting to a subsequence, we may assume that $\{P_m\}_{m\geq 1}$ is generic which means that every subsequence is Zariski dense in X. This is possible, as follows. First we enumerate all the countably many strictly proper irreducible $\overline{\mathbb{Q}}$-subvarieties of X as $\{Z_1, Z_2, \dots\}$. Then we let $m_0 := 0$, let P_{m_1} be the first point in the sequence $\{P_m\}_{m>m_0}$ which is not contained in Z_1, let P_{m_2} be the first point in the sequence $\{P_m\}_{m>m_1}$ that is not contained in $Z_1 \cup Z_2$, and so on. The subsequence $\{P_{m_k}\}_{k\geq 1}$ is generic in X.

If for some $i \in \{s+1, \dots, n\}$, the projection from X to the i-th coordinate axis $\mathbb{A}^1$ is constant, then X is contained in the periodic hypersurface $x_i = a_{1,i}$ and we are done.

So, from now on, we may assume that each projection of X on the coordinate axes $x_{s+1}, \dots, x_n$ is not constant.

In particular, this means that for every $i \in \{s+1, \dots, n\}$ and any f_i-periodic point ζ, there are at most finitely many m such that $a_{m,i} = \zeta$; otherwise an infinite subsequence of $\{P_m\}$ is contained in the hypersurface $\{x_i = \zeta\}$. Since $\{P_m\}_m$ is generic, X is also contained in $\{x_i = \zeta\}$, as desired.

Claim 5.2. *Theorem 5.1 holds when $s = 1$.*

Proof. Since $s = 1$, each V_m is of the form

$$\mathbb{A}^1 \times (a_{m,2}, \dots, a_{m,n}).$$

We project X to the last $n-1$ coordinate axes and thus obtain a proper subvariety $X_1 \subset \mathbb{A}^{n-1}$ (note that $X \subset \mathbb{A}^n$ has codimension at least equal to 2). Furthermore, according to our hypothesis, X_1 contains a Zariski dense set of periodic points $(a_{i,2}, \dots, a_{i,n})$; thus Theorem 3.1 yields that X_1 is periodic. Therefore, X is contained in the periodic, proper, irreducible subvariety $\mathbb{A}^1 \times X_1 \subset \mathbb{A}^n$, as desired. □

From now on, we assume $2 \le s \le n$. Furthermore, as argued in the proof of Theorem 4.1, we may assume that for $j = 1, \dots, s-1$, we have $\deg(g_{m,j}) \to \infty$ as $m \to \infty$.

Claim 5.3. *Theorem 5.1 holds if X does not project dominantly onto the s-th coordinate $\mathbb{A}^1$ of $\mathbb{A}^n$.*

Proof. Let b_s be the image of the constant projection from X to the s-th coordinate $\mathbb{A}^1$ and let $\pi_{(s)}$ be the projection from X to the remaining $n-1$ coordinates $\mathbb{A}^{n-1}$. Let $X_{(s)}$ be the Zariski closure of $\pi_{(s)}(X)$.

For each m we have that $V_m \cap X$ contains some point $(b_{m,1}, \dots, b_{m,n})$ such that for $i = 1, \dots, s-1$, we have

$$\hat{h}_f(b_{m,i}) = \frac{\hat{h}_f(b_s)}{\prod_{j=i}^{s-1} \deg(g_{m,j})} \to 0, \qquad \text{as } m \to \infty.$$

Since for $i = s+1, \dots, n$ we have $\hat{h}_f(b_{m,i}) = \hat{h}_f(a_{m,i}) = 0$, we conclude that $X_{(s)}$ contains a Zariski dense set of points of canonical height converging to 0. Thus Theorem 1.7 yields that $X_{(s)}$ is preperiodic. A more careful analysis shows that X is contained in a proper, irreducible, *periodic* subvariety, as follows.

Since $\dim(X_{(s)}) = \dim(X) \le n-2$, we have that $X_{(s)} \subset \mathbb{A}^{n-1}$ is a proper, preperiodic subvariety. By Remark 2.6, there exist $i < j$ in $\{1, \dots, s-1, s+1, \dots, n\}$ and an irreducible curve C in $\mathbb{A}^2$ that is preperiodic under $(x_i, x_j) \mapsto (f_i(x_i), f_j(x_j))$ such that $X_{(s)} \subseteq \pi^{-1}(C)$ where π is the projection to the i-th and j-th coordinate axes, i.e.,

$$(x_1, \dots, x_{s-1}, x_{s+1}, \dots, x_n) \to (x_i, x_j). \tag{5.4}$$

We have several cases (note that the projection from X to each of the ℓ-th coordinate $\mathbb{A}^1$ for $\ell \in \{s+1, \dots, n\}$ is nonconstant):

(i) $i, j \in \{s+1, \dots, n\}$. Then the curve C contains the Zariski dense set of periodic points $\{(a_{m,i}, a_{m,j})\}_m$. By Theorem 3.1, C is periodic. Hence $\pi^{-1}(C)$ is periodic and X is contained in the periodic hypersurface $\pi_{(s)}^{-1}(\pi^{-1}(C))$.

(ii) $i, j \in \{1, \dots, s-1\}$ and the curve C is fibered. Hence there exists an f-preperiodic point γ such that X is contained in the hypersurface $x_i = \gamma$, say. From $b_s = b_{m,s} = g_{m,s-1} \circ \cdots \circ g_{m,i}(\gamma)$ and Corollary 2.10, by choosing sufficiently large m, we have that b_s is f-periodic. Hence X is contained in the periodic hypersurface $\{x_s = b_s\}$.

(iii) $i, j \in \{1, \dots, s-1\}$ and the curve C is nonfibered. By Remark 2.5, C satisfies an equation $g(x_i) = G(x_j)$ where g and G commute with an iterate of f. As in Case A in Section 4 (see Lemma 4.3), we have that $b_{m,j}$ is f-periodic when m is sufficiently large (see Lemma 4.3). Then $b_s = b_{m,s} = g_{m,s-1} \circ \cdots \circ g_{m,j}(b_{m,j})$ is f-periodic and we are done.

(iv) $i \in \{1, \dots, s-1\}$, $j \in \{s+1, \dots, n\}$, and the curve C is fibered. We can use the same arguments as in Case (ii) above since we know C must project dominantly onto the x_j coordinate axis and therefore, we must have that the curve C is given by an equation of the form $x_i = \gamma$, for a preperiodic point γ.

(v) $i \in \{1, \dots, s-1\}$, $j \in \{s+1, \dots, n\}$, and the curve C is nonfibered. Then $f_i \approx f_j$. By Lemma 2.2, we may assume that $f_j = f_i = f$. Now C satisfies an equation $g(x_i) = G(x_j)$ as in Case (iii). Hence $g(b_{m,i}) = G(a_{m,j})$ is f-periodic. By choosing m sufficiently large such that $\deg(g_{m,s-1} \circ \cdots \circ g_{m,i}) \geq \deg(g)$, we conclude that $b_s = b_{m,s} = g_{m,s-1} \circ \cdots \circ g_{m,i}(b_{m,i})$ is periodic.

This finishes the proof of Claim 5.3. ☐

From now on, in the proof of Theorem 5.1 we assume that X projects dominantly onto the s-th axis.

Let $\pi_{(s)}$ and $X_{(s)}$ be as in the proof of Claim 5.2. We still have 2 more cases: $\dim(X_{(s)}) = \dim(X) - 1$ or $\dim(X_{(s)}) = \dim(X)$.

Claim 5.5. *Theorem 5.1 holds if* $\dim(X_{(s)}) = \dim(X) - 1$.

Proof. In this case, we have that $X = X_{(s)} \times \mathbb{A}^1$ (where the factor $\mathbb{A}^1$ comes from the s-th coordinate). Furthermore, by our assumption, we know that $X_{(s)}$ has a Zariski dense intersection with periodic curves of $\mathbb{A}^{n-1}$ given by the equations

$$x_2 = g_{m,1}(x_1), x_3 = g_{m,2}(x_2), \dots, x_{s-1} = g_{m,s-1}(x_{s-2})$$

and the equations

$$x_{s+1} = a_{m,s+1}, x_{s+2} = a_{m,s+2}, \dots, x_n = a_{m,n}.$$

In other words, $X_{(s)}$ has a dense intersection with $\mathrm{Per}^{[n-2]} \subset \mathbb{A}^{n-1}$. By the inductive hypothesis, we conclude that $X_{(s)}$ is contained in a strictly proper periodic subvariety of $\mathbb{A}^{n-1}$, and so is $X \subset \mathbb{A}^n$. ☐

From now on, in the proof of Theorem 5.1 we may assume $\dim(X_{(s)}) = \dim(X)$**.**

Then there is a strictly smaller Zariski closed subset $Y_{(s)}$ of $X_{(s)}$ such that for $Y := \pi^{-1}(Y_{(s)})$, the induced morphism from $X \setminus Y$ to $X_{(s)} \setminus Y_{(s)}$ is finite. At the expense of removing finitely many pairs

(P_m, V_m) for which $P_m \in Y$, we may assume that $P_m \in V_m \cap (X \setminus Y)$ for every m (note that the sequence of points $\{P_m\}$ is generic in X).

Since the map from $X \setminus Y$ to $X_{(s)} \setminus Y_{(s)}$ is finite, by [Ghioca and Nguyen 2016, Corollary 3.4] there are constants $c_0, \dots, c_{s-1}, c_{s+1}, \dots c_n$ such that for each $m \in \mathbb{N}$ we have the inequality

$$\hat{h}_f(b_{m,s}) \le c_0 + \sum_{\substack{1 \le i \le n \\ i \ne s}} c_i \hat{h}_f(b_{m,i}). \tag{5.6}$$

Using the fact that for each $i = 1, \dots, s-1$, we have

$$\hat{h}_f(b_{m,i}) = \frac{\hat{h}_f(b_{m,s})}{\prod_{j=i}^{s-1} \deg(g_{m,j})}, \tag{5.7}$$

while for each $i = s+1, \dots, n$, we have that $\hat{h}_f(b_{m,i}) = \hat{h}_f(a_{m,i}) = 0$. Combining (5.7) with (5.6) and with the fact that $\deg(g_{m,i}) \to \infty$ as $m \to \infty$ for each $i = 1, \dots, s-1$, we conclude that

$$\lim_{m \to \infty} \hat{h}_f(b_{m,i}) = 0, \quad \text{for each } i = 1, \dots, s-1. \tag{5.8}$$

So, $X_{(s)}$ contains a Zariski dense set of points of small height, i.e., the points

$$(b_{m,1}, \dots, b_{m,s-1}, b_{m,s+1}, \dots, b_{m,n}).$$

Then Theorem 1.7 yields that $X_{(s)}$ is preperiodic.

As in the proof of Claim 5.3, there exist $i < j$ in $\{1, \dots, s-1, s+1, \dots, n\}$ and a preperiodic curve C in $\mathbb{A}^2$ such that $X_{(s)}$ is contained in $\pi^{-1}(C)$ where π is the projection to the i-th and j-th coordinate axes, as in (5.4). We have cases (i)–(v) as in the proof of Claim 5.3. Case (i) can be handled by the exact same arguments. On the other hand, cases (ii) and (iv) cannot occur under the hypothesis that X projects dominantly onto the s-th coordinate axis. Indeed, in both those two cases (ii) and (iv) we would have that C is fibered, given by some equation $x_i = \gamma$ (or $x_j = \gamma$) for some i (or j) in $\{1, \dots, s-1\}$ and some preperiodic point γ. But then (without loss of generality) $b_{m,i} = \gamma$ for each m and so,

$$b_{m,s} = (g_{m,s-1} \circ \dots \circ g_{m,i})(b_{m,i}) = (g_{m,s-1} \circ \dots \circ g_{m,i})(\gamma)$$

takes only finitely many values as we vary m by Corollary 2.10. However, the points $\{P_m\}$ are dense in X and X projects dominantly onto the s-th coordinate axis, contradiction. Therefore, we are left to analyze only cases (iii) and (v) appearing in the proof of Claim 5.3.

In cases (iii) and (v), we have that $b_{m,s}$ is periodic when m is large; by removing finitely many m, we may assume that $b_{m,s}$ is periodic for every m. For any $k \in \{1, \dots, s-1\}$, from $b_{m,s} = g_{m,s-1} \circ \dots \circ g_{m,k}(b_{m,k})$, we have that $b_{m,k}$ is f-preperiodic. Therefore, using again that each $b_{m,k} = a_{m,k}$ is periodic for $k > s$, Theorem 1.7 yields that X is preperiodic because it contains a Zariski dense set of preperiodic points. From the discussion in Section 2, we know that X is a product of preperiodic curves. Since $\dim(X) = \dim(X_{(s)})$ and $X_{(s)} \subseteq C \times \mathbb{A}^{n-3}$ (the factor $\mathbb{A}^{n-3}$ comes from all the ℓ-axes where $\ell \in \{1, \dots, n\} \setminus \{i, j, s\}$), we only have two possibilities.

Case F: The first possibility is that $X \subseteq C' \times \mathbb{A}^{n-3}$ where C' is a preperiodic curve in $\mathbb{A}^3$ which is also the projection from X to the i-th, j-th, and s-th axes (hence C is the projection from C' to the i-th and j-th axes $\mathbb{A}^2$). Now in both cases (iii) and (v) from the proof of Claim 5.3, we have that $b_{m,j}$ is periodic for all (sufficiently large) m. Consequently, the projection from X to the j-th axis together with the s-th axis is a curve containing the Zariski dense set of periodic points $(b_{m,j}, b_{m,s})_m$. Therefore this projection is a periodic curve by Theorem 3.1. Hence X lies in the periodic hypersurface which is the inverse image in $\mathbb{A}^n$ of this periodic plane curve under the projection map $(x_1, \ldots, x_n) \mapsto (x_j, x_s)$.

Case G: The second possibility is that there exist $\ell \in \{1, \ldots, n\} \setminus \{i, j, s\}$ such that $X \subseteq C \times C'' \times \mathbb{A}^{n-4}$ where C'' is a preperiodic curve in $\mathbb{A}^2$ which is also the projection from X to the s-th and ℓ-th axes and the factor $\mathbb{A}^{n-4}$ comes from the k-th axes for $k \in \{1, \ldots, n\} \setminus \{i, j, s, \ell\}$. Now if $\ell \in \{s+1, \ldots, n\}$ then we have $b_{m,\ell} = a_{m,\ell}$ is periodic, hence the curve C'' contains the Zariski dense set of periodic points $(b_{m,s}, b_{m,\ell})_m$. From Theorem 3.1, we have that C'' is periodic and we are done since then X is contained in the periodic hypersurface $\mathbb{A}^2 \times C'' \times \mathbb{A}^{n-4}$.

From now on, in the proof of Theorem 5.1 we assume that $\ell \in \{1, \ldots, s\}$.

If the projection from C'' to the ℓ-th coordinate is constant then we derive a contradiction. Indeed, then $x_\ell = \gamma$ where γ is f-preperiodic. From $b_{m,s} = g_{m,s-1} \circ \ldots \circ g_{m,\ell}(\gamma)$, we obtain that the s-th coordinates $b_{m,s}$ of the points P_m must belong to a finite set, contradicting thus the fact that these points are dense in X, which is a variety projecting dominantly onto the s-th coordinate axis.

So, from now on, we may assume that C'' is nonfibered (note that we are already working under the assumption that X projects dominantly onto the s-th coordinate axis).

Therefore C'' satisfies an equation $U(x_s) = T(x_\ell)$ where U and T commute with an iterate of f. It remains to treat case (iii) or case (v) in the proof of Claim 5.3. In either case, we may assume that $f_j = f$ and C satisfies an equation $g(x_i) = G(x_j)$ where g and G commute with an iterate of f. As in the proof of Claim 5.3, we have that $b_{m,j}$ is f-periodic for all large m. Hence both $T(b_{m,\ell}) = U(b_{m,s})$ and $g(b_{m,i}) = G(b_{m,j})$ are f-periodic for all large m.

If $i < \ell$, we have $b_{m,\ell} = g_{m,\ell-1} \circ \ldots \circ g_{m,i}(b_{m,i})$. Therefore when m is large enough so that $\deg(g_{m,\ell-1} \circ \ldots \circ g_{m,i}) \geq \deg(g)$, we have that $b_{m,\ell}$ is periodic (see Lemma 4.3). Consequently, the curve C'' is periodic since it contains a Zariski dense set of periodic points $(b_{m,\ell}, b_{m,s})$. Similarly, if $\ell < i$, when m is large so that

$$\deg(g_{m,i-1} \circ \ldots \circ g_{m,\ell}) \geq \deg(T),$$

we have $b_{m,i}$ is periodic (again using Lemma 4.3), hence C is periodic because it contains a Zariski dense set of periodic points $(b_{m,i}, b_{m,j})$. This finishes the proof of Theorem 5.1. □

Acknowledgements

We thank Tom Tucker for his comments on our paper. We are also grateful to both referees for their very useful suggestions.

References

[Bell et al. 2016] J. P. Bell, D. Ghioca, and T. J. Tucker, *The dynamical Mordell–Lang conjecture*, Mathematical Surveys and Monographs **210**, American Mathematical Society, Providence, RI, 2016. MR Zbl

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs **4**, Cambridge University Press, 2006. MR Zbl

[Bombieri et al. 1999] E. Bombieri, D. Masser, and U. Zannier, "Intersecting a curve with algebraic subgroups of multiplicative groups", *Internat. Math. Res. Notices* 20 (1999), 1119–1140. MR Zbl

[Bombieri et al. 2006] E. Bombieri, D. Masser, and U. Zannier, "Intersecting curves and algebraic subgroups: conjectures and more results", *Trans. Amer. Math. Soc.* **358**:5 (2006), 2247–2257. MR Zbl

[Bombieri et al. 2007] E. Bombieri, D. Masser, and U. Zannier, "Anomalous subvarieties—structure theorems and applications", *Int. Math. Res. Not.* **2007**:19 (2007), Art. ID rnm057, 33. MR Zbl

[DeMarco et al. 2017] L. DeMarco, D. Ghioca, H. Krieger, K. D. Nguyen, T. J. Tucker, and H. Ye, "Bounded height in families of dynamical systems", preprint, 2017. To appear in *Int. Math. Res. Not.* arXiv

[Ghioca and Nguyen 2016] D. Ghioca and K. D. Nguyen, "Dynamical anomalous subvarieties: structure and bounded height theorems", *Adv. Math.* **288** (2016), 1433–1462. MR Zbl

[Ghioca and Nguyen 2017] D. Ghioca and K. D. Nguyen, "Dynamics of split polynomial maps: uniform bounds for periods and applications", *Int. Math. Res. Not.* **2017**:1 (2017), 213–231. MR

[Ghioca et al. 2011] D. Ghioca, T. J. Tucker, and S. Zhang, "Towards a dynamical Manin–Mumford conjecture", *Int. Math. Res. Not.* **2011**:22 (2011), 5109–5122. MR Zbl

[Ghioca et al. 2015] D. Ghioca, K. D. Nguyen, and H. Ye, "The Dynamical Manin–Mumford Conjecture and the Dynamical Bogomolov Conjecture for split rational maps", preprint, 2015. To appear in *J. Eur. Math. Soc.* (*JEMS*). arXiv

[Ghioca et al. 2018] D. Ghioca, K. D. Nguyen, and H. Ye, "The dynamical Manin–Mumford conjecture and the dynamical Bogomolov conjecture for endomorphisms of $(\mathbb{P}^1)^n$", *Compos. Math.* **154**:7 (2018), 1441–1472. MR

[Medvedev and Scanlon 2014] A. Medvedev and T. Scanlon, "Invariant varieties for polynomial dynamical systems", *Ann. of Math.* (2) **179**:1 (2014), 81–177. MR Zbl

[Nguyen 2015] K. Nguyen, "Some arithmetic dynamics of diagonally split polynomial maps", *Int. Math. Res. Not.* **2015**:5 (2015), 1159–1199. MR Zbl

[Pink $\geq$ 2018] R. Pink, "A common generalization of the conjectures of André–Oort, Manin–Mumford, and Mordell-Lang", preprint, Available at https://people.math.ethz.ch/~pink/ftp/AOMMML.pdf.

[Silverman 2007] J. H. Silverman, *The arithmetic of dynamical systems*, Graduate Texts in Mathematics **241**, Springer, 2007. MR Zbl

[Xie 2017] J. Xie, "The dynamical Mordell–Lang conjecture for polynomial endomorphisms of the affine plane", pp. vi+110 in *Journées de Géométrie Algébrique d'Orsay*, Astérisque **394**, Société Mathématique de France, Paris, 2017. MR Zbl

[Zannier 2012] U. Zannier, *Some problems of unlikely intersections in arithmetic and geometry*, Annals of Mathematics Studies **181**, Princeton University Press, 2012. MR Zbl

[Zhang 2006] S.-W. Zhang, "Distributions in algebraic dynamics", pp. 381–430 in *Surveys in differential geometry, Vol. X*, edited by S. T. Yau, Surv. Differ. Geom. **10**, International Press, Somerville, MA, 2006. MR Zbl

[Zilber 2002] B. Zilber, "Exponential sums equations and the Schanuel conjecture", *J. London Math. Soc.* (2) **65**:1 (2002), 27–44. MR Zbl

Communicated by Shou-Wu Zhang
Received 2017-11-01 Revised 2018-04-06 Accepted 2018-06-06

dghioca@math.ubc.ca Department of Mathematics, University of British Columbia, Vancouver BC, Canada

dangkhoa.nguyen@ucalgary.ca Department of Mathematics and Statistics, The University of Calgary, Calgary AB, Canada

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1773

Homogeneous length functions on groups

D. H. J. Polymath

A pseudolength function defined on an arbitrary group $G = (G, \cdot, e, ()^{-1})$ is a map $\ell : G \to [0, +\infty)$ obeying $\ell(e) = 0$, the symmetry property $\ell(x^{-1}) = \ell(x)$, and the triangle inequality $\ell(xy) \le \ell(x) + \ell(y)$ for all $x, y \in G$. We consider pseudolength functions which saturate the triangle inequality whenever $x = y$, or equivalently those that are *homogeneous* in the sense that $\ell(x^n) = n\ell(x)$ for all $n \in \mathbb{N}$. We show that this implies that $\ell([x, y]) = 0$ for all $x, y \in G$. This leads to a classification of such pseudolength functions as pullbacks from embeddings into a Banach space. We also obtain a quantitative version of our main result which allows for defects in the triangle inequality or the homogeneity property.

1. Introduction

Let $G = (G, \cdot, e, ()^{-1})$ be a group (written multiplicatively, with identity element e). A *pseudolength function* on G is a map $\ell : G \to [0, +\infty)$ that obeys the properties

- $\ell(e) = 0$,
- $\ell(x^{-1}) = \ell(x)$,
- $\ell(xy) \le \ell(x) + \ell(y)$

for all $x, y \in G$. If in addition we have $\ell(x) > 0$ for all $x \in G \setminus \{e\}$, we say that ℓ is a *length function*. By setting $d(x, y) := \ell(x^{-1}y)$, it is easy to see that pseudolength and length functions are in bijection with left-invariant pseudometrics and metrics on G, respectively.

From the above properties it is clear that one has the upper bound

$$\ell(x^n) \le |n|\ell(x)$$

for all $x \in G$ and $n \in \mathbb{Z}$. Let us say that a pseudolength function $\ell : G \to [0, +\infty)$ is *homogeneous* if equality is always attained here, in that one has

$$\ell(x^n) = |n|\ell(x) \tag{1.1}$$

for all $x \in G$ and any $n \in \mathbb{Z}$. Using the axioms of a pseudolength function, it is not difficult to show that the homogeneity condition (1.1) is equivalent to the triangle inequality holding with equality whenever $x = y$ (i.e., that (1.1) holds for $n = 2$); see [Gajda and Kominek 1991, Lemma 1].

MSC2010: primary 20F12; secondary 20F65.
Keywords: homogeneous length function, pseudolength function, quasimorphism, Banach space embedding.

