Algebra & Number Theory

Volume 14
2020
No. 1

msp

Algebra & Number Theory

msp.org/ant

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

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

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 14:1 (2020)
dx.doi.org/10.2140/ant.2020.14.1

Gorenstein-projective and semi-Gorenstein-projective modules

Claus Michael Ringel and Pu Zhang

Let A be an artin algebra. An A-module M will be said to be semi-Gorenstein-projective provided that $\mathrm{Ext}^i(M, A) = 0$ for all $i \geq 1$. All Gorenstein-projective modules are semi-Gorenstein-projective and only few and quite complicated examples of semi-Gorenstein-projective modules which are not Gorenstein-projective have been known. One of the aims of the paper is to provide conditions on A such that all semi-Gorenstein-projective left modules are Gorenstein-projective (we call such an algebra left weakly Gorenstein). In particular, we show that in case there are only finitely many isomorphism classes of indecomposable left modules which are both semi-Gorenstein-projective and torsionless, then A is left weakly Gorenstein. On the other hand, we exhibit a 6-dimensional algebra Λ with a semi-Gorenstein-projective module M which is not torsionless (thus not Gorenstein-projective). Actually, also the Λ-dual module M^* is semi-Gorenstein-projective. In this way, we show the independence of the total reflexivity conditions of Avramov and Martsinkovsky, thus completing a partial proof by Jorgensen and Şega. Since all the syzygy-modules of M and M^* are 3-dimensional, the example can be checked (and visualized) quite easily.

1. Introduction

1.1. Notations and definitions. Let A be an artin algebra. All modules will be finitely generated. Usually, the modules we are starting with will be left modules, but some constructions then yield right modules. Let $\mathrm{mod}\, A$ be the category of all finitely generated left A-modules and $\mathrm{add}(A)$ the full subcategory of all projective modules.

If M is a module, let PM be a projective cover of M, and ΩM the kernel of the canonical map $PM \to M$. The modules $\Omega^t M$ with $t \geq 0$ are called the *syzygy* modules of M. A module M is said to be Ω-*periodic* provided that there is some $t \geq 1$ with $\Omega^t M = M$.

The right A-module $M^* = \mathrm{Hom}(M, A)$ is called the A-*dual* of M. Let $\phi_M : M \to M^{**}$ be defined by $\phi_M(m)(f) = f(m)$ for $m \in M$, $f \in M^*$. A module M is said to be *torsionless* provided that M is a submodule of a projective module, or, equivalently, provided that ϕ_M is injective. A module M is called *reflexive* provided that ϕ_M is bijective.

Supported by NSFC 11971304.
MSC2010: primary 16E65; secondary 16E05, 16G10, 16G50, 20G42.
Keywords: Gorenstein-projective module, semi-Gorenstein-projective module, left weakly Gorenstein algebra, torsionless module, reflexive module, t-torsionfree module, Frobenius category, Ʊ-quiver.

Let $\operatorname{Tr} M$ be the cokernel of f^*, where f is a minimal projective presentation of M (this is the canonical map $P(\Omega M) \to PM$). Note that $\operatorname{Tr} M$ is a right A-module, called the *transpose of M*.

A complete projective resolution is a (double infinite) exact sequence

$$P^\bullet : \cdots \to P^{-1} \to P^0 \xrightarrow{d^0} P^1 \to \cdots$$

of projective left A-modules, such that $\operatorname{Hom}_A(P^\bullet, A)$ is again exact. A module M is *Gorenstein-projective*, if there is a complete projective resolution $P^\bullet$ with M isomorphic to the image of d^0.

A module M will be said to be *semi-Gorenstein-projective* provided that $\operatorname{Ext}^i(M, A) = 0$ for all $i \geq 1$. All Gorenstein-projective modules are semi-Gorenstein-projective. If M is semi-Gorenstein-projective, then so is ΩM. Denote by $\operatorname{gp}(A)$ the class of all Gorenstein-projective modules and by ${}^\perp A$ the class of all semi-Gorenstein-projective modules. Then $\operatorname{gp}(A) \subseteq {}^\perp A$. We propose to call an artin algebra A *left weakly Gorenstein* provided that ${}^\perp A = \operatorname{gp}(A)$, i.e., any semi-Gorenstein-projective module is Gorenstein-projective. (And A is said to be *right weakly Gorenstein* if its opposite algebra A^{op} is left weakly Gorenstein.)

The first aim of the paper is to provide a systematic treatment of the relationship between semi-Gorenstein-projective modules and Gorenstein-projective modules, see theorems 1.2 to 1.4. Some of these results are (at least partially) known or can be obtained from the literature, in particular see the paper [B3] by Beligiannis, but we hope that a unified, elementary and direct presentation may be appreciated.

1.2. First, we have various characterizations of the left weakly Gorenstein algebras.

Theorem. *Let A be an artin algebra. The following statements are equivalent:*

(1) *A is left weakly Gorenstein.*

(2) *Any semi-Gorenstein-projective module is torsionless.*

(3) *Any semi-Gorenstein-projective module is reflexive.*

(4) *For any semi-Gorenstein-projective module M, the map ϕ_M is surjective.*

(5) *For any semi-Gorenstein-projective module M, the module M^* is semi-Gorenstein projective.*

(6) *Any semi-Gorenstein-projective module M satisfies* $\operatorname{Ext}^1(M^*, A_A) = 0$.

(7) *Any semi-Gorenstein-projective module M satisfies* $\operatorname{Ext}^1(\operatorname{Tr} M, A_A) = 0$.

The equivalence of (1) and (2) was published by Huang–Huang [HH, Theorem 4.2].

1.3. The next result concerns artin algebras with finitely many indecomposable semi-Gorenstein-projective modules or with finitely many indecomposable torsionless modules.

Theorem. *If the number of isomorphism classes of indecomposable modules which are both semi-Gorenstein-projective and torsionless is finite, then A is left weakly Gorenstein and any indecomposable nonprojective semi-Gorenstein-projective module is Ω-periodic.*

This combines two different directions of thoughts. First of all, Yoshino [Y] has shown that for certain commutative rings R (in particular all artinian commutative rings) the finiteness of the number of isomorphism classes of indecomposable semi-Gorenstein-projective R-modules implies that R is left weakly Gorenstein. This was generalized to artin algebras by Beligiannis [B3, Corollary 5.11]. Second, according to Marczinzik [M1], all torsionless-finite artin algebras (these are the artin algebras with only finitely many isomorphism classes of torsionless indecomposable modules) are left weakly Gorenstein. Note that a lot of interesting classes of artin algebras are torsionless-finite, see 3.6.

1.4. If $\mathcal{C}$ is an extension-closed full subcategory of mod A, then the embedding of $\mathcal{C}$ into mod A provides an exact structure on $\mathcal{C}$, called its *canonical* exact structure (for the basic properties of exact structures, see for example [K, Appendix A]). An exact category $\mathcal{F}$ is called a *Frobenius category* provided that it has enough projective and enough injective objects and that the projective objects in $\mathcal{F}$ are just the injective objects in $\mathcal{F}$. We denote by $\mathcal{P}(\mathcal{F})$ (and by $\mathcal{I}(\mathcal{F})$) the full subcategory of the projective (respectively injective) objects in $\mathcal{F}$.

The subcategories $\mathrm{gp}(A)$ and ${}^{\perp}A$ are extension-closed, and with its canonical exact structure $\mathrm{gp}(A)$ is Frobenius with $\mathcal{P}(\mathrm{gp}(A)) = \mathrm{add}\, A$ [B2, Prop. 3.8]. Thus, if A is left weakly Gorenstein, then $\mathcal{F} = \mathrm{gp}(A)$ is an extension-closed full subcategory of mod A which is Frobenius with the canonical exact structure and satisfies $\mathcal{P}(\mathcal{F}) \subseteq {}^{\perp}A \subseteq \mathcal{F}$. The following result shows that these properties characterize left weakly Gorenstein algebras.

Theorem. *Let A be an artin algebra and $\mathcal{F}$ an extension-closed full subcategory of* mod A *such that $\mathcal{F}$ is a Frobenius category with respect to its canonical exact structure. If $\mathcal{P}(\mathcal{F}) \subseteq {}^{\perp}A \subseteq \mathcal{F}$, then A is left weakly Gorenstein and $\mathcal{F} = \mathrm{gp}(A)$.*

A full subcategory $\mathcal{C}$ of mod A is said to be *resolving* provided that it contains all the projective modules and is closed under extensions, direct summands and kernels of surjective maps. Note that ${}^{\perp}A$ and $\mathrm{gp}(A)$ are resolving subcategories.

Corollary 1. *Let A be an artin algebra and $\mathcal{F}$ a resolving subcategory of* mod A *with ${}^{\perp}A \subseteq \mathcal{F}$. Assume that $\mathcal{F}$ with its canonical exact structure is a Frobenius subcategory. Then A is left weakly Gorenstein and $\mathcal{F} = \mathrm{gp}(A)$.*

Taking $\mathcal{F} = {}^{\perp}A$ in Theorem 1.4 we get

Corollary 2. *An artin algebra A is left weakly Gorenstein if and only if ${}^{\perp}A$ with its canonical exact structure is a Frobenius subcategory.*

We remark that $\mathrm{gp}(A)$ is the largest resolving Frobenius subcategory of mod A (compare [B1, Prop. 2.13, Theorem 2.11], [B2, p.145], and [B3, p.1989]; also [ZX, Prop. 5.1]). This implies Theorem 1.4 and the two corollaries (as one of the referees has pointed out).

1.5. The $\mho$-quiver of an artin algebra A. The main tool used in the paper are the operator $\mho$, and the $\mho$-quiver of A. Here are the definitions.

Recall that a map $f : M \to M'$ is said to be *left minimal* provided that any map $h : M' \to M'$ with $hf = f$ is an automorphism [AR1]. A left $\mathrm{add}(A)$-approximation will be called *minimal* provided that it is left minimal. We denote by $\mho M$ the cokernel of a minimal left $\mathrm{add}(A)$-approximation of M. (The symbol $\mho$, pronounced “agemo”, should be a reminder that $\mho$ has to be considered as a kind of inverse of Ω.) It turns out that the operator $\mho$ coincides with $\operatorname{Tr}\Omega\operatorname{Tr}$, which has been studied by Auslander and Reiten in [AR2]. See Subsection 4.4, and also 4.7.

Let $\omega : M \to P$ be a minimal left $\operatorname{add}(A)$-approximation with cokernel map $\pi : P \to \mho M$. If M is indecomposable and not projective, then the image of ω is contained in the radical of P, thus π is a projective cover. If M is, in addition, torsionless (so that ω is injective), then $\mho M$ is indecomposable and not projective, and $\Omega\mho M \simeq M$.

The $\mho$-*quiver* of A has as vertices the isomorphism classes $[X]$ of the indecomposable nonprojective modules X and there is an arrow

$$[X] \dashleftarrow [\mho X]$$

for any torsionless (indecomposable, nonprojective) module X. (We hope that the reader is not irritated by the chosen orientation of the arrow: it corresponds to the usual convention when dealing with Ext-quivers.) A component of the $\mho$-quiver will be called an $\mho$-*component;* a path in the $\mho$-quiver will be called an $\mho$-*path.*

In the $\mho$-quiver, an arrow ending at $[X]$ starts at $[\mho X]$, thus for any vertex $[X]$, there is at most one arrow ending in $[X]$. If $[Z]$ is the start of an arrow, say $Z \simeq \mho X$ for some vertex $[X]$, then $X \simeq \Omega\mho X \simeq \Omega Z$ implies that the arrow is uniquely determined. This shows that *at any vertex of the $\mho$-quiver, at most one arrow starts and at most one arrow ends.* As a consequence, we have:

Proposition. *Any $\mho$-component is a linearly oriented quiver $\mathbb{A}_n$ with $n \geq 1$ vertices, or an oriented cycle $\tilde{\mathbb{A}}_n$ with $n+1 \geq 1$ vertices, or of the form $-\mathbb{N}$, or $\mathbb{N}$, or $\mathbb{Z}$.*

Note that we consider any subset I of $\mathbb{Z}$ as a quiver, with arrows from z to $z-1$ (provided that both $z-1$ and z belong to I). For example, the interval $\{1, 2, \ldots, n\}$ is the quiver $\mathbb{A}_n$ with linear orientation (with 1 being the unique sink and n the unique source). Here are the quivers $-\mathbb{N}$ and $\mathbb{N}$:

$$\underset{-\mathbb{N}}{\cdots \dashleftarrow \circ \dashleftarrow \circ \dashleftarrow \circ} \qquad\qquad \underset{\mathbb{N}}{\circ \dashleftarrow \circ \dashleftarrow \circ \dashleftarrow \cdots}$$

As we will see in 7.7, all cases mentioned here can arise as $\mho$-components.