If one has a real or complex Banach space $\mathbb{B} = (\mathbb{B}, \|\ \|)$, and $\phi : G \to \mathbb{B}$ is any homomorphism from G to $\mathbb{B}$ (viewing the latter as a group in additive notation), then the function $\ell : G \to [0, +\infty)$ defined by $\ell(x) := \|\phi(x)\|$ is easily verified to be a homogeneous pseudolength function. Furthermore, if ϕ is injective, then ℓ is in fact a homogeneous length function. For instance, the function $\ell((n, m)) := |n + \sqrt{2}m|$ is a length function on $\mathbb{Z}^2$, where in this case $\mathbb{B} := \mathbb{R}$ and $\phi((n, m)) := n + \sqrt{2}m$. On the other hand, one can easily locate many length functions that are not homogeneous, for instance by taking the square root of the length function just constructed.

The main result of this paper is that such Banach space constructions are in fact the *only* way to generate homogeneous (pseudo-)length functions.

Theorem 1.2 (classification of homogeneous length functions). *Given a group G, let $\ell : G \to [0, +\infty)$ be a homogeneous pseudolength function. Then there exist a real Banach space $\mathbb{B} = (\mathbb{B}, \|\ \|)$ and a group homomorphism $\phi : G \to \mathbb{B}$ such that $\ell(x) = \|\phi(x)\|$ for all $x \in G$. Furthermore, if ℓ is a length function, one can take ϕ to be injective, i.e., an isometric embedding.*

We derive Theorem 1.2 from a more quantitative result bounding the pseudolength of a commutator

$$[x, y] := xyx^{-1}y^{-1}; \tag{1.3}$$

see Proposition 2.1 below. Our arguments are elementary, relying on directly applying the axioms of a homogeneous length function to various carefully chosen words in x and y, and repeatedly taking an asymptotic limit $n \to \infty$ to dispose of error terms that arise in the estimates obtained in this fashion.

An additional advantage of quantifying Theorem 1.2 in Proposition 2.1 is that one can derive from the latter proposition a "quasified" version of Theorem 1.2. See Theorem 4.4 below.[1]

Finally, as one quick corollary of Theorem 1.2, we obtain the following characterization of the groups that admit homogeneous length functions.

Corollary 1.4. *A group admits a homogeneous length function if and only if it is abelian and torsion-free.*

Examples and approaches. We now make some remarks to indicate the nontriviality of Theorem 1.2. Corollary 1.4 implies that there are no nonabelian groups with homogeneous length functions. Whether or not such a striking geometric rigidity phenomenon holds was previously unknown to experts. Moreover, the corollary fails to hold if one or more of the precise conditions in the theorem are weakened. For instance, such length functions indeed exist (i) on nonabelian monoids, and (ii) on balls of finite radius in free groups. We explain these two cases further in Section 4.

Given these cases, one could *a priori* ask if every nonabelian group admits a homogeneous length function. This is not hard to disprove; here are two examples.

[1] A different variant of Theorem 1.2 involves replacing homogeneity by the assumption that ℓ is a pseudolength function on G whose homogenization is positive:

$$\ell_{\text{hom}}(g) := \lim_{n\to\infty} \frac{\ell(g^n)}{n} > 0, \quad \forall g \neq e.$$

(This was studied in [Niemiec 2013, Theorem 2.10(III)] in the special case of abelian (G, ℓ).) In this case we work with (G, ℓ_{hom}) instead of ℓ, to conclude that G maps into a Banach space.

Example 1.5 (nilpotent groups). If G is nilpotent of nilpotency class two (e.g., the Heisenberg group), then $[x, y]^{n^2} = [x^n, y^n]$ for all $x, y \in G$ and integers $n \geq 0$ since the map $(g, h) \mapsto [g, h]$ is now a bihomomorphism $G \times G \to [G, G] \subset Z(G)$. If $[x, y]$ is nontrivial, then any homogeneous length function on G would assign a linearly growing quantity to the right-hand side and a quadratically growing quantity to the left-hand side, which is absurd; thus such groups cannot admit homogeneous length functions. The claim then also follows for nilpotent groups of higher nilpotency class, since they contain subgroups of nilpotency class two.[2]

Example 1.6 (connected Lie groups). As we explain in Remark 2.9, a homogeneous length function ℓ induces a *biinvariant* metric on G. Now if (G, ℓ) is furthermore a connected Lie group, then by [Milnor 1976, Lemma 7.5], $G \cong K \times \mathbb{R}^n$ for some compact Lie group K and integer $n \geq 0$. By (1.1), K cannot have torsion elements, hence must be trivial. But then G is abelian.

Prior to Corollary 1.4, the above examples left open the question of whether *any* nonabelian group admits a homogeneous length function. One may as well consider groups generated by two noncommuting elements. As a prototypical example, let $\boldsymbol{F}_2$ be the free group on two generators a and b. The word length function on $\boldsymbol{F}_2$ is a length function, but it is not homogeneous, since for instance the word length of $(bab^{-1})^n = ba^n b^{-1}$ is $n+2$, which is not a linear function of n. It is however the case that the word length of x^n has linear *growth* in n for any nontrivial x. Similarly for the Levenshtein distance (edit distance) on $\boldsymbol{F}_2$.

Our initial attempts to construct homogeneous length functions on $\boldsymbol{F}_2$ all failed. Of course, this failure is explained by our main result. However, many of these methods apply under minor weakening of the hypotheses, such as working with monoids rather than groups, or weakening homogeneity. Results in these cases are discussed further in Section 4.

Further motivations. We next mention some motivations from functional analysis and probability, or more precisely the study of Banach space embeddings. If G is an additive subgroup of a Banach space $\mathbb{B}$, then clearly the norm on $\mathbb{B}$ restricts to a homogeneous length function on G. In [Cabello Sánchez and Castillo 2002; Gajda and Kominek 1991] one can find several equivalent conditions for a given length function on a given group to arise in this way (studied in the broader context of additive mappings and separation theorems in functional analysis); see also [Niemiec 2013, Theorem 2.10(II)] for an alternative proof. These conditions are summarized in [Khare and Rajaratnam 2016]. For instance, given a group G with a length function ℓ, there exists an isometric embedding from G to a Banach space $\mathbb{B}$ with ℓ induced from the metric on $\mathbb{B}$, if and only if G is amenable and $\ell(x^2) = 2\ell(x)$ for all x.

In view of such equivalences, it is natural to try to characterize the groups possessing a homogeneous length function. This characterization is given in Corollary 1.4, which shows these are precisely the abelian torsion-free groups.

[2] One can also show by relatively simple means that solvable nonabelian groups cannot admit homogeneous length functions either; see the discussion on lamplighter groups in the comments to terrytao.wordpress.com/2017/12/16/.

Groups and semigroups with translation-invariant metrics also naturally arise in probability theory, with the most important "normed" (i.e., homogeneous) examples being Banach spaces [Ledoux and Talagrand 1991]. Notice however that in certain fundamental stochastic settings, formulating and proving results does not require the full Banach space structure. In this vein, a general variant of the Hoffmann-Jørgensen inequality was shown in [Khare and Rajaratnam 2017] in arbitrary metric semigroups — including Banach spaces as well as (nonabelian) compact Lie groups. Similarly in [Khare and Rajaratnam 2016], the authors transferred the (sharp) Khinchin–Kahane inequality from Banach spaces to abelian groups G equipped with a homogeneous length function. To explore extensions of these results to the nonabelian setting (e.g., Lie groups with left-invariant metrics), we need to first understand if such objects exist. As explained above, this question was not answered in the literature; but it is now settled by our main result.

Finally, there may also be a relation to the Ribe program [Naor 2012], which aims to reformulate aspects of Banach space theory in purely metric terms. Indeed, from Corollary 1.4 we see that a metric space X is isometric to an additive subgroup of a Banach space if and only if there is a group structure on X which makes the metric left-invariant and the length function $\ell(x) := d(1, x)$ homogeneous.

2. Key proposition

The key proposition used to prove Theorem 1.2 is the following estimate, which can treat a somewhat more general class of functions than homogeneous pseudolength functions, in which the symmetry hypothesis is dropped and one allows for an error in the homogeneity property, which is now also only claimed for $n = 2$.

Proposition 2.1. *Let $G = (G, \cdot)$ be a group, let $c \in \mathbb{R}$, and let $\ell : G \to \mathbb{R}$ be a function obeying the following axioms*:

(i) *For any $x, y \in G$, one has*

$$\ell(xy) \leq \ell(x) + \ell(y). \tag{2.2}$$

(ii) *For any $x \in G$, one has*

$$\ell(x^2) \geq 2\ell(x) - c. \tag{2.3}$$

Then for any $x, y \in G$, one has

$$\ell([x, y]) \leq 5c, \tag{2.4}$$

where the commutator $[x, y]$ was defined in (1.3).

Notably, we neither assume symmetry $\ell(x^{-1}) = \ell(x)$, not even up to a constant, nor $\ell(e) = 0$ (although $0 \leq \ell(e) \leq c$ follows from the axioms); we also allow ℓ to take on negative values. The reader may however wish to restrict attention to homogeneous length functions, and set $c = 0$ and $\ell \geq 0$ for a first reading of the arguments below. The factor of 5 is probably not optimal here, but the crucial feature of the bound (2.4) for our main application is that the right-hand side vanishes when $c = 0$ (the right-hand side is also independent of x and y, which we use in other applications).

We define a *semilength function* to be a function $\ell : G \to \mathbb{R}$ such that for all $x, y \in G$, $\ell(xy) \le \ell(x) + \ell(y)$, i.e., ℓ satisfies (2.2). Every pseudolength function is a semilength function. A semilength function that satisfies (2.3) for some $c \in \mathbb{R}$ is called *quasihomogeneous*.

Remark 2.5. Suppose $\ell : G \to \mathbb{R}$ and there is a constant k such that $\ell(xy) \le \ell(x) + \ell(y) + k$ for all $x, y \in G$. Then the function $\ell'(x) := \ell(x) + k$ is a semilength function. Further, ℓ' satisfies (2.3) with c replaced by $c' := k + c$, whenever ℓ satisfies (2.3) on the nose. Thus Proposition 2.1 continues to hold if (2.2) is replaced by the condition $\ell(xy) \le \ell(x) + \ell(y) + k$ for all $x, y \in G$, with the bound in the conclusion (2.4) becoming $5c + 4k$.

We now turn to the proof. For the remainder of this section, let G, c, and ℓ satisfy the hypotheses of the proposition. Our task is to establish the bound (2.4). We shall now use (2.2) and (2.3) repeatedly to establish a number of further inequalities relating the semilengths $\ell(x)$ of various elements x of G, culminating in (2.4). Many of our inequalities will involve terms that depend on an auxiliary parameter n, but we will be able to eliminate several of them by the device of passing to the limit $n \to \infty$. It is because of this device that we are able to obtain a bound (2.4) whose right-hand side is completely uniform in x and y.

From (2.2) and induction we have the upper homogeneity bound

$$\ell(x^n) \le n\ell(x) \tag{2.6}$$

for any natural number $n \ge 1$. Similarly, from (2.3) and induction one has the lower homogeneity bound

$$\ell(x^n) \ge n\ell(x) - \log_2(n)c \ge n\ell(x) - nc$$

whenever n is a power of two. It is convenient to rearrange this latter inequality as

$$\ell(x) \le \frac{\ell(x^n)}{n} + c. \tag{2.7}$$

This inequality, particularly in the asymptotic limit $n \to \infty$, will be the principal means by which the hypothesis (2.3) is employed.

We remark that by further use of (2.6) one can also obtain a similar estimate to (2.7) for natural numbers n that are not powers of two, but the powers of two will suffice for the arguments that follow.

Lemma 2.8 (approximate conjugation invariance). *For any $x, y \in G$, one has*

$$\ell(yxy^{-1}) \le \ell(x) + c.$$

Remark 2.9. Setting $c = 0$, we conclude that any homogeneous pseudolength function is conjugation invariant, and thus determines a biinvariant metric on G. It should not be surprising that this observation is used in the proof of Theorem 1.2, since it is a simple consequence of that theorem.

Proof of Lemma 2.8. From (2.7) with x replaced by yxy^{-1}, one has

$$\ell(yxy^{-1}) \le \frac{\ell(yx^n y^{-1})}{n} + c$$

whenever n is a power of two. On the other hand, from (2.6) and (2.2) one has

$$\ell(yx^n y^{-1}) \leq \ell(y) + n\ell(x) + \ell(y^{-1})$$

and thus

$$\ell(yxy^{-1}) \leq \ell(x) + c + \frac{\ell(y) + \ell(y^{-1}) - c}{n}.$$

Sending $n \to \infty$, we obtain the claim. ☐

Lemma 2.10 (splitting lemma). *Let $x, y, z, w \in G$ be such that x is conjugate to both wy and zw^{-1}. Then one has*

$$\ell(x) \leq \tfrac{1}{2}(\ell(y) + \ell(z)) + \tfrac{3}{2}c. \tag{2.11}$$

Proof. If we write $x = swys^{-1} = tzw^{-1}t^{-1}$ for some $s, t \in G$, then from (2.7) we have

$$\ell(x) \leq \frac{\ell(x^n x^n)}{2n} + c = \frac{\ell(s(wy)^n s^{-1} t(zw^{-1})^n t^{-1})}{2n} + c$$

whenever n is a power of two. From Lemma 2.8 and (2.2) one has

$$\begin{aligned}\ell((wy)^{k+1} s^{-1} t(zw^{-1})^{k+1}) &= \ell(wy(wy)^k s^{-1} t(zw^{-1})^k zw^{-1}) \\ &\leq \ell(y(wy)^k s^{-1} t(zw^{-1})^k z) + c \\ &\leq \ell((wy)^k s^{-1} t(zw^{-1})^k) + \ell(y) + \ell(z) + c\end{aligned}$$

for any $k \geq 0$, and hence by induction

$$\ell((wy)^n s^{-1} t(zw^{-1})^n) \leq \ell(s^{-1}t) + n(\ell(y) + \ell(z) + c).$$

Inserting this into the previous bound for $\ell(x)$ via two applications of (2.2), we conclude that

$$\ell(x) \leq \frac{\ell(y) + \ell(z) + c}{2} + \frac{\ell(s) + \ell(s^{-1}t) + \ell(t^{-1})}{2n} + c;$$

sending $n \to \infty$, we obtain the claim. ☐

Corollary 2.12. *If $x, y \in G$, let $f = f_{x,y} : \mathbb{Z}^2 \to \mathbb{R}$ denote the function*

$$f(m, k) := \ell(x^m [x, y]^k).$$

Then for any $m, k \in \mathbb{Z}$, we have

$$f(m, k) \leq \frac{f(m-1, k) + f(m+1, k-1)}{2} + 2c. \tag{2.13}$$

Proof. Observe that $x^m[x, y]^k$ is conjugate to both $x(x^{m-1}[x, y]^k)$ and to $(y^{-1}x^m[x, y]^{k-1}xy)x^{-1}$, hence by (2.11) one has

$$\ell(x^m[x, y]^k) \leq \tfrac{1}{2}\big[\ell(x^{m-1}[x, y]^k) + \ell(y^{-1}x^m[x, y]^{k-1}xy)\big] + \tfrac{3}{2}c.$$

Since $y^{-1}x^m[x, y]^{k-1}xy$ is conjugate to $x^{m+1}[x, y]^{k-1}$, the claim now follows from Lemma 2.8. ☐

We now prove Proposition 2.1. Let $x, y \in G$. We can write the inequality (2.13) in probabilistic form as

$$f(m,k) \le \boldsymbol{E} f\big((m, k-\tfrac{1}{2}) + Y(1, -\tfrac{1}{2})\big) + 2c$$

where $Y = \pm 1$ is a Bernoulli random variable that equals 1 or -1 with equal probability. The key point here is the drift of $(0, -\frac{1}{2})$ in the right-hand side. Iterating this inequality, we see that

$$f(0,n) \le \boldsymbol{E} f\big((Y_1 + \cdots + Y_{2n})(1, -\tfrac{1}{2})\big) + 4cn,$$

where $n \ge 0$ and $Y_1, \dots, Y_{2n}$ are independent copies of Y (so in particular $Y_1 + \cdots + Y_{2n}$ is an even integer).

From (2.2) and (2.6) one has the inequality

$$f(m,k) \le |m|(\max(\ell(x), \ell(x^{-1}))) + |k|(\max(\ell([x,y]), \ell([x,y]^{-1}))) + \ell(e)$$

for all integers m and k, where the final term $\ell(e)$ is used when $m = k = 0$, but can also be added in the remaining cases since it is nonnegative. We conclude that

$$f\big((Y_1 + \cdots + Y_{2n})(1, -\tfrac{1}{2})\big) \le A|Y_1 + \cdots + Y_{2n}| + \ell(e)$$

where A is a quantity independent of n; more explicitly, one can take

$$A := \max(\ell(x), \ell(x^{-1})) + \tfrac{1}{2}\max(\ell([x,y]), \ell([x,y]^{-1})).$$

Taking expectations, since the random variable $Y_1 + \cdots + Y_{2n}$ has mean zero and variance $2n$, we see from the Cauchy–Schwarz inequality or Jensen's inequality that

$$\boldsymbol{E}|Y_1 + \cdots + Y_{2n}| \le (\boldsymbol{E}|Y_1 + \cdots + Y_{2n}|^2)^{1/2} = \sqrt{2n}$$

and hence

$$f(0,n) \le A\sqrt{2n} + \ell(e) + 4cn.$$

But from (2.7), if n is a power of 2 then we have

$$\ell([x,y]) \le \frac{f(0,n)}{n} + c.$$

Combining these two bounds and sending $n \to \infty$, we obtain Proposition 2.1. □

Remark 2.14. One can deduce a "local" version of Proposition 2.1 as follows: notice that the constant c can be described in terms of ℓ from (2.3), to yield

$$\ell([x,y]) \le 5 \sup_{z \in G}(2\ell(z) - \ell(z^2)) \tag{2.15}$$

for any group G and function $\ell : G \to \mathbb{R}$ for which this supremum exists, and any $x, y \in G$. (Both sides are zero when G is a Banach space and ℓ is the norm, so equality is obtained in that case.) It is also enough to consider the supremum over the subgroup of G generated by x and y without loss of generality, which may lead to a better bound on $\ell([x, y])$ than taking the supremum over all of G. Notice also that the constant c must be nonnegative, from (2.3) and (2.2) with $x = y = e$:

$$c \geq 2\ell(e) - \ell(e^2) = \ell(e) \geq \ell(e^2) - \ell(e) = 0. \tag{2.16}$$

In fact, this reasoning and our results imply that the only way to get $c = 0$ on the right-hand side of (2.4) is when ℓ arises from pulling back the norm of a Banach space $\mathbb{B}$ along a group homomorphism $G \to \mathbb{B}$, or equivalently along a group homomorphism from the torsion-free abelianization of G to $\mathbb{B}$.

3. Completing the proof of Theorem 1.2

With Proposition 2.1 in hand, it is not difficult to conclude the proof of Theorem 1.2. Suppose that G is a group with a homogeneous semilength function $\ell : G \to [0, +\infty)$. Applying Proposition 2.1 with $c = 0$, we conclude that $\ell([x, y]) = 0$ for all $x, y \in G$, thus by the triangle inequality ℓ vanishes on the commutator subgroup $[G, G]$, and therefore factors through the abelianization $G_{\mathrm{ab}} := G/[G, G]$ of G. Observe that this already establishes part of one implication of Corollary 1.4. Factoring out by $[G, G]$ like this, we may now assume without loss of generality that G is abelian. To reflect this, we now use additive notation for G, thus for instance $\ell(nx) = |n|\ell(x)$ for each $x \in G$ and $n \in \mathbb{Z}$, and one can also view G as a module over the integers $\mathbb{Z}$.

At this point we repeat the arguments in [Khare and Rajaratnam 2016, Theorem B], which treated the case when G was separable, though it turns out that this separability hypothesis is unnecessary.

If x is a torsion element of G, i.e., $nx = 0$ for some n, then the homogeneity condition forces $\ell(x) = 0$. Thus ℓ vanishes on the torsion subgroup of G; factoring out by this subgroup, we may thus assume without loss of generality that G is not only abelian, but is also torsion-free.

We can view G as a subgroup of the $\mathbb{Q}$-vector space $G \otimes_{\mathbb{Z}} \mathbb{Q}$, the elements of which can be formally expressed as $\frac{1}{n}x$ for natural numbers n and elements $x \in G$ (with two such expressions $\frac{1}{n}x$, $\frac{1}{m}y$ identified if and only if $mx = ny$, and the $\mathbb{Q}$-vector space operations defined in the obvious fashion); the fact that this is well defined as a $\mathbb{Q}$-vector space follows from the hypotheses that G is abelian and torsion-free. We can then define the map $\|\ \|_{\mathbb{Q}} : G \otimes_{\mathbb{Z}} \mathbb{Q} \to [0, +\infty)$ by setting

$$\left\|\tfrac{1}{n}x\right\|_{\mathbb{Q}} := \tfrac{1}{n}\ell(x)$$

for any $x \in G$ and natural number n; the linear growth condition ensures that $\|\ \|_{\mathbb{Q}}$ is well defined. It is not difficult to verify that $\|\ \|_{\mathbb{Q}}$ is indeed a seminorm over the $\mathbb{Q}$-vector space $G \otimes_{\mathbb{Z}} \mathbb{Q}$.

The norm $\|\ \|_{\mathbb{Q}}$ on $G \otimes_{\mathbb{Z}} \mathbb{Q}$ gives a metric $d(x, y) = \|x - y\|_{\mathbb{Q}}$. Consider the metric completion $\mathbb{B}$ of $G \otimes_{\mathbb{Z}} \mathbb{Q}$ with this metric. It is easy to see that the $\mathbb{Q}$-vector space structure on $G \otimes_{\mathbb{Z}} \mathbb{Q}$ extends to an $\mathbb{R}$-vector space structure on $\mathbb{B}$, and the norm $\|\ \|_{\mathbb{Q}}$ on $G \otimes_{\mathbb{Z}} \mathbb{Q}$ extends to a norm $\|\ \|_{\mathbb{R}}$ on $\mathbb{B}$. As $\mathbb{B}$ is complete by construction, it is a Banach space. The inclusion of G in $G \otimes_{\mathbb{Z}} \mathbb{Q}$ gives a homomorphism $\phi : G \to \mathbb{B}$ as required.

This concludes the proof of Theorem 1.2. Since the homomorphism $\phi : G \to \mathbb{B}$ can only be injective for abelian torsion-free G, we obtain the "only if" portion of Corollary 1.4. Conversely, if a group G is abelian and torsion-free, by the above constructions it embeds into a real vector space $\mathbb{B} := G \otimes_{\mathbb{Z}} \mathbb{R}$; now by Zorn's lemma $\mathbb{B}$ has a norm (e.g., consider the ℓ^1 norm with respect to a Hamel basis of $\mathbb{B}$), which restricts to the desired homogeneous length function on G. We remark that $G \otimes_{\mathbb{Z}} \mathbb{R}$ is the construction of the smallest, "enveloping" vector space containing a copy of the abelian, torsion-free group G.

Remark 3.1. The above arguments also show that homogeneous pseudolength functions on G are in bijection with seminorms on the real vector space $G_{\mathrm{ab},0} \otimes_{\mathbb{Z}} \mathbb{R}$, where $G_{\mathrm{ab},0}$ denotes the torsion-free abelianization of G.

4. Further remarks and results

If we weaken any of several conditions in Corollary 1.4, then examples of nonabelian structures with generalized length functions do, in fact, often exist. However, the generality of Proposition 2.1 allows us to obtain nontrivial information in some of these cases. Here we mention several such cases and discuss other related problems.