An indecomposable nonprojective module M will be said to be of $\mho$-*type* Δ where $\Delta \in \{\mathbb{A}_n, \tilde{\mathbb{A}}_n, -\mathbb{N}, \mathbb{N}, \mathbb{Z}\}$ in case the $\mho$-component containing $[M]$ is of the form Δ.

Let us collect what can be read out about an indecomposable nonprojective module when looking at its position in the $\mho$-quiver. Recall that a module M is said to be t-*torsionfree,* provided $\operatorname{Ext}^i(\operatorname{Tr} M, A_A) = 0$ for $1 \leq i \leq t$ (and ∞-*torsionfree,* provided $\operatorname{Ext}^i(\operatorname{Tr} M, A_A) = 0$ for all $i \geq 1$); the definition is due to Auslander (see [A1, Br, AB]).

Theorem. *Let M be an indecomposable nonprojective module.*

(0) *$[M]$ is an isolated vertex if and only if* $\operatorname{Ext}^1(M, A) \neq 0$ *and M is not torsionless.*

(1) *$[M]$ is the start of a path of length $t \geq 1$ if and only if* $\operatorname{Ext}^i(M, A) = 0$ *for $1 \leq i \leq t$. In particular, $[M]$ is the start of an arrow if and only if* $\operatorname{Ext}^1(M, A) = 0$.

(1′) *$[M]$ is the start of an infinite path if and only if M is semi-Gorenstein-projective.*

(1″) *$[M]$ is of $\mho$-type $-\mathbb{N}$ if and only if M is semi-Gorenstein-projective, but not Gorenstein-projective.*

(2) *$[M]$ is the end of a path of length $t \geq 1$ if and only if M is t-torsionfree for $1 \leq i \leq t$, if and only if $\mho^{i-1}M$ is torsionless for $1 \leq i \leq t$. In particular, $[M]$ is the end of an arrow if and only if M is torsionless; and $[M]$ is the end of a path of length 2 if and only if M is reflexive.*

(2′) *$[M]$ is the end of an infinite path if and only if M is ∞-torsionfree, if and only if M is reflexive and M^* is semi-Gorenstein-projective.*

(2″) *$[M]$ is of $\mho$-type $\mathbb{N}$ if and only if M is ∞-torsionfree, but not Gorenstein-projective.*

(3) *$[M]$ is the start of an infinite path and also the end of an infinite path if and only if M is Gorenstein-projective. M is of $\mho$-type $\mathbb{Z}$ if and only if M is Gorenstein-projective and not Ω-periodic. M is of $\mho$-type $\tilde{\mathbb{A}}_n$ for some $n \geq 0$ if and only if M is Gorenstein-projective and Ω-periodic.*

(4) *A-duality provides a bijection between the isomorphism classes of the reflexive indecomposable A-modules of $\mho$-type $\mathbb{A}_n$ and the isomorphism classes of the reflexive indecomposable $A^{\rm op}$-modules of $\mho$-type $\mathbb{A}_n$. Thus, for any $n \geq 3$, A has $\mho$-components of form $\mathbb{A}_n$ if and only if $A^{\rm op}$ has $\mho$-components of form $\mathbb{A}_n$.*

(5) *A-duality provides a bijection between the isomorphism classes of the reflexive indecomposable A-modules of $\mho$-type $\mathbb{N}$ and the isomorphism classes of the reflexive indecomposable $A^{\rm op}$-modules of $\mho$-type $-\mathbb{N}$. Thus, A has $\mho$-components of form $\mathbb{N}$ if and only if $A^{\rm op}$ has $\mho$-components of form $-\mathbb{N}$.*

Remark 1. Characterizations of Gorenstein-projective modules. The $\mho$-quiver shows nicely that an indecomposable module M is Gorenstein-projective if and only if both M and $\operatorname{Tr} M$ are semi-Gorenstein-projective, if and only if M is reflexive and both M and M^* are semi-Gorenstein projective: See (1′), (2′) and (3).

Remark 2. Symmetry. The $\mho$-quiver shows a symmetry between the semi-Gorenstein-projective modules and the ∞-torsionfree modules: An indecomposable nonprojective module M is semi-Gorenstein-projective provided there is an infinite $\mho$-path starting in M; and M is ∞-torsionfree, provided there is an infinite $\mho$-path ending in M.

Remark 3. Weakly Gorenstein algebras. *An artin algebra A is left weakly Gorenstein if and only if there are no modules of $\mho$-type $-\mathbb{N}$,* see (1″). Similarly, *A is right weakly Gorenstein if and only if there are no modules of $\mho$-type $\mathbb{N}$,* see (2″) and (5).

1.6. The first example of a semi-Gorenstein-projective module which is not Gorenstein-projective was constructed by Jorgensen and Şega [JS] in 2006, for a commutative algebra of dimension 8. Recently, Marczinzik [M2] presented some noncommutative algebras with semi-Gorenstein-projective modules which are not Gorenstein-projective. In 6.1, we exhibit a class of 6-dimensional k-algebras $\Lambda(q)$ with parameter $q \in k \setminus \{0\}$ and a family $M(\alpha)$ of 3-dimensional indecomposable $\Lambda(q)$-modules (with $\alpha \in k$) in order to find new examples:

Theorem. *Let $\Lambda(q)$ be the 6-dimensional algebra defined in 6.1. If the multiplicative order of q is infinite, then the Λ-modules $M(q)$ and $M(q)^*$ both are semi-Gorenstein-projective, but $M(q)$ is not torsionless, thus not Gorenstein-projective; all the syzygy modules $\Omega^t M(q)$ and $\Omega^t(M(q)^*)$ with $t \geq 0$ are 3-dimensional and indecomposable; the module $M(q)^{**} \simeq \Omega M(1)$ is also 3-dimensional, but decomposable.*

Addendum. *For any q, the $\Lambda(q)$-modules $M(\alpha)$ with $\alpha \in k \setminus q^{\mathbb{Z}}$ are Gorenstein-projective. Thus, if k is infinite, then there are infinitely many isomorphism classes of 3-dimensional Gorenstein-projective modules.*

1.7. Independence of the total reflexivity conditions. It was asked by Avramov and Martsinkowsky [AM] whether the following conditions which characterize the Gorenstein-projective modules, are independent.

(G1) The A-module M is semi-Gorenstein-projective.

(G2) The A-dual $M^* = \mathrm{Hom}(M, {}_AA)$ of M is semi-Gorenstein-projective.

(G3) The A-module M is reflexive.

Theorem. *For artin algebras, the conditions* (G1), (G2) *and* (G3) *are independent.*

Proof. Theorem 1.6 provides a $\Lambda(q)$-module M (namely $M = M(q)$) satisfying the conditions (G1), (G2) and not (G3). It remains to use the following proposition. □

Proposition. *If a module M is semi-Gorenstein-projective and not Gorenstein-projective, then $\Omega^2 M$ satisfies* (G1) *and* (G3)*, but not* (G2).

If a module M' satisfies (G1) *and* (G3)*, but not* (G2)*, then $N = (M')^*$ is a right module satisfying* (G2) *and* (G3)*, but not* (G1).

Proof. Let M be semi-Gorenstein-projective and not Gorenstein-projective. Then $\Omega^2 M$ is reflexive and semi-Gorenstein-projective. By Lemma 2.5, $(\Omega^2 M)^* = \mathrm{Tr}\, M$. Thus $(\Omega^2 M)^*$ is not semi-Gorenstein-projective (otherwise, M is Gorenstein-projective).

If M' satisfies (G1) and (G3), but not (G2), then $(M')^*$ is reflexive and $(M')^{**} = M'$ is semi-Gorenstein-projective, i.e., $N = (M')^*$ satisfies (G2) and (G3), but not (G1). □

Actually, for our example $A = \Lambda(q)$, there is also an A-module which satisfies (G2), (G3), but not (G1), namely the module $M(1)$, see 7.5.

In [JS], where Jorgensen and Şega present the first example of a semi-Gorenstein-projective module which is not Gorenstein-projective, they also exhibited modules satisfying (G1), (G3), but not (G2), and modules satisfying (G2), (G3), but not (G1). The algebra A considered in [JS] is commutative. It is an open problem whether there exists a **commutative** ring A with a module M satisfying (G1), (G2), but not (G3). The forthcoming paper [RZ4] will be devoted to a better understanding of the modules M with both M and M^* being semi-Gorenstein-projective.

1.8. Outline of the paper. The proofs of theorems 1.2, 1.3 and 1.4 are given in sections 2, 3 and 5, respectively. We use what we call (as a shorthand) *approximation sequences*, namely exact sequences

$$0 \to X \to Y \to Z \to 0$$

with Y projective and $\mathrm{Ext}^1(Z, A) = 0$, see section 2. Of special interest are the approximation sequences with both X and Z indecomposable and nonprojective; in this case, we have $X = \Omega Z$ and $Z = \mho X$, and we call them $\mho$-*sequences,* see section 3.

Section 4 deals with the $\mho$-quiver of A. An essential ingredient in this setting seems to be Corollary 4.4. The corresponding Remark 1 in 4.4 asserts that the kernel of the canonical map $\mho^t M \to (\mho^t M)^{**}$ is equal to $\mathrm{Ext}^{t+1}(\mathrm{Tr}\, M, A_A)$, and its cokernel is equal to $\mathrm{Ext}^{i+2}(\mathrm{Tr}\, M, A_A)$, for all $t \geq 0$.

In sections 6 and 7, we present the 6-dimensional algebra $\Lambda = \Lambda(q)$ depending on a parameter $q \in k \setminus \{0\}$, which we need for Theorem 1.6. We analyze some 3-dimensional representations which we denote by $M(\alpha)$ with $\alpha \in k$. Essential properties of the modules $M(\alpha)$ can be found in 6.3 to 6.5; they are labeled by (1) to (9). The properties (1) to (5) in 6.3 are those which are needed in order to exhibit a module, namely $M(q)$, which is semi-Gorenstein-projective, but not torsionless, provided the multiplicative order of q is infinite (see 6.4). The remaining properties (6) to (9) in 6.5 show, in particular, that in case the multiplicative order of q is infinite, also the Λ-dual $M(q)^*$ of $M(q)$ is semi-Gorenstein-projective. The proof of Theorem 1.6 and its Addendum is given in 6.7 and 6.8. In 7.1 and 7.2, some components of the $\mho$-quivers of the algebras Λ and Λ^{op} are described.

The final sections 8 and 9 add remarks and mention some open questions.

1.9. Terminology. We end the introduction with some remarks concerning the terminology and its history. The usual reference for the introduction of Gorenstein-projective modules (under the name *modules of Gorenstein dimension zero*) is the Memoirs by Auslander and Bridger [AB] in 1969. Actually, in his thesis [Br], Bridger attributes the concept of the Gorenstein dimension to Auslander: In January 1967, Auslander gave four lectures at the Séminaire Pierre Samuel (see the notes [A1] by Mangeney, Peskine and Szpiro). In these lectures, he discussed the class of all reflexive modules M such that both M and M^* are semi-Gorenstein-projective modules and denoted it by $G(A)$ [A1, Definition 3.2.2]. Thus $G(A)$ is the class of the Gorenstein-projective modules and the conditions (G1), (G2) and (G3) served as the first definition. In [AB, Proposition 3.8], it is shown that a module M belongs to $G(A)$ if and only if both M and $\mathrm{Tr}\, M$ are semi-Gorenstein-projective. Of course, we should stress the following: Whereas some formulations in [AB] assume that the ring A in question is a commutative noetherian ring, all the essential considerations in [A1, Br, AB] are shown in the setting of an abelian category with enough projectives, and of the category of finitely generated modules over a, not necessarily commutative, noetherian ring. Enochs and Jenda [EJ1, EJ2] have reformulated the definition of Gorenstein-projective modules in terms of complete projective resolutions, see also [Chr]. Several other names for the Gorenstein-projective modules are in use, they are also called “totally reflexive” modules [AM], and “maximal Cohen–Macaulay” modules [Buch] and “Cohen–Macaulay” modules [B2].

We should apologize that we propose a new name for the modules M with $\mathrm{Ext}^i(M, A) = 0$ for all $i \geq 1$, namely *semi-Gorenstein-projective*. These modules have been called for example “Cohen–Macaulay modules” or “stable modules”. However, the name “Cohen–Macaulay module” is in conflict with its established use for commutative rings, and, in our opinion, the wording “stable” may be too vague as a proper identifier. We hope that the name *semi-Gorenstein-projective* describes well what is going on: that there is something like a half of a complete projective resolution (“semi” means “half”). We also propose the name *left weakly Gorenstein* for an algebra A with $\mathrm{gp}(A) = {}^{\perp}A$ (in contrast to “nearly Gorenstein” in [M2]); of course, a Gorenstein algebra A satisfies $\mathrm{gp}(A) = {}^{\perp}A$, but the algebras with $\mathrm{gp}(A) = {}^{\perp}A$ seem to be quite far away from being Gorenstein. The left weakly Gorenstein algebras have also been called “algebras with condition (GC)” in [CH].

2. Approximation sequences. Proof of Theorem 1.2

2.1. Lemma. *Let* $\epsilon : 0 \to X \xrightarrow{\omega} Y \xrightarrow{\pi} Z \to 0$ *be an exact sequence with Y projective. Then the following conditions are equivalent:*

(i) *ω is a left* $\mathrm{add}(A)$*-approximation.*

(ii) $\mathrm{Ext}^1(Z, A) = 0$.

(iii) *The A-dual sequence ϵ^* of ϵ is exact.*

An exact sequence $0 \to X \to Y \to Z \to 0$ with Y projective satisfying the equivalent properties will be called in this paper an *approximation sequence* (this is just a shorthand, since it is too vague to be used in general). One may observe that the conditions (i), (ii) and (iii) are equivalent for any exact sequence $\epsilon : X \xrightarrow{\omega} Y \to Z \to 0$ with Y projective, even if ω is not injective, but we are only interested in the short exact sequences.

Proof of the equivalence of the properties. Since Y is projective, applying $\mathrm{Hom}(-, A)$ to ϵ we get the exact sequence $0 \to Z^* \xrightarrow{\pi^*} Y^* \xrightarrow{\omega^*} X^* \to \mathrm{Ext}^1(Z, A) \to 0$. Note that ω is a left $\mathrm{add}(A)$-approximation if and only if ω^* is surjective. From this we get the equivalence of (i) and (ii) and the equivalence of (ii) and (iii). □

2.2. Also the following basic lemma is well-known (see, for example [R]).

Lemma. *Let* $P_{-1} \xrightarrow{f} P_0 \xrightarrow{g} P_1$ *be an exact sequence of projective modules and let $g = up$ be a factorization with $p : P_0 \to I$ epi and $u : I \to P_1$ mono. Then* $P_{-1}^* \xleftarrow{f^*} P_0^* \xleftarrow{g^*} P_1^*$ *is exact if and only if u is a left* $\mathrm{add}(A)$*-approximation.*

For the convenience of the reader, we insert the proof.

Proof. Since $f^*g^* = (gf)^* = 0$, we have $\mathrm{Im}\, g^* \subseteq \mathrm{Ker}\, f^*$. Assume now that u is a left $\mathrm{add}(A)$-approximation and let $h \in \mathrm{Ker}\, f^*$, thus $hf = 0$. Since p is a cokernel of f, there is h' with $h = h'p$. Since u is a left $\mathrm{add}(A)$-approximation, there is h'' with $h' = h''u$. Thus $h = h'p = h''up = h''g = g^*(h'')$ belongs to the image of g^*, there also $\mathrm{Ker}\, f^* \subseteq \mathrm{Im}\, g^*$.

Conversely, we assume that $\operatorname{Im} g^* = \operatorname{Ker} f^*$ and let $h : I \to A$ be a map. Then $hpf = 0$, so that $f^*(hp) = 0$. Therefore hp belongs to $\operatorname{Ker} f^*$, thus to $\operatorname{Im} g^*$. There is $h'' \in P_1^*$ with $hp = g^*(h'') = h''g = h''up$, and therefore $h = h''u$. □

This Lemma will be used in various settings, see 4.3.

2.3. *A semi-Gorenstein-projective and Ω-periodic module is Gorenstein-projective.*

Proof. Let M be semi-Gorenstein-projective and assume that $\Omega^t M = M$ for some $t \geq 1$. Let $\cdots \to P_i \to \cdots \to P_0 \to M \to 0$ be a minimal projective resolution of M. Then

$$0 \to \Omega^t M \to P_{t-1} \to \cdots \to P_0 \to M \to 0 \qquad (+)$$

is the concatenation of approximation sequences. Since $\Omega^t M = M$, we can concatenate countably many copies of $(+)$ in order to obtain a double infinite acyclic chain complex of projective modules. As a concatenation of approximation sequences, it is a complete projective resolution. Therefore, M is Gorenstein-projective. □

2.4. Here are two essential observations.

(a) *Let $0 \to X \to Y \to Z \to 0$ be an approximation sequence. Then ϕ_X is surjective if and only if Z is torsionless.* We can also say: *X is reflexive if and only if Z is torsionless.*

(b) *Let $0 \to X \to Y \to Z \to 0$ be an approximation sequence. Then* $\operatorname{Ext}^1(X^*, A_A) = 0$ *if and only if ϕ_Z is surjective.*

Proof of (a) and (b). Since $0 \to X \xrightarrow{\omega} Y \xrightarrow{\pi} Z \to 0$ is an approximation sequence, it follows that

$$0 \to Z^* \xrightarrow{\pi^*} Y^* \to X^* \to 0$$

is an exact sequence of right A-modules. This induces an exact sequence

$$0 \to X^{**} \to Y^{**} \xrightarrow{\pi^{**}} Z^{**} \to \operatorname{Ext}^1_A(X^*, A_A) \to 0$$

of left A-modules, and we obtain the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & X \arrow[r] \arrow[d, "\phi_X"] & Y \arrow[r, "\pi"] \arrow[d, equal] & Z \arrow[r] \arrow[d, "\phi_Z"] & 0 & \\
0 \arrow[r] & X^{**} \arrow[r] & Y^{**} \arrow[r, "\pi^{**}"] & Z^{**} \arrow[r] & \operatorname{Ext}^1(X^*, A_A) \arrow[r] & 0.
\end{tikzcd}$$

By the Snake lemma, the kernel of ϕ_Z is isomorphic to the cokernel of ϕ_X, Thus ϕ_Z is a monomorphism if and only if ϕ_X is an epimorphism. Since X is torsionless, X is reflexive if and only if ϕ_X is surjective. This is (a).

By the commutative diagram above, we see that ϕ_Z is epic if and only if so is π^{**}, and if and only if $\operatorname{Ext}^1_A(X^*, A_A) = 0$. This is (b). □

Corollary. *A module X is reflexive if and only if both X and $\mho X$ are torsionless.*

Proof. If X is reflexive, then it is torsionless. Thus we may assume from the beginning that X is torsionless. Any minimal left $\mathrm{add}(A)$-approximation $X \to Y$ is injective and its cokernel is $\mho X$. The exact sequence $0 \to X \to Y \to \mho X \to 0$ is an approximation sequence, and 2.4(a) asserts that X is reflexive if and only if $\mho X$ is torsionless. $\square$

Remark. The assertion of the corollary can be strengthened as follows. For any module X, let us denote by $\mathrm{K}\,X$ the kernel of the map $\phi_X : X \to X^{**}$. Of course, $\mathrm{K}\,X$ is the kernel of any left $\mathrm{add}(A)$-approximation of X. Therefore X is torsionless if and only if $\mathrm{K}\,X = 0$. Claim: *The cokernel of the map $\phi_X : X \to X^{**}$ is isomorphic to* $\mathrm{K}\,\mho X$.

Proof. Let $u : X \to Y$ be a minimal $\mathrm{add}(A)$ approximation, say with cokernel $p : Y \to \mho X$. The A-dual of the exact sequence $X \xrightarrow{u} Y \xrightarrow{p} \mho X \to 0$ is $0 \leftarrow X^* \xleftarrow{u^*} Y^* \xleftarrow{p^*} (\mho X)^* \leftarrow 0$, since u is an $\mathrm{add}(A)$-approximation. Using again A-duality, we obtain the exact sequence $0 \to X^{**} \xrightarrow{u^{**}} Y^{**} \xrightarrow{p^{**}} (\mho X)^{**}$. Thus there is the following commutative diagram with exact rows:

$$\begin{tikzcd}
 & X \arrow[r, "u"] \arrow[d, "\phi_X"] & Y \arrow[r, "\pi"] \arrow[d, "\phi_Y"] & \mho X \arrow[r] \arrow[d, "\phi_{\mho X}"] & 0 \\
0 \arrow[r] & X^{**} \arrow[r, "u^{**}"] & Y^{**} \arrow[r, "\pi^{**}"] & \mho X^{**}. &
\end{tikzcd}$$

Since ϕ_Y is an isomorphism, the snake lemma yields $\mathrm{Cok}\,\phi_X \simeq \mathrm{Ker}(\phi_{\mho X}) = \mathrm{K}\,\mho X$. $\square$

In 4.4, we will rewrite both $\mathrm{K}\,X$ and $\mathrm{K}\,\mho X$ in order to obtain the classical Auslander–Bridger sequence (see Corollary and Remark 1 in 4.4).

2.5. Lemma. *Let M be a module with* $\mathrm{Ext}^i(M, A) = 0$ *for $i = 1, 2$. Then* $\mathrm{Tr}\,M \simeq (\Omega^2 M)^*$ *and there is a projective module Y such that* $M^* \simeq \Omega^2\,\mathrm{Tr}\,M \oplus Y$.

Proof. Let $\pi : PM \to M$ and $\pi' : P\Omega M \to \Omega M$ be projective covers with inclusion maps $\omega : \Omega M \to PM$ and $\omega' : \Omega^2 M \to P\Omega M$. Then $\omega\pi'$ is a minimal projective presentation of M. By definition, $\mathrm{Tr}\,M$ is the cokernel of $(\omega\pi')^*$. Since $\mathrm{Ext}^i(M, A) = 0$ for $i = 1, 2$, the exact sequences

$$0 \to \Omega^2 M \xrightarrow{\omega'} P\Omega M \xrightarrow{\pi'} \Omega M \to 0, \quad 0 \to \Omega M \xrightarrow{\omega} PM > \pi >> M \to 0$$

are approximation sequences. As a consequence, the corresponding A-dual sequences

$$0 \leftarrow (\Omega^2 M)^* \xleftarrow{(\omega')^*} (P\Omega M)^* \xleftarrow{(\pi')^*} (\Omega M)^* \leftarrow 0, \quad 0 \leftarrow (\Omega M)^* \xleftarrow{\omega^*} (PM)^* \xleftarrow{\pi^*} M^* \leftarrow 0$$

are exact. The concatenation

$$\epsilon : \qquad 0 \leftarrow (\Omega^2 M)^* \xleftarrow{(\omega')^*} (P\Omega M)^* \xleftarrow{(\omega\pi')^*} (PM)^* \xleftarrow{\pi^*} M^* \leftarrow 0$$

shows that $(\Omega^2 M)^*$ is a cokernel of $(\omega\pi')^*$, thus $\mathrm{Tr}\,M \simeq (\Omega^2 M)^*$. In addition, ϵ shows that $\Omega^2\,\mathrm{Tr}\,M = \Omega^2(\Omega^2 M)^*$ is the direct sum of M^* and a projective module Y. $\square$

2.6. *Proof of Theorem 1.2*

(1) implies (2) to (7): This follows directly from well-known properties of Gorenstein-projective modules. Namely, assume (1) and let M be Gorenstein-projective. Then M is reflexive, this yields (3), but, of course, also (2) and (4). Second, M^* is Gorenstein-projective, thus semi-Gorenstein-projective, therefore we get (5) and (6). Finally, $\operatorname{Tr} M$ is Gorenstein-projective, thus semi-Gorenstein-projective, therefore we get (7).