Monoids and embeddings. Our first weakening is to replace "groups" by the more primitive structures "monoids" or "semigroups". In this case, Robert Young (private communication) described to us nonabelian monoids with homogeneous, biinvariant length functions: consider the free monoid $\mathrm{FMon}(X)$ on any alphabet X of size at least 2, with the edit distance $d(v, w)$ between strings $v, w \in \mathrm{FMon}(X)$ being the least number of single generator insertions and deletions to get from v to w. The triangle inequality and positivity are easily verified, while homogeneity of the corresponding length function $\ell(x) := d(e, x)$ is trivial. Moreover, the metric $d(\cdot\,,\cdot)$ turns out to be biinvariant:

$$d(gxh, gyh) = d(x, y) \quad \text{for all } g, h, x, y \in \mathrm{FMon}(X).$$

This specializes to left- and right-invariance upon taking $g \in X$ and $h = e$, or $h \in X$ and $g = e$, respectively.

Note moreover that $\mathrm{FMon}(X)$ embeds into the free group $\mathrm{FGp}(X)$ generated by X and X^{-1}, where X^{-1} is the collection of symbols defined to be inverses of elements of X. In particular, $\mathrm{FMon}(X)$ is cancellative. While this trivially addresses the embeddability issue, notice that a more refined version of embeddability fails. Namely, by our main theorem, $\mathrm{FMon}(X)$ does not embed into any group in the category $\mathcal{C}_{\mathrm{biinv,hom}}$ with cancellative semigroups with homogeneous biinvariant metrics as objects and isometric semigroup maps as morphisms. Thus, one may reasonably ask what is a sufficiently small category in which the embeddability works. The following proposition shows that we just need to drop homogeneity.

Proposition 4.1. *Let $\mathcal{C}_{\mathrm{biinv}}$ denote the category whose objects are cancellative semigroups with biinvariant metrics, and morphisms are isometric semigroup maps. Then* $\mathrm{FMon}(X)$ *embeds isometrically into* $\mathrm{FGp}(X)$ *in $\mathcal{C}_{\mathrm{biinv}}$.*

Proof. From above, $\mathrm{FMon}(X)$ is an object of $\mathcal{C}_{\mathrm{biinv}}$; denote the metric by d_{FM}. One can check that $d_{FM}(w, w')$ equals the difference between $\ell(w) + \ell(w')$ and twice the length of the longest common (possibly noncontiguous) substring in w and w'; here, ℓ denotes the length of a word in the alphabet X.

We next claim $\mathrm{FGp}(X)$ is also an object of $\mathcal{C}_{\mathrm{biinv}}$. Namely, for a word $w = x_1 x_2 \cdots x_m$ in the free group, we consider *noncrossing matchings* in w, i.e., sets M of pairs of letters in $\{1, 2, \ldots m\}$ such that the following hold:

- If $(i, j) \in M$, then $i < j$ and $x_j = x_i^{-1}$.
- If $(i, j), (k, l) \in M$, then either $(i, j) = (k, l)$ or i, j, k, l are distinct.
- If $i < k < j < l$ and $(i, j) \in M$, then $(k, l) \notin M$.

Given a matching M as above, consider the set $U = U(M)$ of indices k, $1 \leq k \leq m$, which are not part of a pair in M. Define the *deficiency* of the matching M as the cardinality of the set $U(M)$, and define the length $\ell_{\mathrm{wc}}(w)$ of the word w as the infimum of the deficiency over all noncrossing matchings in w (the subscript in ℓ_{wc} stands for Watson–Crick). This length was previously studied in [Gadgil 2009], including checking that it is well defined on all of $\mathrm{FGp}(X)$; moreover, $\ell_{\mathrm{wc}}(w)$ equals the smallest number of conjugates of elements in $X \sqcup X^{-1}$ whose product is w. Now define $d_{FG}(w, w') := \ell_{\mathrm{wc}}(w^{-1}w')$. It is easy to see that ℓ_{wc} is a conjugacy invariant length function.

We claim that $d_{FG} \equiv d_{FM}$ on $\mathrm{FMon}(X)$, which proves the result. It is easy to show that if two words in $\mathrm{FMon}(X)$ differ by a single insertion or deletion, then their distance in $\mathrm{FGp}(X)$ is at most one, hence exactly one. In the other direction, we claim that a noncrossing matching on $w^{-1}w'$, with w and w' containing only positive generators (in X), is just a "rainbow", i.e., nested arches with one end in w^{-1} and the other in w'. But then $d_{FG}(w, w')$ equals $\ell(w) + \ell(w')$ minus twice the length of a common substring, which is maximal by the minimality of the deficiency. Hence $d_{FG}(w, w') = d_{FM}(w, w')$, completing the proof. □

Note that given weights $\ell(a)$ and $\ell(b)$, there is a natural weighted version $\ell_{\mathrm{wc};a,b}$ where the letters of U as above are taken with these weights (symmetrically under inversion). This corresponds to the weighted edit distance, with different costs for editing different letters.

Quasimorphisms and commutator lengths. We now investigate potential applications of Proposition 2.1 with $c > 0$. A *quasimorphism* on a group G is a map $f : G \to \mathbb{R}$ whose *defect* is bounded,

$$D(f) := \sup_{x,y \in G} |f(xy) - f(x) - f(y)| < +\infty.$$

Every quasimorphism induces a pseudolength function (in particular a semilength function) by setting

$$\ell(x) := |f(x)| + D(f), \tag{4.2}$$

where we can take $c = 2D(f)$ as a bound on the homogeneity defect. In this case, Proposition 2.1 makes a rather trivial statement: a homogeneous quasimorphism is bounded on commutators,

$$|f([x, y])| \leq 10D(f).$$

In fact, as observed in [Bavard 1991, Lemme 1.1], for homogeneous quasimorphisms one can improve the constant from 10 to 3, and a quasimorphism can always be homogenized by replacing it by $\lim_{n\to\infty} f(x^n)/n$ [Bavard 1991, p.135], which differs from the original f by at most $D(f)$.

Nevertheless, quasimorphisms can be utilized to construct interesting pseudolength functions, for example satisfying homogeneity on specific commutators. The following quasimorphism is due to Brooks [Fujiwara 2009, §2]. For a given word w in the free group $\boldsymbol{F}_2$, written in reduced form, let $f_w : \boldsymbol{F}_2 \to \mathbb{R}$ be the function which assigns to every other $g \in \boldsymbol{F}_2$, also written in reduced form, the maximum number of times that w occurs in g without overlaps, minus the analogous maximal number of times that w^{-1} can occur in g. Since $f_w(w^n) = n f_w(w)$, using (4.2) results in a pseudolength function that grows linearly on the powers of w. For example with w being the commutator of the generators of $\boldsymbol{F}_2$, we see that although the pseudolength function must be bounded on commutators by Proposition 2.1, it can nevertheless grow linearly on the powers of a fixed commutator.

Thus, there exist examples of quasihomogeneous semilength functions on free groups that are not induced by norms. Nevertheless, we will now see that for a large class of groups, including amenable groups and $G = \mathrm{SL}(n, \mathbb{Z})$ for $n \geq 3$, even all quasihomogeneous semilength functions are induced by norms on Banach spaces. Further, the bound from Proposition 2.1 even in the case of free groups is sharper than that obtained without using homogeneity.

Recall that the *commutator length* $\mathrm{cl}(g)$ of a word in $[G, G]$ is the length k of the shortest expression $g = [a_1, b_1] \cdot [a_2, b_2] \cdots [a_k, b_k]$ of g as a product of commutators. The *stable commutator length* is defined as $\lim_{n\to\infty} \mathrm{cl}(g^n)/n$, where the limit exists by subadditivity of the function $n \mapsto \mathrm{cl}(g^n)$.

Then Proposition 2.1, together with $\ell(e) \leq c$ and (2.7) for n a power of two,

$$\ell(x) \leq \frac{\ell(x^n)}{n} + \log_2(n)c,$$

easily imply the following estimates:

Proposition 4.3. *Let ℓ and c be as in Proposition 2.1. Then for $x \in [G, G]$, $\ell(x) \leq (5\,\mathrm{cl}(x) + 1)c$ and $\ell(x) \leq (5\,\mathrm{scl}(x) + 1)c$.*

We say two semilength functions $\ell_1, \ell_2 : G \to \mathbb{R}$ are *equivalent* if $|\ell_1(x) - \ell_2(x)|$ is bounded in $x \in G$.

For a perfect group G on which the stable commutator length vanishes (such as $\mathrm{SL}(n, \mathbb{Z})$ for $n \geq 3$), it is immediate that any homogeneous semilength function is bounded, and hence equivalent to the trivial semilength function $\ell(g) \equiv 0$.

More generally, for groups G for which the stable commutator length vanishes on $[G, G]$, we can deduce an analogue of Theorem 1.2. Note that there are several interesting examples of such groups, including solvable groups, and more generally, amenable groups.

Theorem 4.4. *Let G be a group such that the stable commutator length vanishes on $[G, G]$ and assume $\ell : G \to \mathbb{R}$ satisfies (2.2) and (2.3). Then there exist a real Banach space $\mathbb{B} = (\mathbb{B}, \|\;\|)$ and a group homomorphism $\phi : G \to \mathbb{B}$ such that ℓ is equivalent to $x \mapsto \|\phi(x)\|$.*

Remark 4.5. As in Remark 2.5, we can replace (2.2) by the *a priori* weaker condition that $\ell(xy) \leq \ell(x) + \ell(y) + k$ for all $x, y \in G$ with k fixed.

Proof. Let $\mathrm{ab}: G \to G_{\mathrm{ab}} = G/[G, G]$ be the abelianization homomorphism. We first construct a homogeneous semilength function $\bar{\ell}$ on G_{ab} so that ℓ is equivalent to $\bar{\ell} \circ \mathrm{ab}$. Let $\eta : G_{\mathrm{ab}} \to G$ be a section of ab and let $\bar{\ell}_0(x) := \ell(\eta(x)) + c$. We show that $\bar{\ell}_0$ is a semilength function. The required $\bar{\ell}$ will be obtained by homogenizing $\bar{\ell}_0$.

By Proposition 4.3, as the stable commutator length vanishes on $[G, G]$, it follows that for $x, y \in G$, if $\mathrm{ab}(x) = \mathrm{ab}(y)$, then $|\ell(x) - \ell(y)| \leq c$. Now, for $\alpha, \beta \in G_{\mathrm{ab}}$, $\mathrm{ab}(\eta(\alpha\beta)) = \mathrm{ab}(\eta(\alpha)\eta(\beta))$, hence

$$|\ell(\eta(\alpha\beta)) - \ell(\eta(\alpha)\eta(\beta)))| \leq c.$$

This together with the triangle inequality (2.2) gives

$$\bar{\ell}_0(\alpha\beta) \leq \bar{\ell}_0(\alpha) + \bar{\ell}_0(\beta) + c,$$

while using (2.3) instead gives the required lower bound for $\bar{\ell}_0(\alpha^2)$.

Next, for $x \in G$, as $\mathrm{ab}(\eta(\mathrm{ab}(x))) = \mathrm{ab}(x)$, we have $|\ell(x) - (\bar{\ell}_0 \circ \mathrm{ab})(x)| \leq c$. Thus ℓ is equivalent to $\bar{\ell}_0 \circ \mathrm{ab}$.

Since $(\alpha\beta)^n = \alpha^n\beta^n$ in G_{ab}, we also have $\bar{\ell}_0((\alpha\beta)^n) \leq \bar{\ell}_0(\alpha^n) + \bar{\ell}_0(\beta^n) + c$. We deduce that the homogenization $\bar{\ell}$ of $\bar{\ell}_0$ is a semilength function on G_{ab}, which is equivalent to $\bar{\ell}_0$ due to the bounds (2.6) and (2.7), applied to $\bar{\ell}_0$. Therefore also ℓ is equivalent to $\bar{\ell} \circ \mathrm{ab}$ on G.

The claim now follows upon applying Theorem 1.2 to $(G_{\mathrm{ab}}, \bar{\ell})$ and taking ϕ to be the composition $G \to G_{\mathrm{ab}} \to \mathbb{B}$. ☐

The following examples of length functions on the free group show that some hypotheses are needed to get bounds as strong as those of the theorem (naturally the stable commutator length does not vanish in the free group). For example, consider the word $[a^k, b^m]$ in the free group $\boldsymbol{F}_2$, generated by a and b, for some integers k and m:

- The norm of such an element with respect to the word metric is $2(|k| + |m|)$.
- If we have a length function ℓ which is symmetric and conjugation-invariant, but not necessarily homogeneous, then we have the bound $\ell([a^k, b^m]) \leq 2\min(|k|\ell(a), |m|\ell(b))$. Furthermore, the $\ell_{\mathrm{wc};a,b}$ from above are conjugation-invariant length functions for which these inequalities hold with equality.

  Further, $\ell([a^k, b^m]) \geq 2\min(|k|\ell(a), |m|\ell(b))$ as, for any matching M for $w = [a^k, b^m]$, if some pair (i, j) corresponds to letters a and a^{-1}, then no pair corresponds to letters b and b^{-1} and conversely. Further, it is easy to find a matching for w for which the deficiency is $\min(|k|\ell(a), |m|\ell(b))$. On the other hand, $\ell_{\mathrm{wc};a,b}$ is not homogeneous; for instance, $\ell([a, b]) = 2$ and $\ell([a, b]^3) = 4$. Similarly, we have $\ell([a^k, b^k]) = 2|k|$ and $\ell([a^k, b^k]^3) \leq 4|k|$, which demonstrates that $2\ell(x) - \ell(x^2)$ is unbounded (as must be the case, according to (2.15)).
- On the other hand, the function ℓ_{cyc} associating to each word the length of its cyclically reduced form is homogeneous, but not a semilength function. For this we have $\ell_{\mathrm{cyc}}([a^k, b^m]) = 2(|k| + |m|)$.

Observe that all of the bounds on $\ell([a^k, b^m])$ here become unbounded as $k, m \to \infty$. This should be compared with Proposition 2.1, which establishes a bound $\ell([a^k, b^m]) \leq 5c$ that is uniform in k and m for any function ℓ satisfying the hypotheses of that proposition.

Finite balls in free groups. From Proposition 2.1 and a standard compactness argument, we can establish the following local version of the theorem.

Theorem 4.6. *For any $\varepsilon > 0$ there exists $R \geq 4$ with the following property: if a, b are two elements of a group G, $B_{a,b}(R) \subset G$ is the collection of all words in a, b, a^{-1}, b^{-1} of length at most R (so in particular $B_{a,b}(R)$ contains $[a, b]$), and the map $\ell : B_{a,b}(R) \to [0, +\infty)$ is a "local semilength function" which obeys the triangle inequality*

$$\ell(xy) \leq \ell(x) + \ell(y) \tag{4.7}$$

whenever $x, y, xy \in B_{a,b}(R)$, with equality when $x = y$, then one has

$$\ell([a, b]) \leq \varepsilon(\ell(a) + \ell(b)).$$

Proof. By pulling back to the free group $\boldsymbol{F}_2$ generated by a and b, we may assume without loss of generality that $G = \boldsymbol{F}_2$. Without loss of generality we may also normalize $\ell(a) + \ell(b) = 1$. If the claim failed, then one could find a sequence $R_n \to \infty$ and local pseudolength functions $\ell_n : B_{a,b}(R_n) \to [0, +\infty)$ such that $\ell_n(a) + \ell_n(b) = 1$, but that $\ell_n([a, b]) \geq \varepsilon$. By the Arzela–Ascoli theorem, we can pass to a subsequence that converges pointwise to a homogeneous pseudolength function $\ell : G \to [0, +\infty)$ such that $\ell([a, b]) \geq \varepsilon$, which contradicts Proposition 2.1. $\square$

Remark 4.8. By carefully refining the arguments in the previous section, choosing n to be various small powers of R instead of sending n to infinity, one can extract an explicit value of R of the form $R = C\varepsilon^{-A}$ for some absolute constants $C, A > 0$; we leave the details to the interested reader.

On the other hand, for any finite R one can construct local length functions $\ell : B(0, R) \to [0, +\infty)$ such that $\ell(x) > 0$ for all $x \in B(0, R) \setminus \{e\}$. One construction is as follows. Any two matrices $U_a, U_b \in \mathrm{SO}(3)$ define a representation $x \mapsto U_x$ of the free group $\boldsymbol{F}_2$ in the obvious fashion. Every U_x is then a rotation around some axis in $\mathbb{R}^3$ by some angle $0 \leq \theta_x \leq \pi$ in one of the two directions around that axis; if U_a and U_b are sufficiently close to the identity, then the angle θ_x is at most $\pi/2$ for all $x \in B(0, R)$. We set $\ell(x) := \theta_x$ for $x \in B(0, R)$. Also, if U_a and U_b are chosen generically, the representation is faithful, as follows from the dominance of word maps on simple Lie groups such as SO(3), see [Borel 1983]. Hence $\ell(x) > 0$ for any nonidentity x. From the triangle inequality for angles we thus have (4.7) whenever $x, y, xy \in B(0, R)$, with equality when $x = y$. Note that as one sends $R \to \infty$, the local length functions constructed here converge to zero pointwise, so in the limit we do not get any counterexample to the main theorem.

About this project

This project is an online collaboration that originated from a blog post at

https://terrytao.wordpress.com/2017/12/16,

following the model of the "Polymath" projects [Gowers and Nielsen 2009]. A full list of participants and their grant acknowledgments may be found at

http://michaelnielsen.org/polymath1/index.php?title=linear_norm_grant_acknowledgments.

Acknowledgements

We thank Michal Doucha for useful references and comments, in particular in bringing the paper [Niemiec 2013] to our attention, and the anonymous referee for helpful suggestions.

References

[Bavard 1991] C. Bavard, "Longueur stable des commutateurs", *Enseign. Math.* (2) **37**:1-2 (1991), 109–150. MR Zbl

[Borel 1983] A. Borel, "On free subgroups of semisimple groups", *Enseign. Math.* (2) **29**:1-2 (1983), 151–164. MR Zbl

[Cabello Sánchez and Castillo 2002] F. Cabello Sánchez and J. M. F. Castillo, "Banach space techniques underpinning a theory for nearly additive mappings", *Dissertationes Math.* (*Rozprawy Mat.*) **404** (2002), 73. MR Zbl

[Fujiwara 2009] K. Fujiwara, "Quasi-homomorphisms on mapping class groups", pp. 241–269 in *Handbook of Teichmüller theory*, vol. 2, edited by A. Papadopoulos, IRMA Lect. Math. Theor. Phys. **13**, Eur. Math. Soc., Zürich, 2009. MR Zbl

[Gadgil 2009] S. Gadgil, "Watson–Crick pairing, the Heisenberg group and Milnor invariants", *J. Math. Biol.* **59**:1 (2009), 123–142. MR Zbl

[Gajda and Kominek 1991] Z. Gajda and Z. Kominek, "On separation theorems for subadditive and superadditive functionals", *Studia Math.* **100**:1 (1991), 25–38. MR Zbl

[Gowers and Nielsen 2009] T. Gowers and M. Nielsen, "Massively collaborative mathematics", *Nature* **461** (2009), 879–881.

[Khare and Rajaratnam 2016] A. Khare and B. Rajaratnam, "The Khinchin–Kahane inequality and Banach space embeddings for metric groups", preprint, 2016. arXiv

[Khare and Rajaratnam 2017] A. Khare and B. Rajaratnam, "The Hoffmann-Jørgensen inequality in metric semigroups", *Ann. Probab.* **45**:6A (2017), 4101–4111. MR Zbl

[Ledoux and Talagrand 1991] M. Ledoux and M. Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **23**, Springer, 1991. MR Zbl

[Milnor 1976] J. Milnor, "Curvatures of left invariant metrics on Lie groups", *Advances in Math.* **21**:3 (1976), 293–329. MR Zbl

[Naor 2012] A. Naor, "An introduction to the Ribe program", *Jpn. J. Math.* **7**:2 (2012), 167–233. MR Zbl

[Niemiec 2013] P. Niemiec, "Universal valued Abelian groups", *Adv. Math.* **235** (2013), 398–449. MR Zbl

Communicated by Pham Huu Tiep
Received 2018-01-11 Revised 2018-04-20 Accepted 2018-06-12

fritz@mis.mpg.de *Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany*

gadgil@math.iisc.ernet.in *Department of Mathematics, Indian Institute of Science, Bangalore, India*

khare@iisc.ac.in *Department of Mathematics, Indian Institute of Science, Bangalore, India*
Analysis and Probability Research Group, Indian Institute of Science, Bangalore, India

pace@math.byu.edu *Department of Mathematics, Brigham Young University, Provo, UT, United States*

lior@math.ubc.ca *University of British Columbia, Vancouver BC, Canada*

tao@math.ucla.edu *Department of Mathematics, University of California Los Angeles, Los Angeles, CA, United States*

ALGEBRA AND NUMBER THEORY 12:7 (2018)
dx.doi.org/10.2140/ant.2018.12.1787

When are permutation invariants Cohen–Macaulay over all fields?

Ben Blum-Smith and Sophie Marques

We prove that the polynomial invariants of a permutation group are Cohen–Macaulay for any choice of coefficient field if and only if the group is generated by transpositions, double transpositions, and 3-cycles. This unites and generalizes several previously known results. The "if" direction of the argument uses Stanley–Reisner theory and a recent result of Christian Lange in orbifold theory. The "only if" direction uses a local-global result based on a theorem of Raynaud to reduce the problem to an analysis of inertia groups, and a combinatorial argument to identify inertia groups that obstruct Cohen–Macaulayness.

1. Introduction

The invariant ring of a graded action by a finite group G on a polynomial ring

$$k[\boldsymbol{x}] = k[x_1, \dots, x_n]$$

over a field k is well behaved when the field characteristic is prime to the group order. For example, it is generated in degree $\leq |G|$ (Noether's bound), and it is a Cohen–Macaulay ring (the Hochster–Eagon theorem).

When the characteristic divides the group order (the *modular case*), the situation is much more mysterious. Both of these statements (and many others) can, but do not always, fail. The question of when such pathologies arise has attracted research attention over the last few decades.

In this article we focus on Cohen–Macaulayness. Let $k[\boldsymbol{x}]^G$ be the invariant ring and let

$$p = \operatorname{char} k$$

be the field characteristic. We interpret $k[\boldsymbol{x}]$ as the coordinate ring of $\mathbb{A}_k^n$, so that the action of G on $k[\boldsymbol{x}]$ is induced from an action on $\mathbb{A}_k^n$ by automorphisms. Because the action on $k[\boldsymbol{x}]$ is graded, the corresponding action on $\mathbb{A}_k^n$ is linear, i.e., it arises from a linear representation of G on a k-vector space. Here is a sampling of known results:

- Ellingsrud and Skjelbred [1980] showed that if G is cyclic of order p^m, then $k[\boldsymbol{x}]^G$ is not Cohen–Macaulay unless G fixes a subspace of $\mathbb{A}_k^n$ of codimension ≤ 2.

MSC2010: primary 13A50; secondary 05E40.
Keywords: invariant theory, modular invariant theory, henselization, Stanley–Reisner, Cohen–Macaulay, commutative ring, finite group.

- Larry Smith [1996] showed that if $n = 3$, then $k[\boldsymbol{x}]^G$ is Cohen–Macaulay. (This was known to hold for $n \le 2$.)
- Campbell et al. [1999] showed that if G is a p-group, and if the action of G on $\mathbb{A}_k^n$ is the sum of three copies of the same linear representation, then $k[\boldsymbol{x}]^G$ is not Cohen–Macaulay.
- Gregor Kemper [1999] showed that if G is a p-group and $k[\boldsymbol{x}]^G$ is Cohen–Macaulay, then G is necessarily generated by elements g whose fixed-point sets in $\mathbb{A}_k^n$ have codimension ≤ 2, generalizing [Ellingsrud and Skjelbred 1980] beyond cyclic groups and [Campbell et al. 1999] beyond three-copies representations.

See [Kemper 2012] for a more detailed overview.

A theme uniting these results is that generation of G by elements fixing codimension ≤ 2 subspaces is related to good behavior of $k[\boldsymbol{x}]^G$. Further variations on this theme are found in [Dufresne et al. 2009; Gordeev and Kemper 2003; Kac and Watanabe 1982; Lorenz and Pathak 2001]. The main goal of this paper is a result of this kind for permutation groups $G \subset S_n$, acting on $k[\boldsymbol{x}]$ by permuting the x_i. The result characterizes permutation groups generated in this way, and is not restricted to p-groups.