Both (3) and (4) imply (2): Let M be semi-Gorenstein-projective. Consider the approximation sequence $0 \to \Omega M \to PM \to M \to 0$ and note that ΩM is again semi-Gorenstein-projective. If (3) or just (4) holds, we know that $\phi_{\Omega M}$ is surjective, thus by 2.4 (a), M is torsionless.

Both (6) and (7) imply (2): Let M be semi-Gorenstein-projective. Consider the approximation sequences $0 \to \Omega M \to PM \to M \to 0$ and $0 \to \Omega^2 M \to PM \to \Omega M \to 0$. Since M is semi-Gorenstein-projective, also $\Omega^2 M$ is semi-Gorenstein-projective. If (6) holds, we use (6) for $\Omega^2 M$ in order to see that $\operatorname{Ext}^1((\Omega^2 M)^*, A_A) = 0$. If (7) holds, we use (7) for M in order to see that $\operatorname{Ext}^1(\operatorname{Tr} M, A_A) = 0$. According to 2.5, we see that $\operatorname{Tr} M = (\Omega^2 M)^*$. Thus in both cases (6) and (7), we have $\operatorname{Ext}^1((\Omega^2 M)^*, A_A) = 0$. According to 2.4 (b), it follows from $\operatorname{Ext}^1((\Omega^2 M)^*, A_A) = 0$ that $\phi_{\Omega M}$ is surjective. By 2.4 (a), M is torsionless.

Trivially, (5) implies (6). Altogether we have shown that any one of the assertions (3) to (7) implies (2).

It remains to show that (2) implies (1). Let M be semi-Gorenstein-projective and torsionless. We want to show that M is Gorenstein-projective. Let $M_i = \mho^i M$ for all $i \geq 0$ (with $M_0 = M$). Since M_0 is torsionless, there is an approximation sequence $0 \to M_0 \to P_1 \to M_1 \to 0$, and M_1 is again semi-Gorenstein-projective. By assumption, M_1 is again torsionless. Inductively, starting with a torsionless module M_i, we obtain an approximation sequence $\epsilon_i : 0 \to M_i \to P_{i+1} \to M_{i+1} \to 0$, we conclude that with M_i also M_{i+1} is semi-Gorenstein-projective. By (2) we see that M_{i+1} is torsionless, again. Concatenating a minimal projective resolution of M with these approximation sequences ϵ_i, for $0 \leq i$, we obtain a complete projective resolution of M. $\square$

3. $\mho$-sequences. Proof of Theorem 1.3

3.1. An approximation sequence $0 \to X \to Y \to Z \to 0$ will be called an $\mho$-*sequence* provided that both X and Z are indecomposable and not projective (the relevance of such sequences was stressed already in [RX]).

Lemma. *An approximation sequence is the direct sum of $\mho$-sequences and an exact sequence* $0 \to X' \to Y' \to Z' \to 0$ *with* X', Z' *(thus also* Y'*) being projective.*

Proof. Let $0 \to X \xrightarrow{\omega} Y \xrightarrow{\pi} Z \to 0$ be an approximation sequence. Since Y is projective and π is surjective, a direct decomposition $Z = Z_1 \oplus Z_2$ yields a direct sum decomposition of the sequence. Since ω is a left $\operatorname{add}(A)$-approximation, there is also the corresponding assertion: If $X = X_1 \oplus X_2$, then $X \xrightarrow{\omega} Y$ is the direct sum of two maps $X_1 \to Y_1$ and $X_2 \to Y_2$, thus again we obtain a direct sum decomposition of the sequence. This shows that for an indecomposable approximation sequence $0 \to X \xrightarrow{\omega} Y \xrightarrow{\pi} Z \to 0$, the modules X and Z are indecomposable or zero (and, of course, not both can be zero).

If Z is indecomposable and projective, then the sequence $0 \to X \to Y \to Z \to 0$ splits off $0 \to 0 \to Z \xrightarrow{1} Z \to 0$, thus $X = 0$. Similarly, if X is indecomposable and projective, then the sequence $0 \to X \to Y \to Z \to 0$ splits off $0 \to X \xrightarrow{1} X \to 0 \to 0$, thus $Z = 0$.

It remains that $0 \to X \to Y \to Z \to 0$ is an approximation sequence with both X and Z being indecomposable and nonprojective. $\square$

3.2. Lemma. *Let* $\epsilon : 0 \to X \xrightarrow{\omega} Y \xrightarrow{\pi} Z \to 0$ *be an exact sequence. The following conditions are equivalent:*

(i) ϵ *is an* $\mho$*-sequence.*

(ii) X *and* Z *are indecomposable and not projective,* ω *is a minimal left* $\operatorname{add}(A)$*-approximation,* π *is a projective cover,* $X = \Omega Z$, $Z = \mho X$.

(iii) X *is indecomposable and not projective,* ω *is a minimal left* $\operatorname{add}(A)$*-approximation.*

(iv) Z *is indecomposable and not projective,* π *is a projective cover, and* $\operatorname{Ext}^1(Z, A) = 0$.

(v) $X = \Omega Z$, Y *is projective,* $Z = \mho X$, *and* X *is indecomposable.*

(vi) $X = \Omega Z$, Y *is projective,* $Z = \mho X$, *and* Z *is indecomposable.*

Proof. (i) implies (ii): Let ϵ be an $\mho$-sequence. Then ω has to be minimal, since otherwise ϵ would split off a nonzero sequence of the form $0 \to 0 \to P \xrightarrow{1} P \to 0$ with P projective. Similarly, π has to be a projective cover, since otherwise ϵ would split off a nonzero sequence of the form $0 \to P \xrightarrow{1} P \to 0 \to 0$. Since ω is a minimal left $\operatorname{add}(A)$-approximation and Z is the cokernel of ω, we see that $Z = \mho X$. Since π is a projective cover of Z and X is its kernel, $X = \Omega Z$.

(ii) collects all the relevant properties of an $\mho$-sequence. The condition (iii), (iv), (v) and (vi) single out some of these properties, thus (ii) implies these conditions.

(iii) implies (i): Since X is indecomposable and not projective, ϵ has no direct summand $0 \to P \xrightarrow{1} P \to 0 \to 0$. Since ω is left minimal, ϵ has no direct summand $0 \to 0 \to P \xrightarrow{1} P \to 0$. Similarly, (iv) implies (i).

Both (v) and (vi) imply (i): Since $Z = \mho X$, we have $\operatorname{Ext}^1(Z, A) = 0$. This shows that the sequence is an approximation sequence. Since $X = \Omega Z$, the sequence ϵ has no direct summand of the form $0 \to P \xrightarrow{1} P \to 0 \to 0$. Since $Z = \mho X$, the sequence ϵ has no direct summand of the form $0 \to 0 \to P \xrightarrow{1} P \to 0$. Thus, ϵ is a direct sum of $\mho$-sequences. Finally, since X or Z is indecomposable, ϵ is an $\mho$-sequence. $\square$

3.3. Corollary. *If* M *is indecomposable, nonprojective, semi-Gorenstein-projective, then* ΩM *is indecomposable, nonprojective, semi-Gorenstein-projective and* $M = \mho\Omega M$.

Proof. Since M is semi-Gorenstein-projective module, the canonical sequence $\epsilon : 0 \to \Omega M \to PM \to M \to 0$ is an approximation sequence. Since M is indecomposable and not projective, and $PM \to M$ is a projective cover, ϵ is an $\mho$-sequence, thus ΩM is indecomposable and nonprojective, and $M = \mho\Omega M$, by 3.2. Of course, with M also ΩM is semi-Gorenstein-projective. $\square$

3.4. Lemma. *If the number of isomorphism classes of indecomposable modules which are both semi-Gorenstein-projective and torsionless is finite, then any indecomposable nonprojective semi-Gorenstein-projective module M is Ω-periodic.*

Proof. According to 3.3, the modules $\Omega^t M$ with $t \geq 1$ are indecomposable modules which are torsionless and semi-Gorenstein-projective. Since there are only finitely many isomorphism classes of indecomposable torsionless semi-Gorenstein-projective modules, there are natural numbers $1 \leq s < t$ with $\Omega^s M = \Omega^t M$. Then

$$M = \mho^s \Omega^s M = \mho^s \Omega^t M = \Omega^{t-s} M \quad \text{and} \quad t - s \geq 1,$$

thus M is Ω-periodic. □

3.5. *Proof of Theorem 1.3* We assume that the number of isomorphism classes of indecomposable torsionless semi-Gorenstein-projective modules is finite. According to 3.4, any indecomposable nonprojective semi-Gorenstein-projective module is Ω-periodic. 2.3 shows that any semi-Gorenstein-projective Ω-periodic module is Gorenstein-projective. □

Remark. One of the referees has pointed out that Theorem 1.3 can be improved by replacing the class of all torsionless modules by an arbitrary full subcategory $\mathcal{X}$ which is closed under direct summands, contains $\operatorname{add}(A)$, and contains for any indecomposable module M at least one syzygy module $\Omega^n M$. *If ${}^{\perp}A \cap \mathcal{X}$ contains only finitely many isomorphism classes of indecomposable modules, then Λ is left weakly Gorenstein and any Gorenstein-projective module is Ω-periodic.*

3.6. Torsionless-finite algebras. An artin algebra A is said to be *torsionless-finite* if there are only finitely many isomorphism classes of indecomposable torsionless modules. Theorem 1.3 implies that *any torsionless-finite artin algebra is left weakly Gorenstein,* as Marczinzik [M1] has shown. If Λ is torsionless-finite, also Λ^{op} is torsionless-finite [R], thus a torsionless-finite artin algebra is also right weakly Gorenstein. Note that many interesting classes of algebras are known to be torsionless-finite. In particular, we have

The following algebras are torsionless-finite, hence left and right weakly Gorenstein.

(1) *Algebras A such that $A/\operatorname{soc}({}_A A)$ is representation-finite.*

(2) *Algebras stably equivalent to hereditary algebras, in particular all algebras with radical square zero.*

(3) *Minimal representation-infinite algebras.*

(4) *Special biserial algebras without indecomposable projective-injective modules.*

See for example [R], where also other classes of torsionless-finite algebras are listed.

Chen [Che] has shown that a connected algebra A with radical square zero either is self-injective, or else all the Gorenstein-projective modules are projective. The assertion that algebras with radical square zero are weakly Gorenstein complements this result.

4. The $\mho$-quiver

4.1. We recall that the $\mho$-quiver of A has as vertices the isomorphism classes $[X]$ of the indecomposable nonprojective modules X and *there is an arrow*

$$[X] \dashleftarrow [Z]$$

provided that X is torsionless and $Z = \mho X$, thus provided that there exists an $\mho$-sequence $0 \to X \to Y \to Z \to 0$. We will also write the vertex $[X]$ simply as X.

4.2. The A-dual of an $\mho$-sequence.

Lemma. (a) *Let $\epsilon : 0 \to X \to Y \to Z \to 0$ be an approximation sequence and assume that X is reflexive. Then* $\operatorname{Ext}^1(X^*, A_A) = 0$ *if and only if Z is reflexive, if and only if the A-dual ϵ^* of ϵ is again an approximation sequence.*

(b) *Let $\epsilon : 0 \to X \to Y \to Z \to 0$ be an $\mho$-sequence with X reflexive. Then Z is reflexive, if and only if the A-dual ϵ^* of ϵ is again an $\mho$-sequence.*

Proof. (a) By 2.4(a), we see that Z is always torsionless. Thus 2.4(b) shows that $\operatorname{Ext}^1(X^*, A_A) = 0$ if and only if Z is reflexive. First, assume that Z is reflexive. Then $\operatorname{Ext}^1(X^*, A_A) = 0$, and therefore we see that the A-dual sequence ϵ^* is exact. We dualize a second time: the sequence ϵ^{**} is isomorphic to the sequence ϵ, since the three modules X, Y, Z are reflexive. This means that ϵ^{**} is exact, and therefore ϵ^* is an approximation sequence. Second, conversely, if ϵ^* is an approximation sequence, then it is exact, and therefore $\operatorname{Ext}^1(X^*, A_A) = 0$, thus Z is reflexive.

(b) Assume now that ϵ is an $\mho$-sequence. First, assume that Z is reflexive. Since X, Z both are reflexive, indecomposable and nonprojective, also X^* and Z^* are indecomposable and nonprojective, as we will show below. Thus ϵ^* is an $\mho$-sequence. Conversely, if ϵ is an $\mho$-sequence, then it is an approximation sequence and thus Z is reflexive by (a). $\square$

We have used some basic facts about the A-dual M^* of a module M.

(1) *M^* is always torsionless.*

(2) *If M is nonzero and torsionless, then M^* is nonzero.*

(3) *If M is reflexive, indecomposable and nonprojective, then M^* is reflexive, indecomposable and nonprojective.*

Proof. Here are the proofs (or see for example [L, p.144]). (1) There is a surjective map $p : P \to M$ with P projective. Then $p^* : M^* \to P^*$ is an embedding of M^* into the projective module P^*. The assertion (2) is obvious.

(3) Let M be reflexive, indecomposable and nonprojective. Consider a direct decomposition $M^* = N_1 \oplus N_2$ with $N_1 \neq 0$ and $N_2 \neq 0$. Since M^* is torsionless by (1), both modules N_1 and N_2 are torsionless, therefore $N_1^* \neq 0$, $N_2^* \neq 0$, thus there is a proper direct decomposition $M^{**} = N_1^* \oplus N_2^*$. Since M is

reflexive and indecomposable, this is impossible. Thus M^* has to be indecomposable. If M^* is projective, then also M^{**} is projective. Again, since M is reflexive, this is impossible.

It remains to show that M^* is reflexive. Since M^{**} is isomorphic to M, we see that M^{***} is isomorphic to M^*, thus the canonical map $M^* \to M^{***}$ has to be an isomorphism (since it is a monomorphism of modules of equal length). □

4.3. Lemma 2.2 outlines the importance of left add(A)-approximations when dealing with exact sequences of projective modules. Let us give a unified treatment of the relevance of approximation sequences and of $\mho$-sequences.

(a) *An exact sequence* $\cdots \to P^{-1} \to P^0 \xrightarrow{d^0} P^1 \to \cdots$ *is a complete projective resolution if and only if it is the concatenation of approximation sequences.*

(a′) *An indecomposable nonprojective module M is Gorenstein-projective if and only if $[M]$ is the start of an infinite $\mho$-path and the end of an infinite $\mho$-path.*

(b) *A module M is semi-Gorenstein-projective if and only if a projective resolution (or, equivalently, any projective resolution) is the concatenation of approximation sequences.*

(b′) *An indecomposable nonprojective module M is semi-Gorenstein-projective if and only if $[M]$ is the start of an infinite $\mho$-path.*

(c) *A module M is reflexive and M^* is semi-Gorenstein-projective if and only if there is an exact sequence* $0 \to M \to P^1 \to P^2 \to \cdots$ *which is the concatenation of approximation sequences.*

(c′) *An indecomposable nonprojective module M is reflexive and M^* is semi-Gorenstein-projective if and only if $[M]$ is the end of an infinite $\mho$-path.*

Proof. We use that the A-dual of an approximation sequence is exact, thus the A-dual of the concatenation of approximation sequences is exact.

(a) Let $P^\bullet$ be a double infinite exact sequence of projective modules with maps $d^i : P^i \to P^{i+1}$. Write $d^i = \omega^i \pi^i$ with π^i epi and ω^i mono. If $P^\bullet$ is a complete projective resolution, then the exactness of $(P^\bullet)^*$ at $(P^i)^*$ implies that ω^i is a left add(A)-approximation, see 2.2. Thus $P^\bullet$ is the concatenation of approximation sequences.

(b) Let $\cdots \to P_i \to \cdots \to P_1 \to P_0 \to M \to 0$ be a projective resolution of M. write the map $P_{i+1} \to P_i$ as $\omega_i \pi_i$ with π_i epi and ω_i mono. If the A-dual of the sequence $\cdots \to P_i \to \cdots \to P_0$ is exact, then all the maps ω_i with $i \geq 1$ have to be left add(A)-approximations. This shows that the projective resolution is the concatenation of approximation sequences.

(b′) Let M be indecomposable, nonprojective and semi-Gorenstein-projective. Since $\operatorname{Ext}^1(M, A) = 0$, the sequence $0 \to \Omega M \to PM \to M \to 0$ is an $\mho$-sequence and ΩM is again indecomposable and nonprojective. Also, ΩM is semi-Gorenstein-projective. Thus, we can iterate the procedure and obtain the infinite path

$$(*) \qquad \cdots \dashleftarrow [\Omega^2 M] \dashleftarrow [\Omega M] \dashleftarrow [M]$$

Conversely, assume that there is an infinite path starting with $[M]$, then it is of the form $(*)$. Thus, for all $i \geq 1$, we have $\operatorname{Ext}^i(M, A) \simeq \operatorname{Ext}^1(\Omega^{i-1}M, A) = 0$.

Proof of (c) and (c′). Assume that there are given approximation sequences $\epsilon_i : 0 \to M^i \to P^{i+1} \to M^{i+1} \to 0$ for all $i \geq 0$, with $M^0 = M$. Then all the modules M^i are torsionless, thus reflexive by 2.4(a). In particular, M itself is reflexive. The A-dual of ϵ_i is the sequence

$$\epsilon_i^* : 0 \leftarrow (M^i)^* \leftarrow (P^{i+1})^* \leftarrow (M^{i+1})^* \leftarrow 0,$$

which again is an approximation sequence by 4.2(a). The concatenation of the sequences ϵ_i^* is a projective resolution of $M^* = (M^0)^*$. According to (b), M^* is semi-Gorenstein-projective, since all the sequences ϵ_i^* are approximation sequences.

Conversely, assume that M is reflexive and M^* is semi-Gorenstein-projective. We want to construct a sequence $0 \to M \to P^1 \to P^2 \to \cdots$ which is the concatenation of approximation sequences. It is sufficient to consider the case where M is indecomposable (in general, take the direct sum of the sequences). If M is projective, then $0 \to M \to M \to 0 \to \cdots$ is the concatenation of approximation sequences.

Thus, it remains to consider the case where M is indecomposable and not projective. Since M is torsionless, there is an $\mho$-sequence $\epsilon_0 : 0 \to M \to P^1 \to M^1 \to 0$ (with $M^1 = \mho M$). Note that M^1 is indecomposable, not projective, and that the A-dual $\epsilon_0^* : 0 \leftarrow M^* \leftarrow (P^1)^* \leftarrow (M^1)^* \leftarrow 0$ is exact. Since M is reflexive, M^1 is torsionless by 2.4(a). Since M^* is semi-Gorenstein-projective, $\operatorname{Ext}^1(M^*, A_A) = 0$, therefore ϕ_{M^1} is surjective and ϵ_0^* is an $\mho$-sequence, by 4.2. Altogether we know now that M^1 is reflexive, but also that $(M^1)^* = \Omega(M^*)$. With M^* also $\Omega(M^*)$ is semi-Gorenstein-projective.

Thus M^1 satisfies again the assumptions of being indecomposable, not projective, reflexive and that its A-dual $(M^1)^*$ is semi-Gorenstein-projective. Thus we can iterate the procedure for getting the next $\mho$-sequence $\epsilon_1 : 0 \to M^1 \to P^2 \to M^2 \to 0$, with $M^2 = \mho^2 M$, and so on. Altogether, we obtain the infinite path:

$$[M] \dashleftarrow [\mho M] \dashleftarrow [\mho^2 M] \dashleftarrow \cdots$$

This completes the proof of (c′) and thus also of (c).

(a′) This follows immediately from (b′) and (c′). □

4.4. For any module M, we have denoted by $\mathrm{K}\, M$ the kernel of $\phi_M : M \to M^{**}$. We are going to identify $\mathrm{K}\, M$ with $\operatorname{Ext}^1(\operatorname{Tr} M, A_A)$. Compare [A2, Proposition 6.3]. As a consequence, we see that $\mho M = \operatorname{Tr} \Omega \operatorname{Tr} M$.

Lemma. *Let M be a module. Then* $\operatorname{Ext}^1(\operatorname{Tr} M, A_A) \simeq \mathrm{K}\, M$ *and there is a right module Q such that* $\Omega \operatorname{Tr} M \simeq \operatorname{Tr} \mho M \oplus Q$. *As a consequence,* $\mho M \simeq \operatorname{Tr} \Omega \operatorname{Tr} M$, *thus* $\mho^t(M) \cong \operatorname{Tr} \Omega^t \operatorname{Tr}(M)$ *for* $t \geq 1$.

Proof. Let $P^0 \xrightarrow{f} P^1 \xrightarrow{p} M >\!\!>\!\!> 0$ be a minimal projective presentation of M. Thus $\operatorname{Tr} M$ is the cokernel of f^*. Let $g' : M \to P^2$ be a minimal left $\operatorname{add}(A)$-approximation. Then $\mathrm{K}\, M$ is the kernel of g', thus $g' = uq$, where $q : M \to M/\mathrm{K}\, M$ is the canonical projection and u is injective. Let $g = g'p = uqp$.

The composition

$$P^0 \xrightarrow{f} P^1 \xrightarrow{g} P^2$$

is zero and the homology $H(P^0 \xrightarrow{f} P^1 \xrightarrow{g} P^2)$ is just $\mathrm{K}\,M$, since

$$\mathrm{Ker}(g)/\,\mathrm{Im}(f) \simeq \mathrm{Ker}(qp)/\,\mathrm{Ker}(p) \simeq \mathrm{K}\,M.$$

We claim that the A-dual sequence

$$(P^0)^* \xleftarrow{f^*} (P^1)^* \xleftarrow{g^*} (P^2)^* \tag{$*$}$$

is exact. Since $gf = 0$, we have $f^*g^* = 0$. Conversely, let $h : P^1 \to A$ be in the kernel of f^*, thus $hf = 0$. Therefore h factors through $p = \mathrm{Cok}\, f$, say $h = h'p$ with $h' : M \to A$. Since uq is a left $\mathrm{add}(A)$-approximation, we obtain $h'' : P^2 \to A$ with $h' = h''uq$. Thus $h = h'p = h''uqp = h''g = g^*(h'')$ is in the image of g^*.

Since the cokernel of f^* is $\mathrm{Tr}\,M$, it follows that $(*)$ is the begin of a projective resolution of $\mathrm{Tr}\,M$ and hence $\mathrm{Ext}^1(\mathrm{Tr}\,M, A_A)$ is obtained by applying $\mathrm{Hom}(-, A)$ to $(*)$ and taking the homology at the position 1. Applying $\mathrm{Hom}(-, A)$ to $(*)$ we retrieve the sequence $P^0 \xrightarrow{f} P^1 \xrightarrow{g} P^2$, thus $\mathrm{Ext}^1(\mathrm{Tr}\,M, A_A)$ is equal to $H(P^0 \xrightarrow{f} P^1 \xrightarrow{g} P^2) \simeq \mathrm{K}\,M$. This is the first assertion.

By definition, the cokernel of g' is $\mho M$. Thus the cokernel of g is $\mho M$, and therefore $\mathrm{Cok}\, g^* \simeq \mathrm{Tr}\,\mho M \oplus Q'$ for some projective right module Q'. Now $\mathrm{Cok}\, g^* = \mathrm{Im}\, f^*$, since $(*)$ is exact. Since $\mathrm{Cok}\, f^* = \mathrm{Tr}\,M$, we have $\Omega\,\mathrm{Tr}\,M \simeq \mathrm{Im}\, f^* \oplus Q''$ for some projective right module Q''. This shows that $\Omega\,\mathrm{Tr}\,M \simeq \mathrm{Im}\, f^* \oplus Q'' = \mathrm{Cok}\, g^* \oplus Q'' \simeq \mathrm{Tr}\,\mho M \oplus Q' \oplus Q'' = \mathrm{Tr}\,\mho M \oplus Q$ with $Q = Q' \oplus Q''$. This is the second assertion.