Permutation groups have the feature that the definition of the action is insensitive to the choice of a ground field k. Thus it is natural to ask:

Question 1.1. For which $G \subset S_n$ is $k[\boldsymbol{x}]^G$ Cohen–Macaulay regardless of k?

An additional motivation for this question is that $k[\boldsymbol{x}]^G$ is Cohen–Macaulay for every choice of k if and only if $\mathbb{Z}[\boldsymbol{x}]^G$ is free as a module over the subring $\mathbb{Z}[\boldsymbol{x}]^{S_n}$ of symmetric polynomials, and also if and only if $A[\boldsymbol{x}]^G$ is Cohen–Macaulay for every Cohen–Macaulay ring A. (We will not develop these equivalences here, but see [Blum-Smith 2017, §2.4.1] where the first is worked out in detail, and [Bruns and Herzog 1993, Exercise 5.1.25] for a sketch of the second in a slightly different setting.)

Kemper [2001] gave an if-and-only if criterion that determines Cohen–Macaulayness of a permutation invariant ring when p divides $|G|$ exactly once. This criterion allows one to determine Cohen–Macaulayness for many specific groups and primes, but does not in general answer Question 1.1 because few permutation groups have squarefree order. Some special cases of Question 1.1 are known:

- If G is a Young subgroup (i.e., a product of symmetric groups acting on disjoint sets), then $k[\boldsymbol{x}]^G$ is a polynomial algebra over k, so it is Cohen–Macaulay regardless of k.
- It follows from the result of Kemper [1999] quoted above that if G is a p-group, then $k[\boldsymbol{x}]^G$ cannot be Cohen–Macaulay over all fields unless G is generated by transpositions and double transpositions, or 3-cycles (and $p = 2$ or 3).
- Kemper [1999] also showed that if $G \subset S_n$ is regular (i.e., its action on

$$[n] = \{1, \dots, n\}$$

is free and transitive), then $k[\boldsymbol{x}]^G$ is Cohen–Macaulay over every k if it is isomorphic to C_2, C_3, or $C_2 \times C_2$, but not otherwise. (In fact, in other cases, it is not Cohen–Macaulay for any k with $\operatorname{char} k$ dividing $|G|$.)

- Victor Reiner [1992; 2003] has shown that A_n, and the diagonally embedded $S_n \hookrightarrow S_n \times S_n \subset S_{2n}$, have invariant rings that are Cohen–Macaulay regardless of the field. (These are the S_n-cases of results he found for all finite Coxeter groups.) Patricia Hersh [2003a; 2003b] has shown the same for the wreath product $S_2 \wr S_n \subset S_{2n}$.

Our main objective in this article is to answer Question 1.1 completely. We will prove the following theorem, which unites all of these cases and ties them into the theme mentioned above.

Theorem 1.2. *Let $G \subset S_n$. The ring $k[\boldsymbol{x}]^G$ is Cohen–Macaulay for all choices of k if and only if G is generated by transpositions, double transpositions, and 3-cycles.*

Let N be the subgroup of G generated by transpositions, double transpositions, and 3-cycles. The "if" direction of Theorem 1.2, together with the Hochster–Eagon theorem [Hochster and Eagon 1971, Proposition 13], imply that the characteristics p in which $k[\boldsymbol{x}]^G$ fails to be Cohen–Macaulay must be among those that divide $[G : N]$. This implication will be discussed in more detail in the conclusion (Section 5). The "only if" direction implies that if $[G:N] > 1$, then there is at least one such characteristic p. This p is explicitly constructed in the course of the proof.

The proof of this theorem is methodologically eclectic. The "if" direction uses Stanley–Reisner theory, which relates Cohen–Macaulayness of $k[\boldsymbol{x}]^G$ to the topology of the quotient of a ball by G, and a recent result in orbifold theory by Christian Lange [2016] that characterizes the groups G such that this quotient is a piecewise-linear ball. The "only if" direction is much more algebraic. It is based on a local-global result (Theorem 3.1) reducing the Cohen–Macaulayness of a noetherian invariant ring to that of the invariant rings of its inertia groups acting on strict localizations.

Though Theorem 1.2 is specific to the situation of a polynomial ring $k[\boldsymbol{x}]$ and a permutation group G, a substantial portion of our method for the "only if" direction applies in considerably more generality. Section 2C concerns arbitrary commutative, unital rings, and the local-global result just mentioned only assumes that the invariant ring is noetherian. (Other work on Cohen–Macaulayness of invariants at the generality of noetherian rings includes [Gordeev and Kemper 2003; Lorenz and Pathak 2001].) A secondary goal of this paper is to develop these general tools, which we expect have broader applicability. The fact that Cohen–Macaulayness depends fully on the local action of the inertia groups yields information about Cohen–Macaulayness whenever inertia groups can be accessed directly and are simpler than the whole group, as in the present case.

The method of the "if" direction is similar to the methods used by Reiner [1992; 2003] and Hersh [2003a; 2003b] to prove the results mentioned above. The novelty is the application of Lange's orbifold result [2016] in place of an explicit shelling of a cell complex. The main novelties in the "only if" direction are: the local-global Theorem 3.1, its application to show that certain kinds of inertia p-groups obstruct

Cohen–Macaulayness (Proposition 3.11), and a combinatorial argument that exhibits such an inertia p-group explicitly in the case at hand (Lemma 4.5).

The organization of the paper is as follows. Section 2 collects together the needed background from commutative algebra, Stanley–Reisner theory, and piecewise-linear topology, and introduces notation that is used throughout the article. Section 3 contains the general results on Cohen–Macaulayness and inertia groups that are needed for the "only if" direction of Theorem 1.2, including the local-global Theorem 3.1 and the p-group obstruction Proposition 3.11. Section 4 proves the "if" direction of Theorem 1.2, and then using this, proves the "only if" direction. Finally, Section 5 draws out some implications and poses questions for further inquiry.

2. Background

Throughout this paper, A denotes an arbitrary commutative, unital ring, k denotes a field, p denotes the characteristic of k, $k[\boldsymbol{x}]$ denotes the polynomial ring $k[x_1, \dots, x_n]$, $[n]$ denotes the set $\{1, \dots, n\}$, and G denotes a finite group with a faithful action on $k[\boldsymbol{x}]$ by permutations of the x_i's, or on A by arbitrary automorphisms. In Section 4B, the prime number p will be conceptually prior to k, and k will be chosen to satisfy $\operatorname{char} k = p$.

2A. ***Cohen–Macaulayness.*** Recall that the *depth* of a local noetherian ring is the length of the longest regular sequence contained in the maximal ideal. The depth is always bounded above by the dimension. When equality is achieved, the ring is said to be *Cohen–Macaulay*. A general noetherian ring is defined to be Cohen–Macaulay if its localization at every maximal, or equivalently at every prime, is Cohen–Macaulay [Bruns and Herzog 1993, Definition 2.1.1 and Theorem 2.1.3(b)].

Although there has been work on extending the theory of Cohen–Macaulayness to the nonnoetherian setting [Hamilton and Marley 2007], in this paper we will follow tradition by regarding noetherianity as a requirement of Cohen–Macaulayness.

Cohen–Macaulayness is automatic for artinian rings, since if the dimension is zero, the depth of a localization cannot be strictly lower than this. For example, fields are Cohen–Macaulay. Noetherian regular rings, for example polynomial rings over fields, are also Cohen–Macaulay [Bruns and Herzog 1993, Corollary 2.2.6].

For our purposes it will be necessary to know how the Cohen–Macaulayness of a ring relates to that of a flat extension. The needed fact [Bruns and Herzog 1993, Theorem 2.1.7] is that if $A \to B$ is a flat extension of noetherian rings, then B is Cohen–Macaulay if and only if, for each prime ideal $\mathfrak{q}$ of B and its contraction $\mathfrak{p}$ in A, both $A_{\mathfrak{p}}$ and $B_{\mathfrak{q}}/\mathfrak{p}B_{\mathfrak{q}}$ are Cohen–Macaulay. It is enough to quantify this statement over maximal ideals $\mathfrak{q}$ of B. We will use this fact repeatedly in Section 3.

When a noetherian ring is finite over a regular subring, Cohen–Macaulayness is related to flatness as a module over the subring. In the traditional situation of invariant theory, this fact has a particularly nice formulation. For if $k[\boldsymbol{x}]$ is a polynomial ring over a field, and G acts by graded automorphisms, then $k[\boldsymbol{x}]^G$ is finitely generated and graded, and the Noether normalization lemma guarantees a graded polynomial subring (generated by a *homogeneous system of parameters*) over which $k[\boldsymbol{x}]^G$ is finite. In this situation, $k[\boldsymbol{x}]^G$ is Cohen–Macaulay if and only if it is a *free* module over this subring (the *Hironaka criterion*). We will not build on this fact directly, but we mention it both because it motivates interest in Cohen–Macaulayness, and because we do use a result [Reiner 2003, Theorem A.1] that depends on it, whose proof we outline in the next section.

2B. ***Combinatorial commutative algebra and* PL *topology.*** The proof of the "if" direction of Theorem 1.2 relies on results in combinatorial commutative algebra and some basic facts about PL topology. For motivation, we describe the plan of the proof before recalling these results.

By work of Adriano Garsia and Dennis Stanton [1984], refined by Victor Reiner [2003], Cohen–Macaulayness of the polynomial invariant ring $k[\boldsymbol{x}]^G$ can be deduced from the Cohen–Macaulayness of the *Stanley–Reisner ring* of a certain cell complex (specifically a *boolean complex*) that depends on G. The Cohen–Macaulayness of this Stanley–Reisner ring can in turn be deduced from information about the complex that depends only on the homeomorphism class of its total space. For G generated as in Theorem 1.2, a recent result of Christian Lange [2016] hands us this topological information. This is the structure of the proof, which will be assembled in Section 4A. Here, we recall the needed results and definitions regarding boolean complexes and Stanley–Reisner rings.

Let P be a finite poset and k a field.

Definition 2.1. The *Stanley–Reisner ring* of P over k, written $k[P]$, is the quotient of the polynomial ring $k[\{y_\alpha\}_{\alpha\in P}]$, with indeterminates indexed by the elements of P, by the ideal generated by products $y_\alpha y_\beta$ indexed by incomparable pairs $\alpha, \beta \in P$.

Remark 2.2. This is a special case of a more general definition, which we will not use directly: the Stanley–Reisner ring of a simplicial complex. (We will use a further generalization — see Definition 2.5 below.) The Stanley–Reisner ring of a poset is nothing but the Stanley–Reisner ring of the *chain complex* of the poset, i.e., the simplicial complex with vertex set the elements of the poset, whose simplices are the chains in the poset. It is helpful to keep in mind that the Stanley–Reisner ring of a poset has an underlying simplicial complex as well.

Write $[n] = \{1, \dots, n\}$. Let B_n be the boolean algebra on the set $[n]$, i.e., the set of subsets of $[n]$, ordered by inclusion. Then the Stanley–Reisner ring $k[B_n \setminus \{\varnothing\}]$ is, in a sense that can be made precise, a coarse approximation of the polynomial ring $k[\boldsymbol{x}]$. In particular, it carries a natural action of S_n via the latter's action on the set $[n]$, and if $G \subset S_n$, then $k[\boldsymbol{x}]^G$ is Cohen–Macaulay whenever $k[B_n \setminus \{\varnothing\}]^G$ is Cohen–Macaulay. This is the content of [Reiner 2003, Theorem A.1].

The proof is given in full there, and also in great detail in [Blum-Smith 2017, Section 2.5.3], and in any case is essentially a characteristic-neutral reformulation of an argument of Adriano Garsia and Dennis Stanton [1984], building on Garsia's earlier work [1980]. However, we would like this result to be better-known, so we indicate the line of proof.

As mentioned in Section 2A, a finitely generated graded k-algebra is Cohen–Macaulay if and only if it is free as a module over the subring generated by any homogeneous system of parameters. Thus, Cohen–Macaulayness can be established by showing the existence of a module basis over such a subring. For any $G \subset S_n$, $k[\boldsymbol{x}]^{S_n}$ and $k[B_n \setminus \{\varnothing\}]^{S_n}$ are such subrings, respectively, of $k[\boldsymbol{x}]^G$ and $k[B_n \setminus \{\varnothing\}]^G$, and they are isomorphic. Thus, Cohen–Macaulayness may be passed from $k[B_n \setminus \{\varnothing\}]^G$ to $k[\boldsymbol{x}]^G$ by showing that the existence of a module basis for the former over the common subring $k[B_n \setminus \{\varnothing\}]^{S_n} \cong k[\boldsymbol{x}]^{S_n}$ implies the existence of a basis for the latter. Garsia [1980] introduced a k-linear, S_n-equivariant map $\mathcal{G} : k[B_n \setminus \{\varnothing\}] \to k[\boldsymbol{x}]$ sending

$$y_U \mapsto \prod_{i \in U} x_i,$$

where $U \in B_n \setminus \{\varnothing\}$ is any nonempty subset of $[n]$. The map $\mathcal{G}$ is first extended multiplicatively to all monomials of $k[B_n \setminus \{\varnothing\}]$, and then k-linearly to the whole ring. This map is an isomorphism of k-vector spaces, and also, in a sense made precise in [Blum-Smith 2017, Proposition 2.5.66], a coarse approximation of a ring homomorphism. In particular, for any $G \subset S_n$, if $k[B_n \setminus \{\varnothing\}]^G$ is Cohen–Macaulay, it maps an appropriately chosen $k[B_n \setminus \{\varnothing\}]^{S_n}$-basis of $k[B_n \setminus \{\varnothing\}]^G$ to a $k[\boldsymbol{x}]^{S_n}$-basis of $k[\boldsymbol{x}]^G$. This statement about bases was proven by Garsia and Stanton [1984] with $k = \mathbb{Q}$, in which case both rings are automatically Cohen–Macaulay — Garsia and Stanton's interest was in the explicit construction of bases — but it was observed in [Reiner 2003, Theorem A.1] that the argument is characteristic-neutral and so allows one to deduce Cohen–Macaulayness of $k[\boldsymbol{x}]^G$ from that of $k[B_n \setminus \{\varnothing\}]^G$ in the modular situation.

Remark 2.3. The map $\mathcal{G}$ is referred to as the *transfer map* in [Garsia 1980; Garsia and Stanton 1984; Reiner 2003]. Other authors in invariant theory [Neusel and Smith 2002; Smith 1995] use the same phrase to denote the A^G-linear map

$$\begin{aligned} \mathrm{Tr} : A &\to A^G \\ x &\mapsto \sum_{g \in G} g(x). \end{aligned}$$

While this latter map is also called the *trace*, there are well-established usages of *transfer* to describe maps analogous to Tr in both topology and group theory, so we prefer to call $\mathcal{G}$ the *Garsia map* to avoid competition for the term and to honor Garsia's introduction of it [1980]. The present paper makes no use of the Garsia map except implicitly in quoting [Reiner 2003, Theorem A.1].

The work cited above reduces proving Cohen–Macaulayness of $k[\boldsymbol{x}]^G$ to the analogous statement for $k[B_n \setminus \{\varnothing\}]^G$. The Cohen–Macaulayness of this latter ring can be assessed using a topological criterion, following a general philosophy in Stanley–Reisner theory that the Cohen–Macaulayness of a Stanley–Reisner ring is equivalent to a condition on the homology of the underlying simplicial complex. In the present situation, $k[B_n \setminus \{\varnothing\}]^G$ is not the Stanley–Reisner ring of a poset or simplicial complex, but it turns out to be the Stanley–Reisner ring of a *boolean complex*. We recall the needed definitions:

Definition 2.4. A *boolean complex* is a regular CW complex in which every face has the combinatorial type of a simplex.

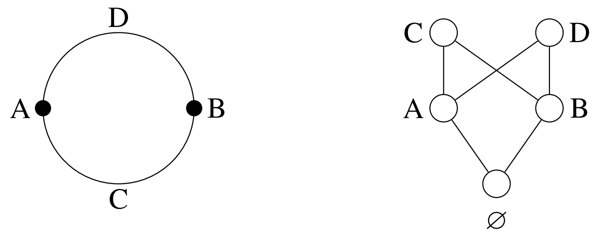

Figure 1. Left: a boolean complex with total space homeomorphic to a circle. Right: its face poset.

This is a mild generalization of a simplicial complex, in which it is possible for two faces to intersect in an arbitrary subcomplex rather than a single subface. (For example, two faces can have all the same vertices.) See Figure 1. The terminology is due to Garsia and Stanton [1984].

The *face poset* of a cell complex is the poset whose elements are the cells (*faces*), and the relation $\alpha \le \beta$ means that α's closure is contained in β's closure. For our purposes it is convenient to modify this definition to include an additional *empty face* $\varnothing$, with $\varnothing \le \alpha$ for all faces α. With this convention, a boolean complex can be characterized as a regular CW complex whose face poset has the property that every lower interval is a finite boolean algebra; this is the etymology of the name *boolean complex*. Face posets of boolean complexes are referred to as *simplicial posets*, a term introduced by Richard Stanley [1986].

Stanley [1991] generalized the notion of a Stanley–Reisner ring to a boolean complex Ω, as follows. Let k be a field and let Q be the face poset of Ω, including the minimal element $\varnothing$. Let $k[\{z_\alpha\}_{\alpha\in Q}]$ be a polynomial ring with indeterminates indexed by the elements of Q. Let I be the ideal of this ring generated by:

(1) The element $z_\varnothing - 1$.

(2) All products $z_\alpha z_\beta$ where $\alpha, \beta \in Q$ have no common upper bound.

(3) All elements of the form
$$z_\alpha z_\beta - z_{\alpha\wedge\beta} \sum_{\gamma\in\mathrm{lub}(\alpha,\beta)} z_\gamma$$
where α and β have at least one common upper bound and $\mathrm{lub}(\alpha, \beta)$ denotes the (consequently nonempty) set of least upper bounds of α and β.

The greatest (common) lower bound $\alpha \wedge \beta$ of α and β exists and is unique in the above formula because, as remarked above, every lower interval, and in particular the lower interval below any common upper bound for α and β, is a boolean algebra and therefore a lattice. Thus whenever α and β have any common upper bound, they have a unique greatest common lower bound in some lower interval containing them both, and thus in the whole poset.

Definition 2.5. The quotient ring $k[\{z_\alpha\}_{\alpha\in P}]/I$ is called the *Stanley–Reisner ring* of Ω and denoted $k[\Omega]$.

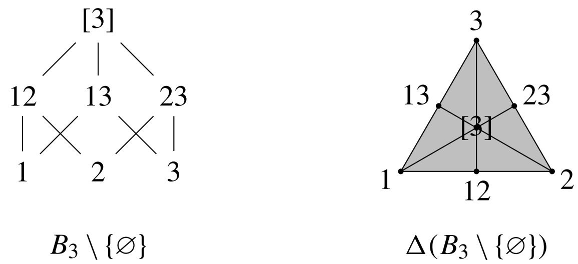

Figure 2. The poset $B_3 \setminus \{\varnothing\}$, and its order complex, which is a 2-ball.

Remark 2.6. Definition 2.5 generalizes Definition 2.1, but in a somewhat subtle way. Given a poset P, one can form its chain complex Ω, regarded as a boolean complex, and then the $k[P]$ of Definition 2.1 will be isomorphic to the $k[\Omega]$ of Definition 2.5; however, the poset Q of the latter definition will not be P. Instead, its elements will be *chains* in P, ordered by inclusion. For example, let $P = B_2 \setminus \{\varnothing\}$. Then the elements of P may be abbreviated 1, 2, and 12, and the only incomparable pair consists of 1 and 2. Thus

$$k[P] = k[y_1, y_2, y_{12}]/(y_1 y_2)$$

according to Definition 2.1. However, Q consists of the six chains in P: the empty chain $\varnothing$, three chains of length 1 (1, 2, and 12), and two chains of length 2 ($1 \subset 12$ and $2 \subset 12$). Thus

$$k[\Omega] = k[z_\varnothing, z_1, z_2, z_{12}, z_{1\subset 12}, z_{2\subset 12}]/I$$

where I is as described above. The isomorphism is given by mapping the z of a given chain to the product of y's corresponding to elements of the chain, for example $z_{1\subset 12} \mapsto y_1 y_{12}$. Indeed, the definition of I becomes much more transparent after considering why this map is an isomorphism.

The ring of interest to us is the invariant ring $k[B_n \setminus \{\varnothing\}]^G$ inside the Stanley–Reisner ring of the poset $B_n \setminus \{\varnothing\}$. This ring can be identified with the Stanley–Reisner ring of a boolean complex using a result of Victor Reiner, as follows. Let Δ be the order complex of $B_n \setminus \{\varnothing\}$, i.e., the simplicial complex whose vertices are the elements of $B_n \setminus \{\varnothing\}$, and whose faces are the chains in $B_n \setminus \{\varnothing\}$. As a simplicial complex, Δ is the barycentric subdivision of an $(n-1)$-simplex, thus it is topologically an $(n-1)$-ball. See Figure 2.

The simplicial complex Δ carries a natural simplicial action of S_n, via the latter's action on $[n]$. The quotient cell complex Δ/G is usually not simplicial, but it is a boolean complex. This is because Δ is a *balanced* complex, and the action of G is a balanced action.

Definition 2.7. A boolean complex of dimension d is *balanced* if there is a labeling of its vertices by $d+1$ labels such that the vertices of any one face have distinct labels. Given such a labeling, a cellular action by a group is a *balanced action* if it preserves the labeling.

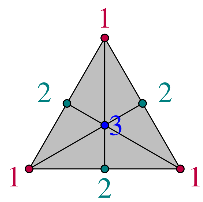

Figure 3. The labeling of the order complex of $B_3 \setminus \{\varnothing\}$, showing it is balanced.

In the present case, the vertices of Δ are the nonempty subsets of $[n]$, and thus Δ is balanced by associating a subset to its cardinality. (Here, $d = n - 1$, so the n possible cardinalities give the right number of labels.) The action of S_n is clearly balanced with respect to this labeling. See Figure 3.

It is straightforward to check that the quotient of a balanced boolean complex by a balanced action is again a balanced boolean complex. (Details are given in [Blum-Smith 2017, Lemma 2.5.86].) Thus Δ/G is a balanced boolean complex.

Victor Reiner [1992, Theorem 2.3.1] showed that if a group G acts cellularly and balancedly on a balanced boolean complex Ω, then the invariant ring $k[\Omega]^G$ inside the Stanley–Reisner ring of Ω is isomorphic to $k[\Omega/G]$, the Stanley–Reisner ring of the quotient boolean complex Ω/G. In the present situation, this gives us

$$k[\Delta/G] \cong k[B_n \setminus \{\varnothing\}]^G. \tag{1}$$

Thus the problem is reduced to showing that $k[\Delta/G]$ is Cohen–Macaulay.

Finally, the Cohen–Macaulayness of $k[\Delta/G]$ can be assessed topologically. In general, the Cohen–Macaulayness of the Stanley–Reisner ring of a boolean complex Ω is equivalent (just as for a simplicial complex) to a condition on $|\Omega|$, the underlying topological space of Ω, that depends only on its homeomorphism class. Namely, $k[\Omega]$ is Cohen–Macaulay if and only if

$$\tilde{H}_i(|\Omega|; k) = 0 \quad \text{and} \quad H_i(|\Omega|, |\Omega| - q; k) = 0 \tag{2}$$

for all points $q \in |\Omega|$ and all $i < \dim \Omega$. (Here, $H_i(|\Omega|, |\Omega| - q; k)$ is relative singular homology and $\tilde{H}_i(|\Omega|; k)$ is reduced singular homology.) This theorem is the product of work of Gerald Reisner (building on work of Melvin Hochster), James Munkres, Richard Stanley, and Art Duval. Reisner [1976] proved that for a *simplicial* complex Ω, Cohen–Macaulayness of $k[\Omega]$ is equivalent to a homological vanishing condition that a priori depends on the simplicial structure and not just the underlying topological space. Munkres [1984] showed that Reisner's condition is equivalent to the purely topological condition stated above. Richard Stanley [1991] showed that the direction

$$\text{(2) is satisfied for all } q \in |\Omega| \text{ and } i < \dim \Omega \Rightarrow k[\Omega] \text{ is Cohen–Macaulay}$$

generalizes to boolean complexes, and Art Duval [1997] showed that this generalization is bidirectional. See [Blum-Smith 2017, §2.5.2] for more details.

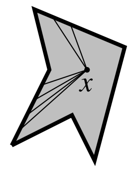

Figure 4. A compact cone neighborhood of a point in $\mathbb{R}^2$. The link K is drawn in bold, and the star S is the entire set, the union of segments from x to the points of K. Some of these segments are also drawn. Note each point of $S \setminus \{x\}$ is on exactly one such segment.

Remark 2.8. Since we only use Stanley–Reisner theory to show the "if" direction of Theorem 1.2 and thus we only need it to deduce Cohen–Macaulayness, and not the *failure* of Cohen–Macaulayness, the proof of Theorem 1.2 only uses Stanley's and not Duval's part of the generalization of (2) to boolean complexes.

Combining the results quoted above, we see that to demonstrate the Cohen–Macaulayness of the ring $k[\boldsymbol{x}]^G$, it is sufficient to prove that the boolean complex $\Omega = \Delta/G$ satisfies the homological vanishing condition (2) for all $x \in |\Delta/G|$ and all $i < n-1$. The proof of the "if" direction of Theorem 1.2 will consist in showing that this condition holds when G is generated by transpositions, double transpositions, and 3-cycles.

This will be accomplished by quoting a recent result of Christian Lange (see Section 4A) that is stated in the language of piecewise-linear (PL) topology, so we also need to recall a few definitions and a basic fact from this field. We follow [Lange 2016, §3.1] and [Rourke and Sanderson 1972, Chapters 1 and 2] for these details. A *polyhedron* is a subset X of $\mathbb{R}^m$ in which each point has a compact cone neighborhood, i.e., given $x \in X$, there is a compact set $K \subset X$ such that (i) the union S of line segments from x to points of K is contained in X, (ii) each point of $S \setminus \{x\}$ is on a unique such line segment from x, and (iii) S is a neighborhood of x in X, i.e., it contains an open subset of X containing x. The set S is called a *star* of x in X, and K is called a *link* of x. See Figure 4.

Remark 2.9. This definition of *polyhedron* is a technical device, used here to define the concepts *piecewise-linear* and *polyhedral star*. It includes the more conventional meaning of a three-dimensional polytope as a special case, but is much, much broader. For example, any open subset of $\mathbb{R}^n$, or of any polytope, is a polyhedron.

More broadly, our use of PL topology in this paper is only to serve a technical need linking Lange's result to our setting.

If $X \subset \mathbb{R}^m$ and $Y \subset \mathbb{R}^n$ are polyhedra, a continuous map $f : X \to Y$ is a *piecewise-linear* (or PL) *map* if its graph $\{(x, f(x)) : x \in X\} \subset \mathbb{R}^{m+n}$ is a polyhedron. A *piecewise-linear* (or PL) *space* is a second-countable, Hausdorff topological space equipped with a covering by open sets U_i, each with a homeomorphism $\varphi_i : X_i \to U_i$ from a polyhedron X_i in some $\mathbb{R}^{m_i}$, such that the transition maps

$$\varphi_j^{-1} \circ \varphi_i|_{\varphi_i^{-1}(U_i \cap U_j)}$$

are PL. A PL space is a PL *manifold* (with or without boundary) if the charts X_i can be taken to be open subsets of $\mathbb{R}^n$ or the half-space $\mathbb{R}^{n-1} \times \mathbb{R}^{\geq 0}$.

A subset P of a PL space Y is called a polyhedron if for each of the charts $\varphi_i : X_i \to U_i \subset Y$, the preimage $\varphi_i^{-1}(P) \subset X_i \subset \mathbb{R}^{m_i}$ is a polyhedron.

If $X \subset \mathbb{R}^n$ is a polyhedron and $x \in X$, one may always find a link and star for x that are polyhedra [Rourke and Sanderson 1972, p.5]. It then follows from the definitions that if Y is a PL space, any point y of Y has a neighborhood S contained in some $U_i \ni y$, such that the preimage $\varphi_i^{-1}(S) \subset X_i$ is both a polyhedron and a star of $\varphi_i^{-1}(y)$ in X_i. We will refer to such an S as a *polyhedral star* of y.

The key fact we need is that if X is a polyhedron and $x \in X$, then any two polyhedral stars of x in X are PL-homeomorphic, in other words the star is a PL-homeomorphism invariant of x [Rourke and Sanderson 1972, pp.20–21]. It follows from the above discussion that the same is true in any PL space.

If Y is a PL manifold, one may take each chart X_i to be an open subset in $\mathbb{R}^n$ or $\mathbb{R}^{n-1} \times \mathbb{R}^{\geq 0}$. In any open subset of $\mathbb{R}^n$, the star of a point $(x_1, \dots, x_n)$ may be taken to be the cube $[x_1 - \varepsilon, x_1 + \varepsilon] \times \cdots \times [x_n - \varepsilon, x_n + \varepsilon]$ for sufficiently small $\varepsilon > 0$; and in $\mathbb{R}^{n-1} \times \mathbb{R}^{\geq 0}$ it can be taken to be the intersection of this cube with the closed half-space $\{x_n \geq 0\}$. In all cases, this is topologically a closed ball. It then follows from the fact quoted in the previous paragraph that *every* polyhedral star in a PL manifold is topologically a ball.

The "if" direction of Theorem 1.2 will be proven by quoting the result of Lange mentioned above to show that if G is generated by transpositions, double transpositions, and 3-cycles, then Δ/G is a polyhedral star of a point in a PL manifold, and therefore a ball. Thus it meets the homological vanishing criterion described above, regardless of the field k.

2C. ***Generalities about group actions on a ring.*** The purpose of this section is to develop the commutative algebra needed to prove the general results in Section 3, which are then used in section Section 4B to prove the "only if" direction of Theorem 1.2.

Let 1 denote the group identity. (In commutative diagrams, let it also denote a trivial group.) Let A^G denote the ring of invariants, and similarly for any subgroup of G. It is well known that A is always integral over A^G [Bourbaki 1964, Chapitre V §1.9, Proposition 22].

Let $\mathfrak{P} \lhd A$ be a prime ideal.

Recall that the *decomposition group* $D_G(\mathfrak{P})$ of $\mathfrak{P}$ is the stabilizer of $\mathfrak{P}$ in G:

$$D_G(\mathfrak{P}) = \{g \in G : g\mathfrak{P} = \mathfrak{P}\}.$$

The decomposition group acts on the integral domain $A/\mathfrak{P}$. The *inertia group* $I_G(\mathfrak{P})$ of $\mathfrak{P}$ is the kernel of this action:

$$I_G(\mathfrak{P}) = \{g \in G : (g-1)A \subset \mathfrak{P}\},$$

where

$$(g-1)A = \{ga - a : a \in A\}.$$

The notations $I_G(\mathfrak{P})$ and $D_G(\mathfrak{P})$ implicitly specify the ring A being acted on by G, since $\mathfrak{P}$ belongs to A.

We recall some basic facts in this setup [Bourbaki 1964, Chapitre V §2.2, Théorème 2], which we use freely in what follows: (i) G acts transitively on the prime ideals of A lying over $\mathfrak{P}^\star = \mathfrak{P} \cap A^G$ and (ii) the extension of residue fields $\kappa(\mathfrak{P})/\kappa(\mathfrak{P}^\star)$ is a normal field extension, and the canonical map from $D_G(\mathfrak{P})$ to the group of $\kappa(\mathfrak{P}^\star)$-automorphisms of $\kappa(\mathfrak{P})$ is a surjection with kernel $I_G(\mathfrak{P})$, i.e., the sequence

$$1 \to I_G(\mathfrak{P}) \to D_G(\mathfrak{P}) \to \mathrm{Aut}_{\kappa(\mathfrak{P}^\star)}(\kappa(\mathfrak{P})) \to 1$$

is exact.

If $N \triangleleft G$ is a normal subgroup, then the quotient group G/N acts on the invariant ring A^N, and the decomposition and inertia groups in G and G/N relate straightforwardly. Note that, by their definitions,

$$I_N(\mathfrak{P}) = I_G(\mathfrak{P}) \cap N \quad \text{and} \quad D_N(\mathfrak{P}) = D_G(\mathfrak{P}) \cap N.$$

Lemma 2.10. *We have*

$$D_{G/N}(\mathfrak{P} \cap A^N) \cong D_G(\mathfrak{P})/D_N(\mathfrak{P})$$

and

$$I_{G/N}(\mathfrak{P} \cap A^N) \cong I_G(\mathfrak{P})/I_N(\mathfrak{P}).$$

We believe this and the next lemma may be well known; however, as we were unable to locate references, we include full proofs.

Proof. The sequences

$$1 \to D_N(\mathfrak{P}) \to D_G(\mathfrak{P}) \xrightarrow{\varphi} D_{G/N}(\mathfrak{P} \cap A^N) \to 1$$

and

$$1 \to I_N(\mathfrak{P}) \to I_G(\mathfrak{P}) \xrightarrow{\psi} I_{G/N}(\mathfrak{P} \cap A^N) \to 1$$

are exact in the first and second positions by the definitions; we have to prove surjectivity of φ and ψ.

Consider φ first. Suppose $g \in G$ is such that its image $\bar{g}$ in G/N lies in $D_{G/N}(\mathfrak{P} \cap A^N)$. Then, setting $\mathfrak{Q} = g\mathfrak{P}$, we have

$$\mathfrak{Q} \cap A^N = \mathfrak{P} \cap A^N.$$

All primes of A that intersect A^N in $\mathfrak{P} \cap A^N$ lie in the same orbit of N. Thus there exists $n \in N$ with $n\mathfrak{Q} = \mathfrak{P}$. Therefore $ng\mathfrak{P} = \mathfrak{P}$, i.e., $ng \in D_G(\mathfrak{P})$, and we have $\varphi(ng) = \bar{g}$. So φ is surjective.

We establish the surjectivity of ψ with a diagram chase. Let $\mathfrak{P}' = \mathfrak{P} \cap A^N$ and let $\mathfrak{P}^\star = \mathfrak{P} \cap A^G$. We have the following commutative diagram:

$$\begin{tikzcd}
 & 1 \arrow[d] & 1 \arrow[d] & 1 \arrow[d] & \\
1 \arrow[r] & I_N(\mathfrak{P}) \arrow[r] \arrow[d] & I_G(\mathfrak{P}) \arrow[r, "\psi"] \arrow[d, "i_G"] & I_{G/N}(\mathfrak{P}') \arrow[d, "i_{G/N}"] & \\
1 \arrow[r] & D_N(\mathfrak{P}) \arrow[r, "\iota_D"] \arrow[d, "p_N"'] & D_G(\mathfrak{P}) \arrow[r, "\varphi"] \arrow[d, "p_G"] & D_{G/N}(\mathfrak{P}') \arrow[r] \arrow[d, "p_{G/N}"] & 1 \\
1 \arrow[r] & \mathrm{Aut}_{\kappa(\mathfrak{P}')}(\kappa(\mathfrak{P})) \arrow[r, "\iota_\kappa"'] \arrow[d] & \mathrm{Aut}_{\kappa(\mathfrak{P}^\star)}(\kappa(\mathfrak{P})) \arrow[r, "\xi"'] \arrow[d] & \mathrm{Aut}_{\kappa(\mathfrak{P}^\star)}(\kappa(\mathfrak{P}')) \arrow[r] \arrow[d] & 1 \\
 & 1 & 1 & 1 &
\end{tikzcd}$$

where $\kappa(\mathfrak{P})$, $\kappa(\mathfrak{P}')$, and $\kappa(\mathfrak{P}^\star)$ are the residue fields. The first and second row are exact by what we have just done. The third row is exact by consideration of the definitions and the fact that $\kappa(\mathfrak{P})$ is normal over $\kappa(\mathfrak{P}')$ (by [Bourbaki 1964, Chapitre V §2.2, Théorème 2(ii)], as recalled above), since field automorphisms always extend to normal extensions. The columns are also exact by the same theorem.

Let $g \in I_{G/N}(\mathfrak{P}')$ be arbitrary and consider $i_{G/N}(g)$. Since φ is surjective, there is a $y \in D_G(\mathfrak{P})$ with $\varphi(y) = i_{G/N}(g)$. Then

$$1 = p_{G/N} \circ i_{G/N}(g) = p_{G/N} \circ \varphi(y) = \xi \circ p_G(y),$$

so that $p_G(y) \in \ker \xi = \operatorname{im} \iota_\kappa$. Thus there is a $z \in \mathrm{Aut}_{\kappa(\mathfrak{P}')}(\kappa(\mathfrak{P}))$ with $\iota_\kappa(z) = p_G(y)$. Since p_N is surjective, we have a $z' \in D_N(\mathfrak{P})$ with $p_N(z') = z$. Now consider

$$y^\star = \iota_D(z')^{-1} y \in D_G(\mathfrak{P}).$$

We have

$$p_G(y^\star) = p_G \circ \iota_D(z')^{-1} p_G(y) = \iota_\kappa \circ p_N(z')^{-1} p_G(y) = \iota_\kappa(z)^{-1} p_G(y) = p_G(y)^{-1} p_G(y) = 1.$$

Thus $y^\star \in \ker p_G = \operatorname{im} i_G$, so there exists $g' \in I_G(\mathfrak{P})$ with $i_G(g') = y^\star$. Then

$$i_{G/N} \circ \psi(g') = \varphi \circ i_G(g') = \varphi(y^\star) = \varphi(\iota_D(z')^{-1} y) = \varphi \circ \iota_D(z')^{-1} \varphi(y) = 1^{-1} i_{G/N}(g) = i_{G/N}(g).$$

Since $i_{G/N}$ is injective, we can conclude $\psi(g') = g$. Thus ψ is surjective. □

The inertia group of a prime that survives a base change remains stable under that base change, and the decomposition group can only shrink:

Lemma 2.11. *Let C be an arbitrary A^G-algebra, and let $B := A \otimes_{A^G} C$. Let G act on B through its action on A and trivial action on C. If there is a prime $\mathfrak{Q}$ of B pulling back to $\mathfrak{P}$ in A, then $D_G(\mathfrak{Q}) \subset D_G(\mathfrak{P})$, and $I_G(\mathfrak{Q}) = I_G(\mathfrak{P})$.*

Proof. Let $\tau : A \to B$ be the canonical map. By construction, τ is G-equivariant. Thus if $g \in G$ stabilizes $\mathfrak{Q} \lhd B$ setwise, it also stabilizes the preimage $\mathfrak{P} \lhd A$ setwise, and it follows that $D_G(\mathfrak{Q}) \subset D_G(\mathfrak{P})$.

When $g \in D_G(\mathfrak{Q})$ and therefore $\in D_G(\mathfrak{P})$, it has an induced action on both $B/\mathfrak{Q}$ and $A/\mathfrak{P}$, and the G-equivariance of τ then implies that the induced map

$$\bar{\tau} : A/\mathfrak{P} \to B/\mathfrak{Q}$$

is $\langle g\rangle$-equivariant. If also $g \in I_G(\mathfrak{Q})$, then its action on $B/\mathfrak{Q}$ is trivial. Since $\mathfrak{P}$ is the full preimage of $\mathfrak{Q}$, $\bar{\tau}$ is an injective map, and it follows that g's action on $A/\mathfrak{P}$ is also trivial, i.e., $g \in I_G(\mathfrak{P})$. Thus $I_G(\mathfrak{Q}) \subset I_G(\mathfrak{P})$.

In the other direction, suppose $g \in I_G(\mathfrak{P})$. By [Liu 2002, Chapter 1, Corollary 1.13], we have a canonical isomorphism

$$B/\tau(\mathfrak{P})B \cong A/\mathfrak{P} \otimes_{A^G} C. \tag{3}$$

Using only the fact that $g \in D_G(\mathfrak{P})$ and the G-equivariance of τ, we already know that g fixes $\mathfrak{P}$ and $\tau(\mathfrak{P})$ setwise, and thus has well-defined actions on $A/\mathfrak{P}$ and $B/\tau(\mathfrak{P})B$ that coincide via (3). But because g is actually in $I_G(\mathfrak{P})$, the action on $A/\mathfrak{P}$ is trivial, and therefore, by (3), the action of g on $B/\tau(\mathfrak{P})B$ is also trivial.

In other words, g fixes the cosets of the additive subgroup $\tau(\mathfrak{P})B$ of B setwise. Since $\mathfrak{Q}$ pulls back to $\mathfrak{P}$, it contains the image of $\mathfrak{P}$, thus we have $\mathfrak{Q} \supset \tau(\mathfrak{P})B$. Then the cosets of $\mathfrak{Q}$ are unions of cosets of $\tau(\mathfrak{P})B$, and therefore g fixes these setwise as well. In other words, g acts trivially on $B/\mathfrak{Q}$, i.e., $g \in I_G(\mathfrak{Q})$. Thus $I_G(\mathfrak{P}) \subset I_G(\mathfrak{Q})$, and we conclude $I_G(\mathfrak{P}) = I_G(\mathfrak{Q})$. □

Remark 2.12. Examining the proof of Lemma 2.11, we see why the analogous equality to $I_G(\mathfrak{P}) = I_G(\mathfrak{Q})$ may fail for decomposition groups. If $g \in D_G(\mathfrak{P})$, then we do have the $\langle g\rangle$-equivariant isomorphism (3), and therefore g does act on the cosets of $\tau(\mathfrak{P})B$ in B, but the only one we know it fixes is $\tau(\mathfrak{P})B$ itself. In particular, $\mathfrak{Q}$, which may be the union of many of these cosets, need not be fixed setwise, so that $g \notin D_G(\mathfrak{Q})$.

Henceforth, let $\mathfrak{p}$ be a prime of A^G. Our goal is to show that, in a suitable sense, the local structure of A^G at $\mathfrak{p}$ is determined by the inertia group of a prime of A lying over $\mathfrak{p}$. The precise statement is Lemma 2.14 below. It is stated by Michel Raynaud [1970, Chapitre X §1, Corollaire 1], with lines of proof indicated. Because it is central to our results, we develop in detail the notation and tools that will be required to state and prove this lemma.

Let $C_{\mathfrak{p}}^{\text{hs}}$ be the strict henselization (see [Raynaud 1970, Chapitre VIII, Definition 4] or [EGA I 1960, Definition 18.8.7]) of A^G at $\mathfrak{p}$, with respect to some embedding of $\kappa(\mathfrak{p})$ in its separable closure. Then $C_{\mathfrak{p}}^{\text{hs}}$ is faithfully flat over $(A^G)_{\mathfrak{p}}$, and of relative dimension zero [EGA I 1960, Proposition 18.8.8(iii)]. Furthermore, $C_{\mathfrak{p}}^{\text{hs}}$ and $(A^G)_{\mathfrak{p}}$ are simultaneously noetherian [EGA I 1960, Proposition 18.8.8(iv)], and

$$A_{\mathfrak{p}}^{\text{hs}} := A \otimes_{A^G} C_{\mathfrak{p}}^{\text{hs}}$$

is integral over $C_{\mathfrak{p}}^{\mathrm{hs}}$ (as it is a base change of the integral morphism $A^G \to A$). Moreover, G acts on $A_{\mathfrak{p}}^{\mathrm{hs}}$ via the first component of the tensor product, so that the map $A \to A_{\mathfrak{p}}^{\mathrm{hs}}$ is G-equivariant, and

$$(A_{\mathfrak{p}}^{\mathrm{hs}})^G = C_{\mathfrak{p}}^{\mathrm{hs}}$$

since $A^G \to (A^G)_{\mathfrak{p}} \to C_{\mathfrak{p}}^{\mathrm{hs}}$ is flat and the functor of invariants commutes with flat base change.

Let $\mathfrak{P}$ be a prime ideal of A lying over $\mathfrak{p}$, and let $\mathfrak{Q}$ be a prime ideal of $A_{\mathfrak{p}}^{\mathrm{hs}}$ lying over the maximal ideal of $C_{\mathfrak{p}}^{\mathrm{hs}}$ corresponding to $\mathfrak{p}$, and pulling back to $\mathfrak{P}$ in A.

From Lemma 2.11, we have that

$$I_G(\mathfrak{Q}) = I_G(\mathfrak{P}).$$

The action of G on $A_{\mathfrak{p}}^{\mathrm{hs}}$ induces an action on its ideals. Since $A_{\mathfrak{p}}^{\mathrm{hs}}$ is integral over $C_{\mathfrak{p}}^{\mathrm{hs}}$, all of its maximal ideals lie over the one maximal of $C_{\mathfrak{p}}^{\mathrm{hs}}$. Because $C_{\mathfrak{p}}^{\mathrm{hs}}$ is the invariant ring under the action of G, this implies [Bourbaki 1964, Chapitre V §2.2, Théorème 2(i)] that the maximal ideals of $A_{\mathfrak{p}}^{\mathrm{hs}}$ comprise a single orbit for the action on ideals. The maximals are therefore finite in number. We denote them by $\mathfrak{M}_1 (= \mathfrak{Q}), \dots, \mathfrak{M}_s$.

The product of canonical localization homomorphisms

$$\phi : A_{\mathfrak{p}}^{\mathrm{hs}} \to \prod_{j=1}^{s} (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{M}_j} \tag{4}$$

is an isomorphism. Indeed, $A_{\mathfrak{p}}^{\mathrm{hs}}$ is the inductive limit of $C_{\mathfrak{p}}^{\mathrm{hs}}$-finite subalgebras (since it is integral over $C_{\mathfrak{p}}^{\mathrm{hs}}$). Since $A_{\mathfrak{p}}^{\mathrm{hs}}$ has only s maximals, there exists a finite subalgebra containing s maximals. Now view $A_{\mathfrak{p}}^{\mathrm{hs}}$ as the inductive limit just of the finite subalgebras that contain this one. For each of them, the analogous product of canonical localization morphisms is an isomorphism because $C_{\mathfrak{p}}^{\mathrm{hs}}$ is henselian [Raynaud 1970, Chapitre I, §1 Définition 1 and Proposition 3]; then the statement about (4) follows because inductive limits commute with finite products.

Lemma 2.13. *If A is a noetherian ring, then $A_{\mathfrak{p}}^{\mathrm{hs}}$ is noetherian too.*

Proof. Because of the isomorphism (4), it suffices to show that the localizations of $A_{\mathfrak{p}}^{\mathrm{hs}}$ at its maximal ideals $\mathfrak{M}_j$ are noetherian rings, and because the action of G on $A_{\mathfrak{p}}^{\mathrm{hs}}$ by automorphisms is transitive on these maximals, it suffices to show this for a single maximal. We will do this by showing that there is a maximal ideal $\mathfrak{M}_j$ of $A_{\mathfrak{p}}^{\mathrm{hs}}$ such that

$$(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{M}_j}$$

is isomorphic to the strict henselization of the noetherian local ring $A_{\mathfrak{P}}$, whereupon the result will follow because strict henselization preserves noetherianity [EGA I 1960, Proposition 18.8.8(iv)].