Applying Tr to the isomorphism $\Omega\,\mathrm{Tr}\,M \simeq \mathrm{Tr}\,\mho M \oplus Q$, one obtains $\mathrm{Tr}\,\Omega\,\mathrm{Tr}\,M \simeq \mathrm{Tr}(\mathrm{Tr}\,\mho M \oplus Q) = \mathrm{Tr}\,\mathrm{Tr}\,\mho M$. Since $\mho M$ has no nonzero projective direct summand, one gets $\mathrm{Tr}\,\mathrm{Tr}\,\mho M \simeq \mho M$. Thus $\mho M \simeq \mathrm{Tr}\,\mathrm{Tr}\,\mho M \simeq \mathrm{Tr}\,\Omega\,\mathrm{Tr}\,M$. □

Corollary. *Let M be a module. Then for all $t \geq 0$ one has*

$$\mathrm{Ext}^{t+1}(\mathrm{Tr}\,M, A_A) \simeq \mathrm{K}(\mho^t M).$$

In particular, $\mho^t M$ is torsionless if and only if $\mathrm{Ext}^{t+1}(\mathrm{Tr}\,M, A_A) = 0$. Also, $\Omega^t\,\mathrm{Tr}\,M \simeq \mathrm{Tr}\,\mho^t M \oplus Q_t$ for some projective right module Q_t.

Proof. By induction on t, one has $\Omega^t\,\mathrm{Tr}\,M \simeq \mathrm{Tr}\,\mho^t M \oplus Q_t$ for some projective right module Q_t. It implies that $\mathrm{Ext}^{t+1}(\mathrm{Tr}\,M, A_A) \simeq \mathrm{Ext}^1(\Omega^t\,\mathrm{Tr}\,M, A_A) \simeq \mathrm{Ext}^1(\mathrm{Tr}\,\mho^t M, A_A)$ and thus $\mathrm{Ext}^1(\mathrm{Tr}\,\mho^t M, A_A) \simeq K(\mho^t M)$. □

Remark 1. *For any $t \geq 0$, there is an exact sequence of the form*

$$0 \to \mathrm{Ext}^{t+1}(\mathrm{Tr}\,M, A_A) \to \mho^t M \xrightarrow{\phi_{\mho^t M}} (\mho^t M)^{**} \to \mathrm{Ext}^{t+2}(\mathrm{Tr}\,M, A_A) \to 0.$$

If $t=0$, it is the classical Auslander–Bridger sequence

$$0 \to \mathrm{Ext}^1(\mathrm{Tr}\, M,\, A_A) \to M \to M^{**} \to \mathrm{Ext}^2(\mathrm{Tr}\, M,\, A_A) \to 0$$

(see [AB], also [ARS]).

Proof. The corollary asserts that the kernel of the map $\phi_{\mho^t M} : \mho^t M \to (\mho^t M)^{**}$ is isomorphic to $\mathrm{Ext}^{t+1}(\mathrm{Tr}\, M,\, A_A)$. On the other hand, the Remark at the end of 2.4 shows that $\mathrm{Cok}\,\phi_{\mho^t M} \simeq \mathrm{K}\,\mho^{t+1}M$. Since $\mathrm{K}\,\mho^{t+1}M \simeq \mathrm{Ext}^1(\mathrm{Tr}\,\mho^{t+1}M,\, A_A) \simeq \mathrm{Ext}^1(\Omega^{t+1}\,\mathrm{Tr}\, M,\, A_A) \simeq \mathrm{Ext}^{t+2}(\mathrm{Tr}\, M,\, A_A)$, it follows that $\mathrm{Cok}\,\phi_{\mho^t M} \simeq \mathrm{Ext}^{t+2}(\mathrm{Tr}\, M,\, A_A)$. □

Remark 2. *If M is any module, $\mho\,\mathrm{Tr}\, M \simeq \mathrm{Tr}\,\Omega M$.*

Proof. There is a projective module P such that $\mathrm{Tr}\,\mathrm{Tr}\, M \oplus P \simeq M$. According to Lemma 4.4 we have $\mho\,\mathrm{Tr}\, M \simeq \mathrm{Tr}\,\Omega\,\mathrm{Tr}\,\mathrm{Tr}\, M = \mathrm{Tr}\,\Omega\,(\mathrm{Tr}\,\mathrm{Tr}\, M \oplus P) \simeq \mathrm{Tr}\,\Omega M$. □

Remark 3. In contrast to the isomorphism given in Remark 2, the right modules $\Omega\,\mathrm{Tr}\, M$ and $\mathrm{Tr}\,\mho M$ discussed in the lemma do not have to be isomorphic. For example, let M be a module with $M^* = 0$. Then $\mho M = 0$, thus $\mathrm{Tr}\,\mho M = 0$. On the other hand, if $f : P_1 \to P(M)$ is a minimal projective presentation of M, then the kernel of f^* is M^*, thus zero, and therefore $\Omega\,\mathrm{Tr}\, M \simeq (P(M))^*$. Thus, we see that the right module Q with $\Omega\,\mathrm{Tr}\, M \simeq \mathrm{Tr}\,\mho M \oplus Q$ may be nonzero.

4.5. Modules at the end of an $\mho$-path of length t.

Proposition. *Let M be any module and $t \geq 1$. The following conditions are equivalent:*

(i) *$\mho^{i-1}M$ is torsionless for $1 \leq i \leq t$.*

(ii) *M is t-torsionfree (thus $\mathrm{Ext}^i(\mathrm{Tr}\, M,\, A_A) = 0$ for $1 \leq i \leq t$).*

If M is indecomposable and not projective, then these conditions are equivalent to

(iii) *M is the end of an $\mho$-path of length t.*

Already the special cases $t=1$ and $t=2$ are of interest (but well-known): *A module M is 1-torsionfree if and only if M is torsionless* (this is case $t=1$); *a module M is 2-torsionfree if and only if both M and ΩM are torsionless, thus if and only if M is reflexive* (this is the case $t=2$, taking into account Corollary 2.4). These special cases $t=1$ and $t=2$ are discussed at several places; let us refer in particular to [ARS], Corollary IV.3.3. Our general proof is inspired by [AB].

Proof of Proposition. For the equivalence of (i) and (ii), see Corollary in 4.4: It asserts for any $i \geq 1$, that $\mho^{i-1}M$ is torsionless if and only if $\mathrm{Ext}^i(\mathrm{Tr}\, M,\, A_A) = 0$.

In order to show the equivalence of (i) and (iii), let M be indecomposable and not projective. If (iii) is satisfied, there is an $\mho$-path of length t ending in M. This path has to be $\mho^t M,\ \mho^{t-1}M,\ \dots,\ \mho M,\ M$. This shows that for any module $\mho^i M$ with $0 \leq i < t$, there is an arrow starting in $\mho^i M$, and therefore $\mho^i M$ has to be torsionless.

Conversely, assume that (i) is satisfied. We show (iii) by induction on t. For any $t \geq 1$, there is the arrow $\mho M \to M$, since M is indecomposable, nonprojective and torsionless. According to 3.2, the module $\mho M$

is again indecomposable and nonprojective. Thus, if $t \geq 2$, we can use induction in order to obtain a path of length $t-1$ ending in $\mho M$, since all the modules $\mho^i(\mho M)$ with $0 \leq i < t-1$ are torsionless. $\square$

4.6. *Proof of Theorem 1.5* (1) follows from the fact that $\operatorname{Ext}^t(M, A) = \operatorname{Ext}^{t-1}(\Omega M, A)$ for $t \geq 2$. For the special case $t = 2$, see Corollary 2.4. (2) is Proposition 4.5. For (1′), (2′) and (3), see 4.3. For (4) and (5), we refer to 4.2(b). Note that in an $\mho$-component of the form $\mathbb{A}_n$ with $n \geq 3$, as well as in those of the form $-\mathbb{N}$, all but precisely two vertices are the isomorphism classes of reflexive modules, whereas any vertex of an $\mho$-component of the form $\mathbb{N}$ is the isomorphism class of a reflexive module. $\square$

4.7. The adjoint functors $\mho$ and Ω. Here we collect some important properties of the construction $\mho$. Some details of the proofs are left to the reader, since the assertions are not needed in the paper.

If $\mathcal{C}' \subseteq \mathcal{C}$ are full subcategories of mod A, let $\mathcal{C}/\mathcal{C}'$ be the category with the same objects as $\mathcal{C}$ such that $\operatorname{Hom}_{\mathcal{C}/\mathcal{C}'}(X, Y)$ is the factor group of $\operatorname{Hom}_{\mathcal{C}}(X, Y)$ modulo the subspace of all maps $X \to Y$ which factor through a direct sum of modules in $\mathcal{C}'$.

(1) *The functor $\mho$ is the left adjoint of the endo-functor Ω of* mod $A/$ add A. Direct verification is easy. But we should also add that Auslander and Reiten have shown in [AR2, Corollary 3.4] that the functor $\operatorname{Tr}\Omega\operatorname{Tr}$ is left adjoint to Ω, and we have identified in 4.4 the functors $\mho$ and $\operatorname{Tr}\Omega\operatorname{Tr}$.

(2) Let $\mathcal{L}(A)$ be the full subcategory of all torsionless modules, and $\mathcal{Z}(A)$ the full subcategory of all modules Z with $\operatorname{Ext}^1(Z, A) = 0$. *For any module M, the module ΩM belongs to $\mathcal{L}(A)$, and the module $\mho M$ belongs to $\mathcal{Z}(A)$; in addition, $\mho M$ has no nonzero projective direct summand.*

(3) *If Z satisfies* $\operatorname{Ext}^1(Z, A) = 0$ *and has no nonzero projective direct summand, then* $\mho\Omega Z \simeq Z$ (see 3.2). *If X is torsionless and has no nonzero projective direct summand, then* $\Omega\mho X \simeq X$ (see 1.5 or also 3.2). In this way, one shows that *the functors Ω and $\mho$ provide inverse categorical equivalences*

$$\mathcal{L}(A)/\operatorname{add}(A) \underset{\Omega}{\overset{\mho}{\rightleftarrows}} \mathcal{Z}(A)/\operatorname{add}(A)$$

(4) Thus, Ω *and* $\mho$ *provide inverse bijections between isomorphism classes as follows:*

$$\left\{\begin{matrix}\text{indecomposable}\\ \text{nonprojective modules } X\\ \text{which are torsionless}\end{matrix}\right\} \underset{\Omega}{\overset{\mho}{\rightleftarrows}} \left\{\begin{matrix}\text{indecomposable}\\ \text{nonprojective modules } Z\\ \text{with } \operatorname{Ext}^1(Z, A) = 0\end{matrix}\right\}$$

The arrows of the $\mho$-quiver visualize this bijection.

4.8. Gorenstein algebras. Recall that an artin algebra A is said to be *d-Gorenstein* provided that the injective dimension of both ${}_AA$ and A_A is equal to d. Of course, any algebra of global dimension d is d-Gorenstein The following result of Beligiannis [B2, Proposition 4.4] yields additional examples of weakly Gorenstein algebras.

Proposition. *Let A be an artin algebra and assume that the injective dimension of ${}_AA$ is at most d. Then A is right weakly Gorenstein and any module of the form $\Omega^d M$ is semi-Gorenstein projective.*

Proof. Since the injective dimension of ${}_AA$ is at most d, one knows that for any module M, the syzygy module $\Omega^d M$ is semi-Gorenstein-projective. [Namely, for all $i \geq 1$, we have $\mathrm{Ext}^i(\Omega^d M, A) = \mathrm{Ext}^{d+i}(M, A) = \mathrm{Ext}^i(M, \Sigma^d A) = 0$; here, ΣN denotes the cokernel of an injective envelope of a module N.] This implies that A cannot have any indecomposable module of $\mho$-type $\mathbb{N}$. [Namely, if M is of $\mho$-type $\mathbb{N}$, then M is ∞-torsionfree and therefore $M = \Omega^d(\mho^d M)$. But as we have seen, this implies that M is semi-Gorenstein-projective, therefore Gorenstein-projective. Thus M is of $\mho$-type $\mathbb{Z}$ and not $\mathbb{N}$.] Therefore A is right weakly Gorenstein. □

Corollary 1. *Let A be d-Gorenstein. If an indecomposable nonprojective module M belongs to an $\mho$-path of length d, then M is Gorenstein-projective. If the global dimension of A is d, then there is no $\mho$-path of length d.*

Proof. Since the inj. dim. ${}_AA = d$, A is right weakly Gorenstein and any module $\Omega^d M$ is semi-Gorenstein-projective. Since inj. dim. A_A is finite, A is also left weakly Gorenstein, thus the modules $\Omega^d M$ are even Gorenstein-projective. □

Corollary 2. *If A is d-Gorenstein, then A has no $\mho$-component of form $-\mathbb{N}$, $\mathbb{N}$ or $\mathbb{A}_n$ with $n > d$. If the global dimension of A is d, then any $\mho$-component is of form $\mathbb{A}_n$ with $n \leq d$.*

5. Proof of Theorem 1.4

Since $\mathrm{add}(A) \subseteq {}^\perp A \subseteq \mathcal{F}$, we see that $\mathrm{add}(A) \subseteq \mathcal{P}(\mathcal{F}) = \mathcal{I}(\mathcal{F})$. Thus $\mathrm{Ext}^1_A(X, A) = 0$, for all $X \in \mathcal{F}$.