Consider the local ring $(A^G)_{\mathfrak{p}}$. By slight abuse of notation, let us call its maximal ideal $\mathfrak{p}$. Note that the residue field $\kappa(\mathfrak{p})$ is the same whether $\mathfrak{p}$ refers to the prime in A^G or in $(A^G)_{\mathfrak{p}}$, so we can write $\kappa(\mathfrak{p})$ without ambiguity. Then the maximal ideals in the ring

$$B := A \otimes_{A^G} (A^G)_{\mathfrak{p}}$$

are in bijection with the prime ideals of A lying over $\mathfrak{p} \lhd A^G$. There are finitely many of these since they are subject to a transitive action by G, so B is semilocal. It is also integral as an extension of $(A^G)_{\mathfrak{p}}$ since this is a base change of the integral extension $A^G \subset A$. One of the prime ideals over $\mathfrak{p}$ in A is $\mathfrak{P}$. By the same abuse of notation, let $\mathfrak{P}$ also refer to the corresponding ideal in B; again, this does not introduce ambiguity when writing $\kappa(\mathfrak{P})$. Note that $B_{\mathfrak{P}} = A_{\mathfrak{P}}$ because B is obtained from A by inverting some but not all of the elements in the complement of $\mathfrak{P}$.

Because B is semilocal and integral over $(A^G)_{\mathfrak{p}}$ (and $\mathfrak{P}$ and $\mathfrak{p}$ are maximal ideals of these rings respectively), if we can show that the extension of residue fields $\kappa(\mathfrak{P})/\kappa(\mathfrak{p})$ has finite separable degree, then it will follow from [EGA I 1960, Proposition 18.8.10 and its proof, and Remarque 18.8.11] that the strict henselization

$$(B_{\mathfrak{P}})^{\mathrm{hs}}$$

of the localization $B_{\mathfrak{P}}$ (with respect to some embedding of its residue field in a separable closure) is isomorphic to the localization of

$$B \otimes_{(A^G)_{\mathfrak{p}}} C_{\mathfrak{p}}^{\mathrm{hs}}$$

at some maximal ideal, since $C_{\mathfrak{p}}^{\mathrm{hs}}$ is a strict henselization of $(A^G)_{\mathfrak{p}}$. But we also have

$$B \otimes_{(A^G)_{\mathfrak{p}}} C_{\mathfrak{p}}^{\mathrm{hs}} = A \otimes_{A^G} (A^G)_{\mathfrak{p}} \otimes_{(A^G)_{\mathfrak{p}}} C_{\mathfrak{p}}^{\mathrm{hs}} = A \otimes_{A^G} C_{\mathfrak{p}}^{\mathrm{hs}} = A_{\mathfrak{p}}^{\mathrm{hs}}.$$

Thus the conclusion from [EGA I 1960, 18.8.10 and 18.8.11] will actually be that

$$(A_{\mathfrak{P}})^{\mathrm{hs}} = (B_{\mathfrak{P}})^{\mathrm{hs}} \cong (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{M}_j}$$

for some maximal ideal $\mathfrak{M}_j$ of $A_{\mathfrak{p}}^{\mathrm{hs}}$. This is the desired conclusion, so it remains to show that $\kappa(\mathfrak{P})/\kappa(\mathfrak{p})$ has finite separable degree.

Now return $\mathfrak{p}$ and $\mathfrak{P}$ to the setting of A^G and A, recalling that the residue fields $\kappa(\mathfrak{p})$ and $\kappa(\mathfrak{P})$ do not change. From [Bourbaki 1964, Chapitre V, §2.2(ii)] we have that $\kappa(\mathfrak{P})/\kappa(\mathfrak{p})$ is a normal field extension, and the group of $\kappa(\mathfrak{p})$-automorphisms of $\kappa(\mathfrak{P})$ is isomorphic to

$$D_G(\mathfrak{P})/I_G(\mathfrak{P}).$$

This is a subquotient of the finite group G and is therefore finite. For a normal field extension, infinite separable degree would imply infinitely many automorphisms. Thus $\kappa(\mathfrak{P})/\kappa(\mathfrak{p})$ is an extension of finite separable degree, and the proof is complete. □

The action of G on $A_{\mathfrak{p}}^{\mathrm{hs}}$ induces, via the isomorphism ϕ of (4), an action on $\prod_1^s (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{M}_j}$: it is the unique action on this ring such that ϕ is G-equivariant. Because ϕ is the product of the canonical localization maps

$$\phi_j : A_{\mathfrak{p}}^{\mathrm{hs}} \to (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{M}_j},$$

it is possible to write down this action explicitly. Via the isomorphism ϕ of (4) we associate uniquely to $a \in A^{\mathrm{hs}}_{\mathfrak{p}}$ the s-tuple

$$\phi(a) = (a_{\mathfrak{M}_1}, \dots, a_{\mathfrak{M}_s}) \in \prod_{j=1}^{s} (A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_j} \tag{5}$$

where each $a_{\mathfrak{M}_j}$ is the image in $(A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_j}$ of a under ϕ_j. If $g \in G$ maps $\mathfrak{M}_i$ to $\mathfrak{M}_j$, then it also induces an isomorphism

$$\begin{aligned} (A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_i} &\xrightarrow{g} (A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_j} \\ a/s &\mapsto ga/gs \end{aligned}$$

of the localizations that makes the following square commutative:

$$\begin{tikzcd} A^{\mathrm{hs}}_{\mathfrak{p}} \arrow[r, "g"] \arrow[d, "\phi_i"'] & A^{\mathrm{hs}}_{\mathfrak{p}} \arrow[d, "\phi_j"] \\ (A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_i} \arrow[r, "g"'] & (A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_j} \end{tikzcd}$$

By such isomorphisms, G acts on the disjoint union of the localizations $(A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_j}$. Given an $\alpha \in (A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_i}$, if one chooses $a \in A^{\mathrm{hs}}_{\mathfrak{p}}$ with $\phi_i(a) = \alpha$, then the commutativity of this square can be rewritten as

$$g\alpha = \phi_j(ga).$$

Note that this statement is true regardless of the choice of a. For any such choice, writing $\alpha = a_{\mathfrak{M}_i}$ and $\phi_j(ga) = (ga)_{\mathfrak{M}_j} = (ga)_{g(\mathfrak{M}_i)}$, this becomes

$$g(a_{\mathfrak{M}_i}) = (ga)_{g(\mathfrak{M}_i)},$$

or equivalently,

$$g(a_{g^{-1}(\mathfrak{M}_j)}) = (ga)_{\mathfrak{M}_j}. \tag{6}$$

Thus, for any $a \in A^{\mathrm{hs}}_{\mathfrak{p}}$, the i-th coordinate of $\phi(a)$ determines the j-th coordinate of $\phi(ga)$, without requiring additional information about a. Then the action of G on $\prod_1^s (A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{M}_j}$ induced by ϕ may be written

$$g(a_{\mathfrak{M}_1}, \dots, a_{\mathfrak{M}_s}) = (g(a_{g^{-1}(\mathfrak{M}_1)}), \dots, g(a_{g^{-1}(\mathfrak{M}_s)})). \tag{7}$$

Indeed, if $a \in A^{\mathrm{hs}}_{\mathfrak{p}}$, then the left side of this formula is $g\phi(a)$, and the right side is $\phi(ga)$ by (6).

Because $I_G(\mathfrak{Q})$ stabilizes $\mathfrak{Q} = \mathfrak{M}_1$, it acts on $(A^{\mathrm{hs}}_{\mathfrak{p}})_{\mathfrak{Q}}$. In this setting, we have the following lemma. As mentioned above, this lemma was stated by Michel Raynaud [1970, Chapitre X §1, Corollaire 1], with the proof sketched. It is the needed statement that the local structure of A^G is determined by the inertia groups. Because it is critical to our results, we give a detailed proof.

Lemma 2.14 (Raynaud). *We have a ring isomorphism*

$$(A^{\mathrm{hs}}_{\mathfrak{p}})^{I_G(\mathfrak{P})}_{\mathfrak{Q}} \cong C^{\mathrm{hs}}_{\mathfrak{p}}.$$

Proof. Recall that $I_G(\mathfrak{P}) = I_G(\mathfrak{Q})$. Let $g_1, \dots, g_s \in G$ be a set of left coset representatives for $G/I_G(\mathfrak{P})$, with g_1 the identity. Since $C_{\mathfrak{p}}^{\text{hs}}$ is strictly henselian, its residue field is separably closed, so there are no nontrivial automorphisms of $\kappa(\mathfrak{Q})$ over it. Since the group of automorphisms of $\kappa(\mathfrak{Q})/\kappa(C_{\mathfrak{p}}^{\text{hs}})$ is isomorphic to $D_G(\mathfrak{Q})/I_G(\mathfrak{Q})$, we have $D_G(\mathfrak{Q}) = I_G(\mathfrak{Q})$, so that $I_G(\mathfrak{Q})$, which equals $I_G(\mathfrak{P})$, is the stabilizer of $\mathfrak{Q}$. Thus, if we put $\mathfrak{M}_j := g_j\mathfrak{Q}$, then the ideals $\mathfrak{M}_1, \dots, \mathfrak{M}_s$ are exactly the maximal ideals of $A_{\mathfrak{p}}^{\text{hs}}$, and all of the above discussion applies.

We claim that if one restricts the canonical localization map

$$\phi_1 : A_{\mathfrak{p}}^{\text{hs}} \to (A_{\mathfrak{p}}^{\text{hs}})_{\mathfrak{Q}}$$

to $C_{\mathfrak{p}}^{\text{hs}}$, one obtains an isomorphism onto $(A_{\mathfrak{p}}^{\text{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{Q})}$. We see this as follows:

The map ϕ_1 is the composition of ϕ with projection to the first coordinate. Because (7) makes ϕ a G-equivariant isomorphism, $a \in A_{\mathfrak{p}}^{\text{hs}}$ is in $C_{\mathfrak{p}}^{\text{hs}} = (A_{\mathfrak{p}}^{\text{hs}})^G$ if and only if

$$(g(a_{g^{-1}(\mathfrak{M}_1)}), \dots, g(a_{g^{-1}(\mathfrak{M}_s)})) = (a_{\mathfrak{M}_1}, \dots, a_{\mathfrak{M}_s}) \tag{8}$$

for all $g \in G$. From (8), we will deduce the following:

(a) If $a \in C_{\mathfrak{p}}^{\text{hs}}$ is an arbitrary G-invariant, then $\phi_1(a)$ is invariant under $I_G(\mathfrak{P})$. Thus $\phi_1(C_{\mathfrak{p}}^{\text{hs}})$ is contained in $(A_{\mathfrak{p}}^{\text{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}$.

(b) If $a \in C_{\mathfrak{p}}^{\text{hs}}$ is an arbitrary G-invariant, then all the coordinates of $\phi(a)$ are determined by the first coordinate. Thus a itself is determined by $\phi_1(a)$. In other words, the restriction of ϕ_1 to $C_{\mathfrak{p}}^{\text{hs}}$ is injective.

(c) If $\alpha \in (A_{\mathfrak{p}}^{\text{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}$ is arbitrary, there exists an $a \in C_{\mathfrak{p}}^{\text{hs}}$ with $\phi_1(a) = \alpha$. Thus the restriction of ϕ_1 to $C_{\mathfrak{p}}^{\text{hs}}$ is surjective.

This will suffice to establish the lemma.

To prove (a), take $g \in I_G(\mathfrak{P})$. The condition in the first coordinate of (8) is

$$g(a_{g^{-1}(\mathfrak{M}_1)}) = a_{\mathfrak{M}_1}.$$

For $g \in I_G(\mathfrak{P}) = D_G(\mathfrak{Q})$, we have $g^{-1}(\mathfrak{M}_1) = \mathfrak{M}_1 = \mathfrak{Q}$, and this condition becomes

$$g(a_{\mathfrak{Q}}) = a_{\mathfrak{Q}}.$$

Thus for the G-invariant a, we have that $a_{\mathfrak{Q}} = \phi_1(a)$ is an $I_G(\mathfrak{P})$-invariant. Therefore, $\phi_1(C_{\mathfrak{p}}^{\text{hs}})$ is contained in $(A_{\mathfrak{p}}^{\text{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}$.

For (b), consider $g = g_j$ for $j = 1, \dots, s$. The condition in the jth coordinate of (8) is

$$g(a_{g^{-1}(\mathfrak{M}_j)}) = a_{\mathfrak{M}_j}.$$

Since $g_j^{-1}(\mathfrak{M}_j) = \mathfrak{Q}$, this becomes

$$g_j(a_{\mathfrak{Q}}) = a_{\mathfrak{M}_j}.$$

Letting $j = 1, \dots, s$, this shows that if a is a G-invariant, then all the coordinates of $\phi(a)$ are determined by $a_{\mathfrak{Q}}$, which is $\phi_1(a)$, so a itself is determined by $\phi_1(a)$. Therefore, the restriction of ϕ_1 to $C_{\mathfrak{p}}^{\mathrm{hs}}$ is injective.

Lastly, for (c), let $\alpha \in (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}$ be arbitrary. We construct a specific $a \in A_{\mathfrak{p}}^{\mathrm{hs}}$ with $\phi_1(a) = \alpha$, and show it lies in $C_{\mathfrak{p}}^{\mathrm{hs}}$. Set

$$a_{\mathfrak{M}_j} := g_j(\alpha)$$

for $j = 1, \dots, s$, and let

$$a := \phi^{-1}(a_{\mathfrak{M}_1}, \dots, a_{\mathfrak{M}_s}) \in A_{\mathfrak{p}}^{\mathrm{hs}}.$$

Note that this a satisfies $\phi_1(a) = a_{\mathfrak{M}_1} = g_1(\alpha) = \alpha$ since g_1 is the identity. To show that it also lies in $C_{\mathfrak{p}}^{\mathrm{hs}} = (A_{\mathfrak{p}}^{\mathrm{hs}})^G$, it is necessary and sufficient to show that $\phi(a)$ satisfies (8) for all $g \in G$, i.e., that

$$g(a_{g^{-1}(\mathfrak{M}_j)}) = a_{\mathfrak{M}_j} \tag{9}$$

for all $g \in G$ and all $j = 1, \dots, s$.

To do this, we first establish that

$$a_{g(\mathfrak{Q})} = g(a_{\mathfrak{Q}}) \tag{10}$$

for all $g \in G$, and then use this to show (9) for all g and all j.

To see (10), first recall that $\alpha = a_{\mathfrak{M}_1} = a_{\mathfrak{Q}}$, and then use this and $\mathfrak{M}_j = g_j(\mathfrak{Q})$ to rewrite the definition of each $a_{\mathfrak{M}_j}$:

$$a_{g_j(\mathfrak{Q})} = g_j(a_{\mathfrak{Q}}).$$

This establishes (10) in the particular case that g is one of $g_1, \dots, g_s$. An arbitrary $g \in G$ has the form $g_j h$ for some g_j and some $h \in I_G(\mathfrak{P})$. Since $\mathfrak{Q}$ and $a_{\mathfrak{Q}} = \alpha$ are both $I_G(\mathfrak{P})$-invariant, we have

$$a_{g(\mathfrak{Q})} = a_{g_j h(\mathfrak{Q})} = a_{g_j(\mathfrak{Q})} = g_j(a_{\mathfrak{Q}}) = g_j h(a_{\mathfrak{Q}}) = g(a_{\mathfrak{Q}}),$$

and (10) is established for all $g \in G$.

Now we deduce (9). If $g \in G$ is arbitrary, then

$$a_{g^{-1}(\mathfrak{M}_j)} = a_{g^{-1}g_j(\mathfrak{Q})}$$

because $g_j(\mathfrak{Q}) = \mathfrak{M}_j$, and

$$a_{g^{-1}g_j(\mathfrak{Q})} = g^{-1}g_j(a_{\mathfrak{Q}})$$

by (10). Thus $a_{g^{-1}(\mathfrak{M}_j)} = g^{-1}g_j(a_{\mathfrak{Q}})$, and applying g to the left on both sides yields

$$g(a_{g^{-1}(\mathfrak{M}_j)}) = g_j(a_{\mathfrak{Q}}) = a_{\mathfrak{M}_j},$$

so condition (9) is met for all g and all j, i.e., (8) is met for all g. Thus

$$a \in (A_{\mathfrak{p}}^{\mathrm{hs}})^G = C_{\mathfrak{p}}^{\mathrm{hs}}.$$

Since $\alpha \in (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}$ was arbitrary, this shows that the restriction of ϕ_1 to $C_{\mathfrak{p}}^{\mathrm{hs}}$ is surjective onto $(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}$, completing the proof of isomorphism. □

3. Inertia groups and Cohen–Macaulayness of invariant rings

Using Lemma 2.14, we can show that the Cohen–Macaulayness of a ring of invariants at a prime ideal $\mathfrak{p}$ can always be tested in a faithfully flat neighborhood of $\mathfrak{p}$, and only depends on the action of the inertia group considered around this neighborhood. The precise statement is Theorem 3.1.

We use this to derive an obstruction to Cohen–Macaulayness for a characteristic p ring that will apply in the situation of Theorem 1.2 to prove the "only if" direction. The statement is Proposition 3.11.

In all of what follows, we use the notation of Section 2C: A is a commutative, unital ring endowed with a faithful action of a finite group G, if $\mathfrak{p}$ is a prime ideal of A^G, then $C_{\mathfrak{p}}^{\mathrm{hs}}$ is the strict henselization of A^G at $\mathfrak{p}$, and $A_{\mathfrak{p}}^{\mathrm{hs}}$ is

$$A \otimes_{A^G} C_{\mathfrak{p}}^{\mathrm{hs}},$$

with G acting through its action on A (and trivially on $C_{\mathfrak{p}}^{\mathrm{hs}}$).

Theorem 3.1. *Assume that A^G is noetherian. Then the following assertions are equivalent*:

(1) *A^G is Cohen–Macaulay.*

(2) *For every prime ideal $\mathfrak{p}$ of A^G, and for every prime ideal $\mathfrak{Q}$ of $A_{\mathfrak{p}}^{\mathrm{hs}}$ lying over $\mathfrak{p}C_{\mathfrak{p}}^{\mathrm{hs}}$ and pulling back to a prime $\mathfrak{P}$ of A lying over $\mathfrak{p}$,*

$$(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}{}^{I_G(\mathfrak{P})}$$

is Cohen–Macaulay.

(3) *For every maximal ideal $\mathfrak{p}$ of A^G, there is some prime ideal $\mathfrak{Q}$ of $A_{\mathfrak{p}}^{\mathrm{hs}}$ lying over $\mathfrak{p}C_{\mathfrak{p}}^{\mathrm{hs}}$ and pulling back to a prime $\mathfrak{P}$ of A lying over $\mathfrak{p}$, such that*

$$(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}{}^{I_G(\mathfrak{P})}$$

is Cohen–Macaulay.

Proof. Clearly (2)⇒(3). We will show that (3)⇒(1) and (1)⇒(2).

(3)⇒(1): Lemma 2.14 states that for each maximal ideal $\mathfrak{p}$ of A^G and for any choice of $\mathfrak{P}$ and $\mathfrak{Q}$ as in (3),

$$C_{\mathfrak{p}}^{\mathrm{hs}} \cong (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}.$$

Thus (3) implies that $C_{\mathfrak{p}}^{\mathrm{hs}}$ is Cohen–Macaulay for each $\mathfrak{p}$. The homomorphism of local noetherian rings

$$(A^G)_{\mathfrak{p}} \to C_{\mathfrak{p}}^{\mathrm{hs}}$$

is flat, so by [Bruns and Herzog 1993, Theorem 2.1.7] quoted in Section 2A, Cohen–Macaulayness of $C_{\mathfrak{p}}^{\mathrm{hs}}$ is equivalent to that of $(A^G)_{\mathfrak{p}}$ plus that of $C_{\mathfrak{p}}^{\mathrm{hs}}/\mathfrak{p}C_{\mathfrak{p}}^{\mathrm{hs}}$. In particular, since $C_{\mathfrak{p}}^{\mathrm{hs}}$ is Cohen–Macaulay, so is $(A^G)_{\mathfrak{p}}$. Since this holds for all maximal ideals $\mathfrak{p}$ of A^G, A^G is Cohen–Macaulay.

(1)⇒(2): Suppose A^G is Cohen–Macaulay. Let $\mathfrak{p}$ be any prime ideal of A^G. It suffices to prove that $C_{\mathfrak{p}}^{\mathrm{hs}}$ is Cohen–Macaulay since, by Lemma 2.14, for any $\mathfrak{P}$ and $\mathfrak{Q}$ as in (2), we have

$$C_{\mathfrak{p}}^{\mathrm{hs}} \cong (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}^{I_G(\mathfrak{P})}.$$

Since $(A^G)_{\mathfrak{p}} \to C_{\mathfrak{p}}^{\mathrm{hs}}$ is flat, we again have by [Bruns and Herzog 1993, Theorem 2.1.7] that the Cohen–Macaulayness of $C_{\mathfrak{p}}^{\mathrm{hs}}$ is equivalent to that of $(A^G)_{\mathfrak{p}}$ plus that of $C_{\mathfrak{p}}^{\mathrm{hs}}/\mathfrak{p}C_{\mathfrak{p}}^{\mathrm{hs}}$. The former ring is Cohen–Macaulay since A^G is, by the hypothesis (1), and the latter is Cohen–Macaulay since it is a field (see Section 2A), namely, the residue field of the local ring $C_{\mathfrak{p}}^{\mathrm{hs}}$. □

Theorem 3.1 allows us to test Cohen–Macaulayness of an invariant ring A^G locally, prime by prime, in terms of the local ring $(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$ and the local group action $I_G(\mathfrak{P})$. For the application we have in mind in Section 4, we will need to carry information about A and G to $(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$ and $I_G(\mathfrak{P})$, so we enunciate a few more lemmas to accomplish this:

Lemma 3.2. *If A is Cohen–Macaulay, then $A_{\mathfrak{p}}^{\mathrm{hs}}$ is Cohen–Macaulay for any prime ideal $\mathfrak{p}$ of A^G.*

Proof. Suppose A is Cohen–Macaulay, thus noetherian, and $\mathfrak{p}$ is a prime of A^G. By Lemma 2.13, $A_{\mathfrak{p}}^{\mathrm{hs}}$ is noetherian.

Let $\mathfrak{Q}$ be any maximal ideal of $A_{\mathfrak{p}}^{\mathrm{hs}}$ and let $\mathfrak{P}$ be its contraction in A. (Note that $\mathfrak{Q}$ lies over $\mathfrak{p}C_{\mathfrak{p}}^{\mathrm{hs}}$, per Section 2C, and therefore $\mathfrak{P}$ lies over $\mathfrak{p}$.) Now

$$A^G \to (A^G)_{\mathfrak{p}} \to C_{\mathfrak{p}}^{\mathrm{hs}}$$

is a flat map. Therefore, base changing by $A^G \to A_{\mathfrak{P}}$,

$$A_{\mathfrak{P}} \to A_{\mathfrak{P}} \otimes_{A^G} C_{\mathfrak{p}}^{\mathrm{hs}} = A_{\mathfrak{P}} \otimes_A A_{\mathfrak{p}}^{\mathrm{hs}}$$

is also a flat map. Since $\mathfrak{Q} \lhd A_{\mathfrak{p}}^{\mathrm{hs}}$ pulls back to $\mathfrak{P}$ in A, $(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$ is a localization of $A_{\mathfrak{P}} \otimes_A A_{\mathfrak{p}}^{\mathrm{hs}}$; thus

$$A_{\mathfrak{P}} \to (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$$

is also flat. Therefore, again by [Bruns and Herzog 1993, Theorem 2.1.7], the Cohen–Macaulayness of $(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$ is equivalent to that of $A_{\mathfrak{P}}$ plus that of $(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}/\mathfrak{P}(A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$. The former is Cohen–Macaulay since A is, while the latter is Cohen–Macaulay since it is an artinian local ring (see Section 2A), which in turn is because $A_{\mathfrak{P}} \to (A_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$ is of relative dimension zero. This itself is because this map is a localization of the base change $A_{\mathfrak{P}} \otimes_{(A^G)_{\mathfrak{p}}} -$ of the map $(A^G)_{\mathfrak{p}} \to C_{\mathfrak{p}}^{\mathrm{hs}}$, which is flat of relative dimension zero because it is a strict henselization [EGA I 1960, Proposition 18.8.8(iii)]. □

For a natural number t, an element $g \in G$ is called a *t-reflection* if the ideal generated by $(g-1)A$ in A is contained in a prime of height $\leq t$. A prime $\mathfrak{P}$ contains $(g-1)A$ if and only if $g \in I_{\mathfrak{P}}(A)$, so another way to say this is that g is a t-reflection if it is in the inertia group of some prime of height $\leq t$.