For $X \in \mathcal{F}$, there is an exact sequence $0 \to K \to Q \to X \to 0$ with $Q \in \mathcal{P}(\mathcal{F})$ and $K \in \mathcal{F}$. By $\mathcal{P}(\mathcal{F}) \subseteq {}^\perp A$ we have $Q \in {}^\perp A$. Thus $\mathrm{Ext}^1_A(X, A) = 0$ and $\mathrm{Ext}^{m+1}_A(X, A) = \mathrm{Ext}^m_A(K, A)$ for $m \geq 1$. So $\mathrm{Ext}^2_A(X, A) = 0$, and in particular $\mathrm{Ext}^2_A(K, A) = 0$. Repeating this process we see that $X \in {}^\perp A$. Thus $\mathcal{F} \subseteq {}^\perp A$, and hence ${}^\perp A = \mathcal{F}$ is Frobenius.

For $L \in \mathcal{P}({}^\perp A)$, consider an exact sequence $0 \to K \to P \to L \to 0$ with $P \in \mathrm{add}(A)$. Since L and P are in ${}^\perp A$, $K \in {}^\perp A$. So $\mathrm{Ext}^1_A(L, K) = 0$, thus the exact sequence splits and $L \in \mathrm{add}(A)$. This shows $\mathcal{P}({}^\perp A) \subseteq \mathrm{add}(A) \subseteq \mathcal{P}({}^\perp A)$, and hence $\mathcal{P}({}^\perp A) = \mathrm{add}(A)$.

Now consider $X \in {}^\perp A$. Since ${}^\perp A$ is Frobenius, there is an exact sequence $0 \to X \to I \to C \to 0$ with $I \in \mathcal{I}({}^\perp A) = \mathcal{P}({}^\perp A) = \mathrm{add}(A)$ and $C \in {}^\perp A$. So X is torsionless. This shows that A is left weakly Gorenstein, according to Theorem 1.2. □

6. An example

Let k be a field and $q \in k \setminus \{0\}$. We consider a 6-dimensional local algebra $\Lambda = \Lambda(q)$. If k is infinite, then we show that there are infinitely many Gorenstein-projective Λ-modules of dimension 3. Let $o(q) = |q^{\mathbb{Z}}|$ be the multiplicative order of q. If $o(q)$ is infinite, we show that there is also a semi-Gorenstein-projective Λ-module of dimension 3 which is not Gorenstein-projective.

6.1. The algebra $\Lambda = \Lambda(q)$. The algebra Λ is generated by x, y, z, subject to the relations:

$$x^2,\ y^2,\ z^2,\ yz,\ xy + qyx,\ xz - zx,\ zy - zx.$$

The algebra Λ has a basis $1,\ x,\ y,\ z,\ yx$, and zx and may be visualized as follows:

$$\Lambda:\quad \begin{tikzcd} & 1 \arrow[dl, "x"'] \arrow[dr, "y"] \arrow[drr, "z"] & & \\ x \arrow[dr, "y"'] \arrow[drrr, "z"] & & y \arrow[dl, dashed, "x"] \arrow[dr, "z"] & z \arrow[d, "x"] \\ & yx & & zx \end{tikzcd}$$

Here, we use the following convention: The vertices are the elements of the basis, the arrows are labeled by x, y, z. A solid arrow $v \to v'$ labeled say by x means that $xv = v'$, a dashed arrow $v \dashrightarrow v'$ labeled by x means that xv is a nonzero multiple of v' (in our case, $xy = -qyx$). If v is a vertex and no arrow starting at v is labeled say by x, then $xv = 0$.

One diamond in the picture has been dotted in order to draw attention to the relation $xy + qyx$; this relation plays a decisive role when looking at ΩM for a given Λ-module M.

We study the following modules $M(\alpha)$ with $\alpha \in k$. The module $M(\alpha)$ has a basis v, v', v'', such that $xv = \alpha v',\ \ yv = v', zv = v''$, and such that v' and v'' are annihilated by x, y, z. That is,

$$M(\alpha):\quad \begin{tikzcd} v \arrow[d, dashed, shift right, "x"'] \arrow[d, shift left, "y"] \arrow[dr, "z"] & \\ v' & v'' \end{tikzcd}$$

The modules $M(\alpha)$ with $\alpha \in k$ are pairwise nonisomorphic indecomposable Λ-modules.

For $\alpha \in k$, we define $m_\alpha = x - \alpha y \in \Lambda$. In order to provide a proof of Theorem 1.5, we now collect some general results for the modules $M(\alpha)$, Λm_α, and the right ideals $m_\alpha\Lambda$ which are needed.

6.2. The module $M(q)$.

Lemma. *The intersection of the kernels of all the homomorphisms $M(q) \to {}_\Lambda\Lambda$ is $zM(q) = kv''$ and $M(q)/zM(q) \simeq \Lambda m_1$. In particular, $M(q)$ is not torsionless and $M(q)^* = (\Lambda m_1)^*$.*

Proof. Let $f : M(q) = \Lambda v \to {}_\Lambda\Lambda$ be a homomorphism. Let $f(v) = c_1x + c_2y + c_3z + c_4yx + c_5zx$ with $c_i \in k$. By $qf(v') = f(xv) = xf(v) = -c_2qyx + c_3zx$ and $f(v') = f(yv) = yf(v) = c_1yx$, we get $c_2 = -c_1$ and $c_3 = 0$. Thus, $f(v) = c_1(x - y) + c_4yx + c_5zx$. It follows that $f(v'') = f(zv) = zf(v) = 0$. This shows that v'' is contained in the kernel of any map $f : M(q) = \Lambda v \to {}_\Lambda\Lambda$. On the other hand, the homomorphism $g : M(q) = \Lambda v \to \Lambda$ given by $g(v) = x - y = m_1$ has kernel kv''. This completes the proof of the first assertion.

The map g provides a surjective map $p : M(q) \to \Lambda m_1$ and $p^* : M(q)^* \to (\Lambda m_1)^*$ is bijective, thus an isomorphism of right Λ-modules. □

6.3. The modules $M(\alpha)$ with $\alpha \in k$. We consider now the modules $M(\alpha)$ in general, and relate them to the left ideals Λm_α, and to the right ideals $m_\alpha\Lambda$. Let us denote by U_α the two-sided ideal generated by m_α, it is 3-dimensional with basis $m_\alpha,\ yx,\ zx$. Actually, for any $\alpha \in k$, the right ideal $m_\alpha\Lambda$ is equal to U_α (but we prefer to write U_α instead of $m_\alpha\Lambda$ when we consider it as a left module). For $\alpha \neq 1$, the left ideal Λm_α is equal to U_α.

If M is a module and $m \in M$, we denote by $r(m) : {}_\Lambda\Lambda \to M$ the right multiplication by m (defined by $r(m)(\lambda) = \lambda m$. Similarly, if N is a right Λ-module and $a \in N$, let $l(a) : \Lambda_\Lambda \to N$ be the left multiplication by a.

We denote by $u_\alpha : \Lambda m_\alpha \to \Lambda$ and $u'_\alpha : m_\alpha\Lambda \to \Lambda$ the canonical embeddings.

(1) *The right ideal $m_\alpha\Lambda$ is 3-dimensional (and equal to U_α), for all $\alpha \in k$.*

(2) *The left ideal Λm_α is 3-dimensional (and equal to U_α), for $\alpha \in k\setminus\{1\}$, whereas Λm_1 is 2-dimensional.*

(3) *We have $M(\alpha) \simeq \Lambda/U_\alpha$ for all $\alpha \in k$.*

Proof. The map $r(v) : \Lambda \to M(\alpha)$ is surjective (thus a projective cover) and

$$r(v)(m_\alpha) = m_\alpha v = (x - \alpha y)v = xv - \alpha yv = \alpha v' - \alpha v' = 0.$$

Thus, $\Lambda m_\alpha \subseteq \mathrm{Ker}(r(v))$. Also, $zx \in \mathrm{Ker}(r(v))$, thus $\mathrm{Ker}(r(v)) = U_\alpha$. This shows that $M(\alpha)$ is isomorphic to Λ/U_α. □

(4) *For $\alpha \in k \setminus \{1\}$, we have $M(q\alpha) \simeq \Lambda m_\alpha$.*

Proof. Consider the map $r(m_\alpha) : \Lambda \to \Lambda m_\alpha$. Since $r(m_\alpha)(m_{q\alpha}) = m_{q\alpha}m_\alpha = 0$, we see that $U_{q\alpha} \subseteq \mathrm{Ker}(r(m_\alpha))$. For $\alpha \neq 1$, the module Λm_α is 3-dimensional, therefore $r(m_\alpha)$ yields an isomorphism $\Lambda/U_{q\alpha} \to \Lambda m_\alpha$. Using (3) for $M(q\alpha)$, we see that $M(q\alpha) \simeq \Lambda/U_{q\alpha} \simeq \Lambda m_\alpha$. □

(5) *For any map $f : \Lambda m_\alpha \to \Lambda$, there is $\lambda \in \Lambda$ with $f = r(\lambda)u_\alpha$, for all $\alpha \in k$. Thus u_α is a left $\mathrm{add}(\Lambda)$-approximation.*

Proof. Let $f : \Lambda m_\alpha \to \Lambda$ be any map. Let $f(m_\alpha) = c_1x + c_2y + c_3z + c_4yx + c_5zx$ with $c_i \in k$. Since $f(ym_\alpha) = f(yx)$ and $yf(m_\alpha) = c_1yx$, we see that $f(yx) = c_1yx$. Since $f(xm_\alpha) = f(-\alpha xy) = q\alpha f(yx) = q\alpha c_1yx$ and $xf(m_\alpha) = c_2xy + c_3zx = -qc_2yx + c_3zx$, we see that $q\alpha c_1yx = -qc_2yx + c_3zx$, therefore $c_2 = -\alpha c_1$ and $c_3 = 0$. Thus, $f(m_\alpha) = c_1(x - \alpha y) + c_4yx + c_5zx$ belongs to $U_\alpha = m_\alpha\Lambda$, say $f(m_\alpha) = m_\alpha\lambda$ with $\lambda \in \Lambda$. Therefore $f(m_\alpha) = m_\alpha\lambda = r(\lambda)u_\alpha(m_\alpha)$, but this means that $f = r(\lambda)u_\alpha$. □

6.4. Lemma. *Let $\alpha \in k \setminus \{1\}$. Then there is an $\mho$-sequence*

$$0 \to M(q\alpha) \to \Lambda \to M(\alpha) \to 0.$$

Proof. According to (3), $M(\alpha) \simeq \Lambda/U_\alpha$. Since $\alpha \neq 1$, we have $U_\alpha = \Lambda m_\alpha$ by (2). Thus, we have the following exact sequence

$$0 \to \Lambda m_\alpha \xrightarrow{u_\alpha} \Lambda \to M(\alpha) \to 0$$

According to (5) the embedding $u_\alpha : \Lambda m_\alpha \to \Lambda$ is a left $\mathrm{add}(\Lambda)$-approximation. Thus, the sequence is an $\mho$-sequence. Finally, (4) shows that $\Lambda m_\alpha \simeq M(q\alpha)$. □

Corollary 1. *The module $M(0)$ is Gorenstein-projective with Ω-period equal to* 1. □

Corollary 2. *If $o(q) = \infty$, then the module $M(q)$ is semi-Gorenstein-projective.*

Proof. We assume that $o(q) = \infty$. Then $q^t \neq 1$ for all $t \geq 1$. By 6.4, all the sequences

$$0 \to M(q^{t+1}) \to \Lambda \to M(q^t) \to 0.$$

with $t \geq 1$ are $\mho$-sequences. They can be concatenated in order to obtain a minimal projective resolution of $M(q)$. This shows that $M(q)$ is semi-Gorenstein-projective. □

6.5. The right Λ-modules $m_\alpha\Lambda$ and $M(\alpha)^*$. We have started in 6.3 to present essential properties of the modules $M(\alpha)$. We look now also at the modules $m_\alpha\Lambda$ and $M(\alpha)^*$. We continue the enumeration of the assertions as started in 6.3.