In the geometric situation (where A is a finitely generated algebra over a field), the ideal generated by $(g-1)A$ corresponds to the fixed point locus of g, so this definition makes a group element a t-reflection if this fixed point locus has codimension at most t. Thus if G is a linear group acting on the coordinate ring of affine space, a 1-reflection is either the identity or a reflection in the classical sense. A 2-reflection has a fixed point locus of codimension 0, 1, or 2. In particular, if G acts by permutations of a basis, then the 2-reflections are exactly the identity, the transpositions, the double transpositions, and the 3-cycles.

Lemma 3.3. *If an element $g \in I_G(\mathfrak{P})$ acts as a t-reflection on $A_{\mathfrak{P}}$, then it acts as a t-reflection on A.*

Proof. Since $g \in I_G(\mathfrak{P})$, we have $(g-1)A \subset \mathfrak{P}$. The primes of A contained in $\mathfrak{P}$ are in containment-preserving bijection with the primes of $A_{\mathfrak{P}}$, with the bijection given by extension along the canonical localization map, and $(g-1)A_{\mathfrak{P}}$ is the image of $(g-1)A$ along this map. Thus if a prime of height t in $A_{\mathfrak{P}}$ contains $(g-1)A_{\mathfrak{P}}$, then its pullback in A is also of height t and contains $(g-1)A$. $\square$

Lemma 3.4. *If A is noetherian, and $g \in I_G(\mathfrak{P}) = I_G(\mathfrak{Q})$ acts as a t-reflection on $A_{\mathfrak{p}}^{\mathrm{hs}}$, then it acts as a t-reflection on A.*

Proof. If g is a t-reflection on $A_{\mathfrak{p}}^{\mathrm{hs}}$, then there is a prime ideal $\mathfrak{S}$ of $A_{\mathfrak{p}}^{\mathrm{hs}}$ of height $\leq t$ and containing $(g-1)A_{\mathfrak{p}}^{\mathrm{hs}}$. Let $\mathfrak{R}$ be $\mathfrak{S}$'s pullback in A. Then $\mathfrak{R}$ contains $(g-1)A$. Since by Section 2C and Lemma 2.13,

$$A \to A_{\mathfrak{p}}^{\mathrm{hs}}$$

is a flat extension of noetherian rings, going-down applies [Eisenbud 1995, Lemma 10.11], so that the height of $\mathfrak{S}$ is at least that of $\mathfrak{R}$. In particular, the height of $\mathfrak{R}$ is $\leq t$, so that g is a t-reflection on A. $\square$

We will also need to take an element of G acting on A but not as a t-reflection, and conclude that it does not act on a certain subring as a t-reflection either:

Lemma 3.5. *If N is the normal subgroup of G generated by the t-reflections, then no element of $G \setminus N$ acts on A^N as a t-reflection.*

Remark 3.6. This lemma does not require a noetherian hypothesis on A.

Proof. Let $g \in G$. We will show that if its image $\bar{g} \in G/N$ acts on A^N as a t-reflection, then actually $g \in N$.

If $\bar{g}$ acts on A^N as a t-reflection, then there is a prime $\mathfrak{p}$ of A^N of height $\leq t$ with $\bar{g} \in I_{G/N}(\mathfrak{p})$. Let $\mathfrak{P}$ be any prime of A lying over $\mathfrak{p}$. The height of $\mathfrak{P}$ is equal to that of $\mathfrak{p}$ (e.g., by [Gordeev and Kemper 2003, Lemma 5.3], which is stated for noetherian A but the argument holds in general); in particular it is $\leq t$. By Lemma 2.10, we have

$$I_{G/N}(\mathfrak{p}) = I_G(\mathfrak{P})/I_N(\mathfrak{P}).$$

In particular, $I_G(\mathfrak{P})$ surjects onto $I_{G/N}(\mathfrak{p})$, so there is an element $g' \in I_G(\mathfrak{P})$ whose image in G/N is $\bar{g}$. Since $\mathfrak{P}$ has height $\leq t$, g' is a t-reflection, so it is contained in N by construction. Then its image $\bar{g}$ must actually be the identity. So g (with the same image) lies in the kernel of $G \to G/N$, i.e., $g \in N$. $\square$

The following lemma allows us to detect a failure of Cohen–Macaulayness locally.

Lemma 3.7. *Let A be a ring containing the prime field $\mathbb{F}_p$, and let G be a p-group. Suppose that A is Cohen–Macaulay, A^G is noetherian, and A is finite over A^G. Further, suppose there is a prime ideal $\mathfrak{P}$ of A such that $G = I_G(\mathfrak{P})$. Then A^G is not Cohen–Macaulay unless G is generated by its 2-reflections.*

Remark 3.8. This statement is closely related to [Gordeev and Kemper 2003, Theorem 5.5], which also applies to non-p-groups and gives some control over how far A^G can be from Cohen–Macaulay. However, a key step in the proof of that result requires the rings to be normal rings that are localizations of algebras finitely generated over fields. As our application will be to rings that do not fulfill this hypothesis, we give an independent proof.

Proof of Lemma 3.7. Let N be the normal subgroup of G generated by the 2-reflections.

Since A is finite over the noetherian ring A^G, it is noetherian as an A^G-module. Since it also contains $\mathbb{F}_p$, [Lorenz and Pathak 2001, Corollary 4.3] applies, which, when specialized to the situation that G is a p-group, states that if both A and A^G are Cohen–Macaulay, then the map

$$\mathrm{Tr}_{G/N} : A^N \to A^G$$

given by

$$x \mapsto \sum_{g \in G/N} gx$$

is surjective onto A^G, where we think of each g as an element of G and the sum is taken over coset representatives of N.

We will show that this map cannot be surjective unless $N = G$. Since A is Cohen–Macaulay by assumption, this will show A^G is not Cohen–Macaulay if $N \neq G$.

If $\mathrm{Tr}_{G/N}$ is surjective, then we have

$$1 = \sum_{g \in G/N} gx$$

for some $x \in A^N$. Since $G = I_G(\mathfrak{P})$, all $g \in G$ satisfy $gx = x \mod \mathfrak{P}$ in A, thus

$$1 = \sum_{g \in G/N} x = [G : N]x \mod \mathfrak{P}$$

in A. Since G is a p-group and A contains $\mathbb{F}_p$, $[G : N]x = 0$ in A unless $N = G$. In particular, $[G : N]x$ cannot be $1 \mod \mathfrak{P}$ unless $N = G$. □

Remark 3.9. The map $\mathrm{Tr}_{G/N}$ is called the *relative trace* or *relative transfer*; see Remark 2.3.

Remark 3.10. The proof uses a result of Lorenz and Pathak [2001, Lemma 4.3], which has as a hypothesis that A is noetherian as an A^G-module; call this ($\star$). Above, we deduced ($\star$) from the assumptions that (1) A^G is noetherian and (2) A is finite over it. Actually, ($\star$) also implies (1) and (2), hence is equivalent to them. Since any ideal of A^G is also an A^G-submodule of A (since A^G embeds in A), ($\star$) implies that all these ideals are finitely generated, thus (1). Meanwhile, A itself is an A^G-submodule of A, so ($\star$) implies it is finitely generated as an A^G-module, thus (2). More generally, if a module M over a ring R has an injective R-module map from R, then noetherianity of M as R-module is equivalent to noetherianity of R as a ring plus finite generation of M over R, by the same arguments.

Combining all of these results, we get an obstruction to Cohen–Macaulayness for a characteristic p ring expressed entirely in terms of the presence of a certain inertia group. The proof of the "only if" direction of Theorem 1.2 will be an application of this proposition.

Proposition 3.11. *Let A be a ring containing $\mathbb{F}_p$ and let G be a finite group of automorphisms of A. Let N be the normal subgroup of G generated by the 2-reflections. Suppose that A^N is Cohen–Macaulay, A^G is noetherian, and A^N is finite over A^G. If there is an inertia group for the action of G/N on A^N that is a nontrivial p-group, then A^G is not Cohen–Macaulay.*

Proof. Note that

$$A^G = (A^N)^{G/N}.$$

Since A^G is noetherian, Theorem 3.1 applies.

Suppose $\mathfrak{P}$ is a prime of A^N whose inertia group $I_{G/N}(\mathfrak{P})$ is a p-group, per the hypothesis. Let

$$\mathfrak{p} = \mathfrak{P} \cap (A^N)^{G/N},$$

let $C_{\mathfrak{p}}^{\mathrm{hs}}$ be the strict henselization of $(A^G)_{\mathfrak{p}} = ((A^N)^{G/N})_{\mathfrak{p}}$, and let

$$(A^N)_{\mathfrak{p}}^{\mathrm{hs}} = A^N \otimes_{A^G} C_{\mathfrak{p}}^{\mathrm{hs}},$$

as in Section 2C.

By assumption, A^N is Cohen–Macaulay. Thus $(A^N)_{\mathfrak{p}}^{\mathrm{hs}}$ is Cohen–Macaulay, by Lemma 3.2, and thus so is

$$((A^N)_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$$

for any $\mathfrak{Q} \lhd (A^N)_{\mathfrak{p}}^{\mathrm{hs}}$, and in particular any $\mathfrak{Q}$ as described in Theorem 3.1.

As A^N is finite over the noetherian ring A^G by assumption, its base change $(A^N)_{\mathfrak{p}}^{\mathrm{hs}}$ is finite over $C_{\mathfrak{p}}^{\mathrm{hs}}$, which is noetherian by [EGA I 1960, Proposition 18.8.8(iv)], as discussed in Section 2C. The localization $((A^N)_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$ is a homomorphic image of $(A^N)_{\mathfrak{p}}^{\mathrm{hs}}$ by the isomorphism (4), so it too is finite over $C_{\mathfrak{p}}^{\mathrm{hs}}$.

By Lemma 2.14, $C_{\mathfrak{p}}^{\mathrm{hs}}$ is the invariant ring for the action of $I_{G/N}(\mathfrak{P})$ on $((A^N)_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$. Since A contains $\mathbb{F}_p$ and therefore so do A^N and $((A^N)_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$, and since $I_{G/N}(\mathfrak{P})$ is a p-group that is equal to $I_{G/N}(\mathfrak{Q})$ which is an inertia group of $((A^N)_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$, we have now verified all the hypotheses of Lemma 3.7 for the action of $I_{G/N}(\mathfrak{P})$ on $((A^N)_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}$. We can conclude from that lemma that the invariant ring cannot be Cohen–Macaulay unless $I_{G/N}(\mathfrak{P})$ is generated by 2-reflections.

However $I_{G/N}(\mathfrak{P})$ is not so generated. By Lemma 3.5, no nontrivial element of G/N acts on A^N as a 2-reflection. In particular, no nontrivial element of $I_{G/N}(\mathfrak{P})$ acts on A^N as a 2-reflection. Since A^N is Cohen–Macaulay, it is noetherian, so Lemma 3.4 applies, and no nontrivial element of $I_{G/N}(\mathfrak{P})$ acts on $(A^N)_{\mathfrak{p}}^{\mathrm{hs}}$ as a 2-reflection either. By Lemma 3.3, the same is true for the action of $I_{G/N}(\mathfrak{P}) = I_{G/N}(\mathfrak{Q})$ on

$$((A^N)_{\mathfrak{p}}^{\mathrm{hs}})_{\mathfrak{Q}}.$$

In particular, the p-group $I_{G/N}(\mathfrak{P})$ is not generated by 2-reflections on this ring, since it is nontrivial. Then Lemma 3.7 implies that

$$((A^N)^{\mathrm{hs}}_{\mathfrak{p}})^{I_{G/N}(\mathfrak{P})}_{\mathfrak{Q}}$$

is not Cohen–Macaulay. Therefore, by Theorem 3.1, neither is $(A^N)^{G/N} = A^G$. □

4. Permutation invariants

In this section we prove the two directions of Theorem 1.2. A schematic diagram of the proof is found in Figure 5.

4A. *The* **if** *direction.*

In this section we prove:

Proposition 4.1. *If G is generated by transpositions, double transpositions, and 3-cycles, then $k[\boldsymbol{x}]^G$ is Cohen–Macaulay regardless of the field k.*

The groundwork has been laid in Section 2B. The remaining piece of the proof is supplied by a recent, beautiful result of Christian Lange, building on earlier work of Marina Mikhaîlova. Let H be a finite subgroup of the orthogonal group $O_d(\mathbb{R})$, acting on $\mathbb{R}^d$. Endow $\mathbb{R}^d$ with its standard piecewise-linear (PL) structure. The topological quotient $\mathbb{R}^d/H$ carries a PL structure such that the quotient map $\mathbb{R}^d \to \mathbb{R}^d/H$ is a PL map, and the main result of [Lange 2016] is that it is a PL manifold (possibly with boundary) if and only if H is generated by 2-reflections. (Lange calls elements of $O_d(\mathbb{R})$ fixing a codimension-2 subspace *rotations* since they rotate a plane and fix its orthogonal complement, so he calls groups generated this way *rotation-reflection groups*.) The bulk of the work in this result lies in the "if" direction. The proof is a delicate induction on the group order, based on a complete classification of rotation-reflection groups. This classification was proven in joint work of Lange and Mikhaîlova [2016].

Proof of Proposition 4.1. Let G act on $\mathbb{R}^n$ by permutations of the axes. Let $x_1, \dots, x_n$ be the coordinates on $\mathbb{R}^n$. The subspace

$$T = \left\{ \sum_{i=1}^{n} x_i = 0 \right\}$$

is G-invariant. Transpositions in G act as reflections on T, while double transpositions and 3-cycles act as rotations. Thus under the hypothesis of the proposition, G acts on T as a rotation-reflection group. By Lange's work [2016], T/G is a PL manifold.

Recall the Δ of Section 2B: it is the order complex of $B_n \setminus \{\varnothing\}$, which is the first barycentric subdivision of an $(n-1)$-simplex. Embed the underlying topological space $|\Delta|$ of Δ in T as follows. First, map the vertices of Δ to the barycenters of the standard simplex

$$\left\{ x_i \geq 0,\ \sum x_i = 1 \right\}$$

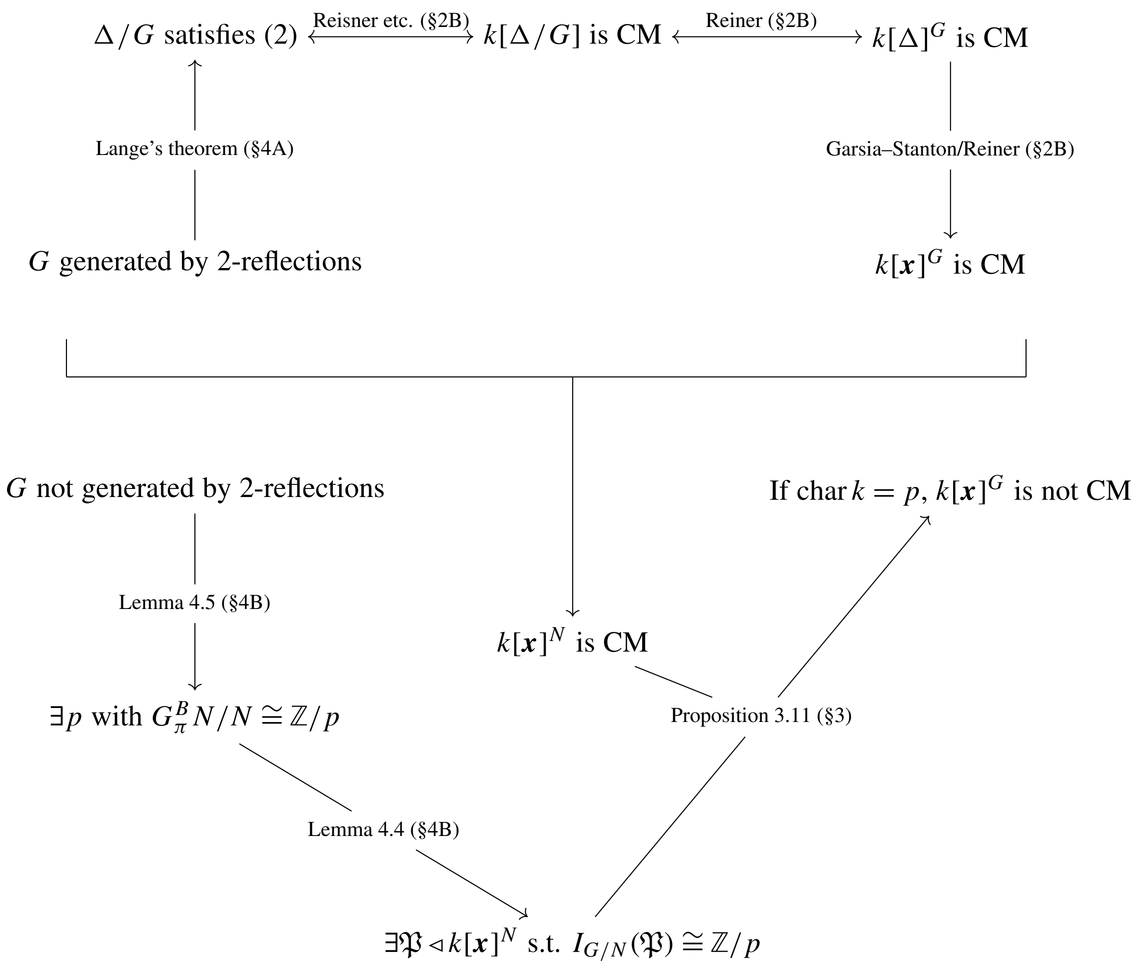

Figure 5. Schematic diagram of the proof of Theorem 1.2. Arrows are implications, and small print above or interrupting an arrow names a result needed for the implication to go through. The §-references indicate where to look for statements and notation definitions. The top half is the "if" direction (Proposition 4.1). The bottom half is the "only if" direction (Proposition 4.2). The group N is the subgroup of G generated by the 2-reflections, so the "if" direction is required to conclude that $k[\boldsymbol{x}]^N$ is Cohen–Macaulay in the bottom half.

in $\mathbb{R}^n$ by mapping each vertex, which by definition is an element $\alpha \in B_n \setminus \{\varnothing\}$, which is itself a nonempty subset of $[n]$, to the barycenter

$$\frac{1}{|\alpha|} \sum_{i \in \alpha} e_i$$

of the set of standard basis vectors $\{e_i\}_{i\in\alpha}$ corresponding to that subset. Then, extend this map to all of Δ by extending linearly from the vertices to each simplex in Δ. Finally, project the affine hyperplane

$\left\{\sum x_i = 1\right\}$ containing the image orthogonally onto T via $(x_1, \ldots, x_n) \mapsto (x_1 - 1/n, \ldots, x_n - 1/n)$. This embedding is G-equivariant for the action of G on $|\Delta|$ induced from its action on $[n]$, and the present action of G on T.

The embedded complex $|\Delta| \subset T$ is evidently a polyhedron, and it is a star of the origin in T since it is the union of closed line segments from the origin to its compact boundary, these segments are disjoint except for the origin itself, and it is a neighborhood of the origin in T (see the definition of a star in Section 2B). Since the action of G is linear, it permutes these segments. Thus $|\Delta|/G = |\Delta/G|$ is also a union of line segments from the (image of the) origin to its compact boundary, and these segments are disjoint except for the origin itself. Also, $|\Delta/G|$ is a neighborhood of the (image of the) origin since $T \to T/G$ is the quotient map by a group of homeomorphisms and is therefore an open map. It is additionally a polyhedron since the quotient map $T \to T/G$ is PL, and the image of a compact polyhedron under a PL map is a compact polyhedron [Rourke and Sanderson 1972, Corollary 2.5]. In other words, $|\Delta/G|$ is a polyhedral star of the image of the origin in the PL $(n-1)$-manifold T/G. It is therefore (per [Rourke and Sanderson 1972, pp.20–21], see the discussion at the end of Section 2B) homeomorphic to a ball. In particular, it is contractible, thus

$$\tilde{H}_i(|\Delta/G|; k) = 0$$

for all i, regardless of the field k; and it is a manifold (with boundary), thus

$$H_i(|\Delta/G|, |\Delta/G| - q; k) = 0$$

for all $i < n - 1$ and all $q \in |\Delta/G|$, regardless of k. Thus it satisfies (2) for all $i < \dim \Delta/G$ and all $q \in |\Delta/G|$, so by the discussion in Section 2B, $k[\boldsymbol{x}]^G$ is Cohen–Macaulay. □

4B. *The* only if *direction.* In this section we complete the proof of Theorem 1.2 by proving the converse of Proposition 4.1.

Proposition 4.2. *If G is not generated by transpositions, double transpositions, and 3-cycles, then there exists a prime p such that for any k of characteristic p, $k[\boldsymbol{x}]^G$ is not Cohen–Macaulay.*

The proof is at the end of the section. Actually we prove somewhat more: for a group G not generated by transpositions, double transpositions, and 3-cycles, we give an explicit construction yielding the prime p. The precise statement is given below as Proposition 4.2b.

In this section, p is conceptually prior to the field k. Our proof will first construct p and then prove that when $\operatorname{char} k = p$, $k[\boldsymbol{x}]^G$ is not Cohen–Macaulay.

We develop the needed machinery for the proof. Let Π_n be the poset of partitions of the set $[n]$, with the order relation given, for any $\pi, \tau \in \Pi_n$, by

$$\pi \leq \tau \iff \pi \text{ refines } \tau.$$

An element $g \in G \subset S_n$ partitions $[n]$ into orbits, and thus determines an element $\pi \in \Pi_n$. This gives a map

$$\begin{aligned} \varphi : G &\to \Pi_n \\ g &\mapsto \pi. \end{aligned}$$

If $\pi \in \Pi_n$, we write G_π^B for the blockwise stabilizer of π in G, i.e., the set of elements of G that act separately on each block of π.

For a given $\pi \in \Pi_n$, let $\mathfrak{P}_\pi^\star$ be the prime ideal of $k[\boldsymbol{x}]$ generated by the binomials $x_i - x_j$ for every pair $i, j \in [n]$ lying in the same block of π. The dimension of $\mathfrak{P}_\pi^\star$ (i.e., the dimension of $k[\boldsymbol{x}]/\mathfrak{P}_\pi^\star$) is the number of blocks of π.

Lemma 4.3. *With this notation, we have*

$$I_G(\mathfrak{P}_\pi^\star) = G_\pi^B.$$

Proof. The ring $k[\boldsymbol{x}]/\mathfrak{P}_\pi^\star$ is the polynomial ring obtained by identifying x_i with x_j for each i and j in the same block of π, so its indeterminates are in bijection with the blocks of π. If $h \in G_\pi^B$, then h acts separately on the x_i's in each block, and therefore h fixes $\mathfrak{P}_\pi^\star$ setwise and the induced action on $k[\boldsymbol{x}]/\mathfrak{P}_\pi^\star$ is trivial. Thus $h \in I_G(\mathfrak{P}_\pi^\star)$. Conversely, if $h \notin G_\pi^B$, then either h fixes π but not blockwise, in which case h fixes $\mathfrak{P}_\pi^\star$ setwise but the action of h on $k[\boldsymbol{x}]/\mathfrak{P}_\pi^\star$ is not trivial, so that $h \in D_G(\mathfrak{P}_\pi^\star)$ but not $I_G(\mathfrak{P}_\pi^\star)$; or else h does not fix π at all, in which case it does not act on $\mathfrak{P}_\pi^\star$, and is not contained in $D_G(\mathfrak{P}_\pi^\star)$, let alone $I_G(\mathfrak{P}_\pi^\star)$. ☐

If N is a normal subgroup of G, denote by $G_\pi^B N/N$ the image of G_π^B in the quotient G/N, and let

$$\mathfrak{P}_\pi = \mathfrak{P}_\pi^\star \cap k[\boldsymbol{x}]^N.$$

Lemma 4.4. *With this notation, we have*

$$I_{G/N}(\mathfrak{P}_\pi) = G_\pi^B N/N.$$

Proof. We have from Lemma 2.10 that

$$I_{G/N}(\mathfrak{P}_\pi) = I_G(\mathfrak{P}_\pi^\star)/I_N(\mathfrak{P}_\pi^\star) = I_G(\mathfrak{P}_\pi^\star)/(N \cap I_G(\mathfrak{P}_\pi^\star)),$$

and from Lemma 4.3 that

$$I_G(\mathfrak{P}_\pi^\star)/(N \cap I_G(\mathfrak{P}_\pi^\star)) = G_\pi^B/(N \cap G_\pi^B) = G_\pi^B N/N. \qquad ☐$$

The following lemma is the device we use to find the characteristic p in which we can prove that $k[\boldsymbol{x}]^G$ fails to be Cohen–Macaulay.

Lemma 4.5. *Let $N \lhd G$ be a proper normal subgroup. Let π be minimal in Π_n among partitions associated (via φ) with elements of G that are not in N. Then*:

(1) *The group $G_\pi^B N/N$ is cyclic of prime order*, *say p.*

(2) *Any element g of $G \setminus N$ whose orbits are given by π has order a power of p.*

(3) *The image of g in G/N generates $G_\pi^B N/N$.*

Proof. Let g be an element of $G \setminus N$ whose orbits are given by π, and let h be any other nontrivial element of G_π^B, in other words a nontrivial element of G whose orbits refine π. (Note that, by minimality of π, either $\varphi(h) = \pi$ or else $h \in N$.) Pick any element $a \in [n]$ acted on nontrivially by h. Then g acts nontrivially on a as well since h's orbits refine g's.

Since h preserves π and g acts transitively on each block of π, there is an $m \in \mathbb{Z}$ such that $g^m(a) = h(a)$. Then $h^{-1}g^m(a) = a$, so that $h^{-1}g^m$ both preserves π and has a fixed point a that g does not have. Thus its orbits properly refine π, and minimality of π among partitions associated to elements of $G \setminus N$ implies that $h^{-1}g^m \in N$. Thus $hN = g^m N$. This shows that g generates the image of G_π^B in G/N, proving (3); thus $G_\pi^B N/N$ is cyclic. Meanwhile, for any prime p dividing the order of g, the orbits of g^p also properly refine those of g, so g^p is in N too; thus the image of g in G/N has order dividing p. Since $g \notin N$ by construction, the order of the image of g in G/N is exactly p. This completes the proof of (1). If q is a hypothetical second prime dividing the order of g in G, then the order of the image of g in G/N is q, for the same reason it is p, and it follows that $q = p$ after all, so there is no such second prime. Therefore g has p-power order in G. This proves (2). □

Proof of Proposition 4.2. Let N be the subgroup of G generated by the transpositions, double transpositions, and 3-cycles (i.e., 2-reflections). By Proposition 4.1, $k[\boldsymbol{x}]^N$ is a Cohen–Macaulay ring. Since $k[\boldsymbol{x}]$ is a finitely generated algebra over k, $k[\boldsymbol{x}]^G$ is also finitely generated as an algebra over k [Bourbaki 1964, Chapitre V §1.9, Théorème 2], so in particular it is noetherian. By the same logic, $k[\boldsymbol{x}]^N$ is finitely generated as an algebra over k, and therefore over $k[\boldsymbol{x}]^G$. Since it is a subring of $k[\boldsymbol{x}]$, which is integral over $k[\boldsymbol{x}]^G$ by [Bourbaki 1964, Chapitre V §1.9, Proposition 22], it is integral over $k[\boldsymbol{x}]^G$ as well, which, together with finite generation as an algebra, implies it is actually finite over the noetherian ring $k[\boldsymbol{x}]^G$. Thus if k is a field of positive characteristic p, then Proposition 3.11 applies, and we can show $k[\boldsymbol{x}]^G$ is not Cohen–Macaulay by exhibiting an inertia group for the action of G/N on $k[\boldsymbol{x}]^N$ that is a nontrivial p-group.

Now if N is a proper subgroup of G per the hypothesis, then we can find a $\pi \in \Pi_n$ that is minimal among all partitions associated (via φ) with elements of $G \setminus N$. Then Lemma 4.5 gives us a prime number p such that $G_\pi^B N/N$ is cyclic of order p, and then Lemma 4.4 gives us a prime ideal $\mathfrak{P}_\pi$ of $k[\boldsymbol{x}]^N$ such that

$$I_{G/N}(\mathfrak{P}_\pi) = G_\pi^B N/N.$$

Thus, for any k of this specific characteristic, we can conclude by Proposition 3.11 that $k[\boldsymbol{x}]^G$ fails to be Cohen–Macaulay. □

An examination of the proof in view of conclusion (2) of Lemma 4.5 shows that we have actually proven the following constructive version of Proposition 4.2 with no additional work:

Proposition 4.2b. *Let N be the subgroup of G generated by the transpositions, double transpositions, and 3-cycles. If $N \subsetneq G$, then for any $g \in G \setminus N$ whose orbits are not refined by the orbits of any other*

$g \in G \setminus N$, the order of g is a prime power p^ℓ, where p has the property that $k[\boldsymbol{x}]^G$ is not Cohen–Macaulay if $\operatorname{char} k = p$. □

5. Conclusion and further questions

In this section we note some implications of the results above, and pose questions for further exploration. Throughout, let N be the subgroup of $G \subset S_n$ generated by the transpositions, double transpositions, and 3-cycles, as at the end of Section 4B.

5A. ***Bad primes; relation to previous work.*** Given a permutation group $G \subset S_n$, let us refer to the set of prime numbers p for which $k[\boldsymbol{x}]^G$ fails to be Cohen–Macaulay if $\operatorname{char} k = p$ as the *bad primes* of G.

It was mentioned in the introduction that the "if" direction of Theorem 1.2 implies that the bad primes of G are a subset of the primes dividing $[G : N]$. We see this as follows: the "if" direction implies that $k[\boldsymbol{x}]^N$ is Cohen–Macaulay. Then, since

$$k[\boldsymbol{x}]^G = (k[\boldsymbol{x}]^N)^{G/N},$$

it follows from the Hochster–Eagon theorem [Hochster and Eagon 1971, Proposition 13] that $k[\boldsymbol{x}]^G$ is Cohen–Macaulay in any characteristic not dividing the order of G/N. Meanwhile, the "only if" direction of Theorem 1.2 implies that if the set of primes dividing $[G : N]$ is nonempty, then so is G's set of bad primes.

It was also mentioned in the introduction that the present work unites and generalizes several previously known results: Reiner's theorem [1992] that the invariant rings of A_n and the diagonally embedded $S_n \hookrightarrow S_n \times S_n$ are Cohen–Macaulay over all fields; Hersh's similar theorem [2003a; 2003b] for the wreath product $S_2 \wr S_n \subset S_{2n}$, and Kemper's theorems [1999] that in the p-group case, the "only if" direction of Theorem 1.2 holds, and that the invariant ring of a *regular* permutation group G is Cohen–Macaulay over all fields if and only if $G = C_2$, C_3, or $C_2 \times C_2$, and in all other cases, every prime dividing $|G|$ is a bad prime for G. Most of these results are immediate implications of the "if" direction of Theorem 1.2:

- The group A_n is generated by 3-cycles.
- The diagonal $S_n \hookrightarrow S_n \times S_n$ is generated by the double transpositions $(i, i+1)(i+n, i+n+1)$ for $i = 1, \dots, n-1$.
- The wreath product $S_2 \wr S_n$ is generated by the transpositions $(2i-1, 2i)$ and the double transpositions $(2i-1, 2i+1)(2i, 2i+2)$ for $i = 1, \dots, n-1$.
- The regular representations of C_2, C_3, and $C_2 \times C_2$ are generated by (in fact, their only nontrivial elements are) transpositions, 3-cycles, and double transpositions, respectively.

Recovering the other half of Kemper's result on regular permutation groups (that *every* prime dividing $|G|$ is bad for G) from the present work requires the constructive version of the "only if" direction given in Proposition 4.2b. Recall that if G acts regularly, i.e., freely and transitively, on $[n]$, then this action is isomorphic to G's left-translation action on its own elements. Then we have $|G| = n$, and every element

g of G splits $[n]$ into orbits of equal length the order of g, because these orbits are in bijection with the right cosets $\langle g\rangle h$, $h \in G$.

If G acts regularly and $|G| = n \geq 5$, then G does not contain any transpositions, double transpositions, or 3-cycles, so N is trivial. If p is any prime dividing $|G|$, then G has an element g of order p, which, by the discussion in the last paragraph, partitions $[n]$ into orbits of equal length p. This partition cannot be refined by any nontrivial partition with parts of equal length since p is prime; thus no element of $G \setminus N = G \setminus \{1\}$ can have orbits refining g's. It follows from Proposition 4.2b that p is a bad prime for G.

The remaining case is $n = 4$ and $G = C_4$. In this case, G is a 2-group not generated by its lone double transposition, so it follows from Theorem 1.2 that 2 is a bad prime for G.

5B. ***Groups generated by transpositions, double transpositions, and 3-cycles.*** Theorem 1.2 calls attention to the family of permutation groups generated by transpositions, double transpositions, and 3-cycles. One may wonder how extensive is this family of groups. It turns out to be very limited. One can extract a classification from Lange and Mikhaîlova's classification of all rotation-reflection groups [2016], but this is more power than is needed. In the case that G is transitive, such groups were already classified in 1979 by W. Cary Huffman [1980, Theorem 2.1]:

(1) If G's transpositions generate a transitive subgroup, then $G = S_n$.

(2) If G contains a transposition but the transpositions do not act transitively, then $n = 2m$ is even and G is isomorphic to the wreath product $S_2 \wr S_m$.

(3) If G does not contain a transposition but does contain a three-cycle, then $G = A_n$.

(4) Otherwise, G contains no transpositions or 3-cycles and is generated by double transpositions. Then we have:

   (a) If G contains a subgroup acting transitively on 5 points and fixing the rest, then either $n = 5$ and $G \cong D_5$ in its usual action on the vertices of a regular pentagon, or else $n = 6$ and $G \cong A_5 \cong \mathrm{PSL}(2, 5)$ in its transitive action on 6 points, e.g., the six points of the projective line over $\mathbb{F}_5$.

   (b) If G contains a subgroup acting transitively on 7 points and fixing the rest, then either $n = 7$ and $G \cong GL(3, 2)$ acting on the nonzero vectors of $\mathbb{F}_2^3$, or else $n = 8$ and $G \cong \mathrm{AGL}(3, 2) = \mathbb{F}_2^3 \rtimes \mathrm{GL}(3, 2)$ acting on the points of $\mathbb{A}^3_{\mathbb{F}_2}$.

   (c) If G does not contain either of these kinds of subgroups, then $n = 2m$ is even, and G is isomorphic to the alternating subgroup of the wreath product $S_2 \wr S_m$.

When one considers intransitive groups G, one does not end up too far beyond direct products of the above, since transpositions and 3-cycles can only act in a single orbit, while double transpositions can only act in two orbits, as a transposition in each. For example, if G has two orbits, the classification begins as follows. If G is not a direct product of the above, it contains a double transposition that acts as a transposition in each orbit. Then its image in each orbit contains a transposition, so is either S_n or $S_2 \wr S_m$ by the above. The possibilities are then highly constrained by Goursat's lemma.

Thus Theorem 1.2 shows that most permutation groups G have at least one bad prime.

5C. ***Further questions.*** Since Theorem 1.2 implies that the set of bad primes of G is contained in the set of prime factors of $[G:N]$ and is nonempty exactly when the latter is nonempty, one might hope that these two sets are always equal. This is not the case. For example, let $G\subset S_7$ be the Frobenius group of order 21 generated by

$$(1234567),\ (124)(365).$$

All the nontrivial elements in this group are 7-cycles or double 3-cycles. Thus N is trivial in this case, and the candidate bad primes are 3 and 7.

Now $\pi=\{1,2,4\}\cup\{3,5,6\}\cup\{7\}$ is a minimal partition as in Lemma 4.5, and thus the corresponding $g=(124)(365)$ generates an inertia group of order 3 for the action of $G/N=G$ on $k[\boldsymbol{x}]^N=k[\boldsymbol{x}]$. Then Proposition 4.2b shows that if k has characteristic 3, $k[\boldsymbol{x}]^G$ fails to be Cohen–Macaulay; i.e., 3 is a bad prime for this G.

On the other hand, 7 is not a bad prime for this G. This can be seen using the criterion given by Kemper [2001, Theorem 3.3], since 7 divides $|G|$ just once. Thus, a prime can divide $[G:N]$ without being bad. (By a computer calculation, no example of this phenomenon occurs below degree 7.)

At the other extreme, one might hope that the bad primes of G are *only* those which are furnished by Proposition 4.2b. This is not true either. Take $G=D_7$, the dihedral group of order 14 acting on the vertices of a heptagon, which is also a Frobenius group. Now, all the nontrivial elements are 7-cycles and triple transpositions, so again, N is trivial, and the candidate bad primes are 2 and 7. This time, they both really are bad primes. One can see this using Kemper's criterion [2001, Theorem 3.3]. For 2 it also follows from Proposition 4.2b, but for 7 it does not, since the 7-cycles have orbits that are properly refined by the triple transpositions.

Thus it remains to be determined, for a given G, exactly which primes are bad. Theorem 1.2 gives us a finite list of candidate bad primes (those dividing $[G:N]$), and, if this list is nonempty, Proposition 4.2b gives us some specific primes that are definitely bad. Among the remaining candidate bad primes, if any divide $|G|$ only once, [Kemper 2001, Theorem 3.3] can be used to determine if they are actually bad. What remains to be determined is whether p is a bad prime if $p^2\mid |G|$ and p is not associated to a $g\in G\setminus N$ with minimal orbits as in Proposition 4.2b.

Question 5.1. How can Cohen–Macaulayness of $k[\boldsymbol{x}]^G$ be assessed when [Kemper 2001, Theorem 3.3] and the present work are both inapplicable, i.e., when $p\mid [G:N]$ and $p^2\mid |G|$, but p does not come from a minimal $g\in G\setminus N$ as in Proposition 4.2b?

Another line of inquiry that flows from the present work has to do with the relationship between the arguments in the "if" and "only if" directions. The proof of the "if" direction is a mildly revised version of an argument given by the first author in his doctoral thesis [Blum-Smith 2017]. In that same work, he also proved the "only if" direction for $k[\Delta/G]$ (see Section 2B for notation), but not for $k[\boldsymbol{x}]^G$. There, the "only if" argument was framed in the same topological language as the "if" argument, which is why it applied to $k[\Delta/G]$ (taking advantage of Stanley–Reisner theory) but not $k[\boldsymbol{x}]^G$. The second author suggested to transfer the "only if" argument from topological into commutative-algebraic language, and much of the present paper sprang from this suggestion.

This transfer was accomplished piecemeal, with an individual search for each commutative-algebraic fact needed to replace each topological fact. For example, Raynaud's theorem (Lemma 2.14) replaced an elementary principle about the relationship between point stabilizers and the local structure in a topological quotient. The well-behavedness of inertia groups with respect to normal subgroups (Lemma 2.10) replaced an elementary fact about group actions on a set. The observation that inertia p-groups obstruct Cohen–Macaulayness if they are not generated by 2-reflections (Lemma 3.7), based on [Lorenz and Pathak 2001, Corollary 4.3], replaced an argument about the homology of links in the quotient of a simplicial complex.

Nonetheless, the authors had the conviction throughout that an overarching principle was at play. It may be fruitful to seek a more comprehensive understanding of the relationship between the topology and the algebra. Stanley–Reisner theory gives a partial answer to this question, but it does not appear to account for the "only if" direction of Theorem 1.2, so a fuller picture is desirable.

Here are two more focused questions that approach this inquiry from various directions:

Question 5.2. Is there a purely algebraic proof of Theorem 1.2, making no use of Stanley–Reisner theory or Lange's result on PL manifolds?

Question 5.3. For a fixed $p = \operatorname{char} k$ as in Question 5.1, can $k[\boldsymbol{x}]^G$ be Cohen–Macaulay without $k[\Delta/G]$ being Cohen–Macaulay?

Acknowledgement

The authors wish to thank Gregor Kemper, Victor Reiner, and Christian Lange for fruitful discussions and comments, Tom Zaslavsky for alerting them to the conventional partial order on Π_n, and an anonymous referee for very helpful comments.

The second author's work on this project was partially supported by a grant from the National Research Foundation of South Africa.

Parts of this work originally appeared as part the first author's doctoral thesis under the supervision of Yuri Tschinkel and Fedor Bogomolov at the Courant Institute of Mathematical Sciences at NYU.

References

[Blum-Smith 2017] B. Blum-Smith, *Two inquiries about finite groups and well-behaved quotients*, Ph.D. thesis, New York University, 2017, Available at https://search.proquest.com/docview/1938647220. MR

[Bourbaki 1964] N. Bourbaki, *Éléments de mathématique: Algèbre commutative, Chapitres 5–6*, Actualités Scientifiques et Industrielles **1308**, Hermann, Paris, 1964. MR Zbl

[Bruns and Herzog 1993] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, 1993. MR Zbl

[Campbell et al. 1999] H. E. A. Campbell, A. V. Geramita, I. P. Hughes, R. J. Shank, and D. L. Wehlau, "Non-Cohen–Macaulay vector invariants and a Noether bound for a Gorenstein ring of invariants", *Canad. Math. Bull.* **42**:2 (1999), 155–161. MR Zbl

[Dufresne et al. 2009] E. Dufresne, J. Elmer, and M. Kohls, "The Cohen–Macaulay property of separating invariants of finite groups", *Transform. Groups* **14**:4 (2009), 771–785. MR Zbl

[Duval 1997] A. M. Duval, "Free resolutions of simplicial posets", *J. Algebra* **188**:1 (1997), 363–399. MR Zbl

[EGA I 1960] A. Grothendieck, "Eléments de géométrie algébrique, I: Le langage des schémas", *Inst. Hautes Études Sci. Publ. Math.* **4** (1960), 5–228. MR Zbl

[Eisenbud 1995] D. Eisenbud, *Commutative algebra with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, 1995. MR Zbl

[Ellingsrud and Skjelbred 1980] G. Ellingsrud and T. Skjelbred, "Profondeur d'anneaux d'invariants en caractéristique p", *Compositio Math.* **41**:2 (1980), 233–244. MR Zbl

[Garsia 1980] A. M. Garsia, "Combinatorial methods in the theory of Cohen–Macaulay rings", *Adv. in Math.* **38**:3 (1980), 229–266. MR Zbl

[Garsia and Stanton 1984] A. M. Garsia and D. Stanton, "Group actions of Stanley–Reisner rings and invariants of permutation groups", *Adv. in Math.* **51**:2 (1984), 107–201. MR Zbl

[Gordeev and Kemper 2003] N. Gordeev and G. Kemper, "On the branch locus of quotients by finite groups and the depth of the algebra of invariants", *J. Algebra* **268**:1 (2003), 22–38. MR Zbl

[Hamilton and Marley 2007] T. D. Hamilton and T. Marley, "Non-Noetherian Cohen–Macaulay rings", *J. Algebra* **307**:1 (2007), 343–360. MR Zbl

[Hersh 2003a] P. Hersh, "Lexicographic shellability for balanced complexes", *J. Algebraic Combin.* **17**:3 (2003), 225–254. MR Zbl

[Hersh 2003b] P. Hersh, "A partitioning and related properties for the quotient complex $\Delta(B_{lm})/S_l \wr S_m$", *J. Pure Appl. Algebra* **178**:3 (2003), 255–272. MR Zbl

[Hochster and Eagon 1971] M. Hochster and J. A. Eagon, "Cohen–Macaulay rings, invariant theory, and the generic perfection of determinantal loci", *Amer. J. Math.* **93** (1971), 1020–1058. MR Zbl

[Huffman 1980] W. C. Huffman, "Imprimitive linear groups generated by elements containing an eigenspace of codimension two", *J. Algebra* **63**:2 (1980), 499–513. MR Zbl

[Kac and Watanabe 1982] V. Kac and K. Watanabe, "Finite linear groups whose ring of invariants is a complete intersection", *Bull. Amer. Math. Soc.* (*N.S.*) **6**:2 (1982), 221–223. MR Zbl

[Kemper 1999] G. Kemper, "On the Cohen–Macaulay property of modular invariant rings", *J. Algebra* **215**:1 (1999), 330–351. MR Zbl

[Kemper 2001] G. Kemper, "The depth of invariant rings and cohomology", *J. Algebra* **245**:2 (2001), 463–531. MR Zbl

[Kemper 2012] G. Kemper, "The Cohen–Macaulay property and depth in invariant theory", pp. 53–63 in *Proceedings of the 33rd Symposium on Commutative Algebra* (Japan), 2012.

[Lange 2016] C. Lange, "Characterization of finite groups generated by reflections and rotations", *J. Topol.* **9**:4 (2016), 1109–1129. MR Zbl

[Lange and Mikhaîlova 2016] C. Lange and M. A. Mikhaîlova, "Classification of finite groups generated by reflections and rotations", *Transform. Groups* **21**:4 (2016), 1155–1201. MR Zbl

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics **6**, Oxford University Press, 2002. MR Zbl

[Lorenz and Pathak 2001] M. Lorenz and J. Pathak, "On Cohen–Macaulay rings of invariants", *J. Algebra* **245**:1 (2001), 247–264. MR Zbl

[Munkres 1984] J. R. Munkres, "Topological results in combinatorics", *Michigan Math. J.* **31**:1 (1984), 113–128. MR Zbl

[Neusel and Smith 2002] M. D. Neusel and L. Smith, *Invariant theory of finite groups*, Mathematical Surveys and Monographs **94**, American Mathematical Society, Providence, RI, 2002. MR Zbl

[Raynaud 1970] M. Raynaud, *Anneaux locaux henséliens*, Lecture Notes in Math. **169**, Springer, 1970. MR Zbl

[Reiner 1992] V. Reiner, "Quotients of Coxeter complexes and P-partitions", pp. vi+134 Mem. Amer. Math. Soc. **460**, Amer. Math. Soc., Providence, RI, 1992. MR Zbl

[Reiner 2003] V. Reiner, Appendix to [Hersh 2003b], pages 269–271 in *J. Pure Appl. Algebra* **178**:3 (2003), 255–272.

[Reisner 1976] G. A. Reisner, "Cohen–Macaulay quotients of polynomial rings", *Advances in Math.* **21**:1 (1976), 30–49. MR Zbl

[Rourke and Sanderson 1972] C. P. Rourke and B. J. Sanderson, *Introduction to piecewise-linear topology*, Ergebnisse der Mathematik **69**, Springer, Berlin, 1972. Reprinted in the series Springer Study Edition **69**, Springer, Berlin, 1982. MR Zbl

[Smith 1995] L. Smith, *Polynomial invariants of finite groups*, Research Notes in Mathematics **6**, A K Peters, Ltd., Wellesley, MA, 1995. MR Zbl

[Smith 1996] L. Smith, "Some rings of invariants that are Cohen–Macaulay", *Canad. Math. Bull.* **39**:2 (1996), 238–240. MR Zbl

[Stanley 1986] R. P. Stanley, *Enumerative combinatorics, I*, Wadsworth & Brooks/Cole, Monterey, CA, 1986. MR Zbl

[Stanley 1991] R. P. Stanley, "f-vectors and h-vectors of simplicial posets", *J. Pure Appl. Algebra* **71**:2-3 (1991), 319–331. MR Zbl

Communicated by Victor Reiner
Received 2018-02-26 Revised 2018-05-16 Accepted 2018-06-17

ben@cims.nyu.edu *Department of Natural Sciences and Mathematics, Eugene Lang College, the New School for Liberal Arts, New York City, NY, United States*

sophie.marques@uct.ac.za *Department of Mathematics and Applied Mathematics, University of Cape Town, Cape Town, South Africa*

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 12 No. 7 2018

1937-0652(2018)12:7;1-A