(6) $\Omega(m_{q\alpha}\Lambda) = m_\alpha\Lambda$ *for all* $\alpha \in k$.

Proof. We consider the composition of the following right Λ-module maps

$$\Lambda_\Lambda \xrightarrow{l(m_\alpha)} \Lambda_\Lambda \xrightarrow{l(m_{q\alpha})} \Lambda_\Lambda$$

Since $m_{q\alpha}m_\alpha = 0$, the composition is zero. The image of $l(m_\alpha)$ is the right ideal $m_\alpha\Lambda$, the image of $l(m_{q\alpha})$ is the right ideal $m_{q\alpha}\Lambda$. Both right ideals are 3-dimensional, thus the sequence is exact. Thus $m_\alpha\Lambda = \mathrm{Ker}(p)$, for a surjective map $p : \Lambda_\Lambda \to m_{q\alpha}\Lambda$. Now p is a projective cover, thus $\mathrm{Ker}(p) = \Omega(m_{q\alpha}\Lambda)$, and therefore $\Omega(m_{q\alpha}\Lambda) \simeq m_\alpha\Lambda$. □

(7) $(\Lambda m_\alpha)^* = m_\alpha\Lambda$ *for all* $\alpha \in k$.

Proof. First, let us show that $(\Lambda m_\alpha)^*$ is 3-dimensional. On the one hand, besides u_α, there are homomorphisms $\Lambda m_\alpha \to \Lambda$ with image kyx and with image kzx, which shows that $(\Lambda m_\alpha)^*$ is at least 3-dimensional. According to (5), any homomorphism $\Lambda m_\alpha \to \Lambda$ maps into $\Lambda m_\alpha\Lambda = U_\alpha$. Since U_α is 3-dimensional, we have $\dim \mathrm{Hom}({}_\Lambda\Lambda, U_\alpha) = 3$, therefore $\dim(\Lambda m_\alpha)^* = \dim \mathrm{Hom}(\Lambda m_\alpha, \Lambda) = \dim \mathrm{Hom}(\Lambda m_\alpha, U_\alpha) \leq \dim \mathrm{Hom}({}_\Lambda\Lambda, U_\alpha) = 3$.

Second, using again (5), we see that $(\Lambda m_\alpha)^*$ is, as a right Λ-module, generated by u_α. Thus, there is a surjective homomorphism $\theta_\alpha : \Lambda_\Lambda \to (\Lambda m_\alpha)^*$ of right Λ-modules defined by $\theta_\alpha(1) = u_\alpha$. We have

$$(\theta_\alpha(m_{q^{-1}\alpha}))(m_\alpha) = (\theta_\alpha(1)m_{q^{-1}\alpha})(m_\alpha) = (u_\alpha m_{q^{-1}\alpha})(m_\alpha) = m_\alpha m_{q^{-1}\alpha} = 0,$$

therefore $\theta_\alpha(m_{q^{-1}\alpha}) = 0$. It follows that θ_α yields a surjective map $\Lambda_\Lambda/m_{q^{-1}\alpha}\Lambda \to (\Lambda m_\alpha)^*$. Actually, this map has to be an isomorphism, since $m_{q^{-1}\alpha}\Lambda$ is 3-dimensional. Therefore $\Lambda_\Lambda/m_{q^{-1}\alpha}\Lambda \simeq (\Lambda m_\alpha)^*$. By (6), $\Lambda_\Lambda/m_{q^{-1}\alpha}\Lambda \simeq m_\alpha\Lambda$. This completes the proof. □

(8) $M(q\alpha)^* = m_\alpha\Lambda$ *for all* $\alpha \in k$.

Proof. For $\alpha \neq 1$, we have $M(q\alpha) \simeq \Lambda m_\alpha$ by (4), thus we use (7). For $\alpha = 1$, we use 6.2 and then (7). □

Let us stress that (7) and (8) show that $M(q)^*$ and $(\Lambda m_1)^*$ are isomorphic, namely isomorphic to $m_1\Lambda$, whereas $M(q)$ and Λm_1 themselves are not isomorphic.

(9) *Let* $\alpha \in k \setminus \{1, q\}$. *For any homomorphism* $g : m_\alpha\Lambda \to \Lambda$ *there is* $\lambda \in \Lambda$ *with* $g = l(\lambda)u'_\alpha$. *Thus,* u'_α *is a left* $\mathrm{add}(\Lambda)$*-approximation.*

Proof. Let $g : m_\alpha\Lambda \to \Lambda_\Lambda$ be a homomorphism. We claim that $g(m_\alpha) \in \Lambda m_\alpha$. Let $g(m_\alpha) = c_1x + c_2y + c_3z + c_4yx + c_5zx$ with $c_i \in k$. Now, $g(m_\alpha x) = g(-\alpha yx) = -\alpha g(yx)$ and $g(m_\alpha)x = c_2xy + c_3zx$. Also, $g(m_\alpha y) = g(xy) = -qg(yx)$, and $g(m_\alpha)y = c_1xy + c_3zx = -c_1qyx + c_3zx$, thus $g(yx) = -q^{-1}g(m_\alpha y) = -q^{-1}(-c_1qyx + c_3zx) = c_1yx - q^{-1}c_3zx$. It follows that $c_2yx + c_3zx = -\alpha g(yx) = -\alpha(c_1yx - q^{-1}c_3zx) = -\alpha c_1yx + \alpha q^{-1}c_3zx$. Therefore $c_2 = -\alpha c_1$ and $c_3 = \alpha q^{-1}c_3$. Since we assume that $\alpha \neq q$, it follows that $c_3 = 0$. Therefore $g(m_\alpha) = c_1x - \alpha c_1y + c_3z + c_4yx + c_5zx = c_1(x - \alpha y) + c_4yx + c_5zx$ belongs to U_α. Since we also assume that $\alpha \neq 1$, we have $U_\alpha = \Lambda m_\alpha$. Thus $g(m_\alpha) \in \Lambda m_\alpha$.

As a consequence, there is $\lambda \in \Lambda$ with $g(m_\alpha) = \lambda m_\alpha$, therefore $g(m_\alpha) = \lambda m_\alpha = l(\lambda)u'_\alpha(m_\alpha)$. It follows that $g = l(\lambda)u'_\alpha$. □

6.6. Lemma. *Let $\alpha \in k \setminus \{1, q\}$. Then there is an $\mho$-sequence of right Λ-modules*

$$0 \to m_\alpha\Lambda \xrightarrow{u'_\alpha} \Lambda_\Lambda \to m_{q\alpha}\Lambda \to 0.$$

Proof. This is 6.5(6) and (9). □

6.7. *Proof of Theorem 1.6* According to 6.5(8), we have $M(q)^* = m_1\Lambda$. As we know from 6.2, $M(q)$ is not torsionless.

We assume now that $o(q) = \infty$. The Corollary 2 in 6.4 shows that $M(q)$ is semi-Gorenstein-projective. Since $q^{-t} \neq 1$ for all $t \geq 1$, the sequences

$$0 \to m_{q^{-t}}\Lambda \xrightarrow{u'_\alpha} \Lambda_\Lambda \to m_{q^{-t+1}}\Lambda \to 0$$

with $t \geq 1$ are $\mho$-sequences, by 6.6. They can be concatenated in order to obtain a minimal projective resolution of $m_1\Lambda$ and show that $m_1\Lambda$ is semi-Gorenstein-projective.

Finally, we want to show that $M(q)^{**} = \Omega M(1)$. According to 6.3(5), the map $u_1 : \Lambda m_1 \to \Lambda$ is a minimal left add(Λ)-approximation, thus we may consider as in 2.4(a) the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \Lambda m_1 & \xrightarrow{u_1} & \Lambda & \xrightarrow{\pi_1} & \Lambda/\Lambda m_1 & \longrightarrow & 0 \\
 & & \downarrow & & \| & & \downarrow{\scriptstyle\phi} & & \\
0 & \longrightarrow & (\Lambda m_1)^{**} & \longrightarrow & \Lambda & \xrightarrow{\pi_1^{**}} & (\Lambda/\Lambda m_1)^{**} & \longrightarrow & \mathrm{Ext}^1(M'(q)^*, \Lambda_\Lambda)
\end{array}$$

where $\phi = \phi_{\Lambda/\Lambda m_1}$. The submodule $zx(\Lambda/\Lambda m_1)$ belongs to the kernel of any map $\Lambda/\Lambda m_1 \to \Lambda$, and it is the kernel of the map $p : \Lambda/\Lambda m_1 \to M(1)$ defined by $p(\bar{1}) = v$. This shows that $zx(\Lambda/\Lambda m_1)$ is the kernel of ϕ, thus the image of ϕ is just $M(1)$. But the image of ϕ coincides with the image of π_1^{**}. In this way, we see that $(\Lambda m_1)^{**}$ is the kernel of a projective cover of $M(1)$, thus equal to $\Omega M(1)$.

Of course, $\Omega M(1)$ is decomposable, namely isomorphic to $\Lambda m_1 \oplus kzx$. □

6.8. Proof of Addendum 1.6. We denote by $q^{\mathbb{Z}}$ the set of elements of k which are of the form q^i with $i \in \mathbb{Z}$. Assume that $\alpha \in k \setminus q^{\mathbb{Z}}$, thus $q^t\alpha \neq 1$ for all $t \in \mathbb{Z}$. According to 6.4, all the sequences

$$0 \to M(q^{t+1}\alpha) \to \Lambda \to M(q^t\alpha) \to 0$$

with $t \in \mathbb{Z}$ are $\mho$-sequences. They can be concatenated in order to obtain a complete projective resolution for $M(\alpha)$, thus $M(\alpha)$ is Gorenstein-projective.

The following (well-known) lemma shows that there are infinitely many elements $\alpha \in k \setminus q^{\mathbb{Z}}$.

Lemma. *Assume that k is an infinite field and $q \in k$. Then $k \setminus q^{\mathbb{Z}}$ is an infinite set.*

We include a proof.

Proof. The assertion is clear if $o(q)$ is finite. Thus, let $o(q)$ be infinite (in particular, $q \neq 0$). Assume that the multiplicative group $k^* = k \setminus \{0\}$ is cyclic, say $k^* = w^{\mathbb{Z}}$. Then $o(w) = \infty$, and each element in k^* different from 1 has infinite multiplicative order. Since $(-1)^2 = 1$, we see that k is of characteristic 2. Now $w + 1 \neq 0$ shows that $w + 1 = w^n$ for some $n > 1$, thus w is algebraic over the prime field $\mathbb{Z}_2$. Thus $k = \mathbb{Z}_2(w)$ is a finite field, a contradiction. Since k^* is not cyclic, there is $a \in k^* \setminus q^{\mathbb{Z}}$. Then $a \cdot q^{\mathbb{Z}}$ is an infinite subset of $k^* \setminus q^{\mathbb{Z}}$. $\square$

7. Further details for $\Lambda = \Lambda(q)$

7.1. The $\mho$-components involving modules $M(\alpha)$**.** *The only $\mho$-sequences which involve a module of the form $M(\alpha)$ with $\alpha \in k$ are those exhibited in* 6.6.

Proof. We have to show that there is no $\mho$-sequence ending in $M(1)$ and no $\mho$-sequence starting in $M(q)$. Since $\Omega M(1)$ is decomposable, there is no $\mho$-sequence ending in $M(1)$. By 6.2, the module $M(q)$ is not torsionless, thus no $\mho$-sequence starts in $M(q)$. $\square$

We now want to determine the $\mho$-type of the modules $M(\alpha)$. According to Corollary 1 in 6.4, $M(0)$ is of $\mho$-type $\tilde{\mathbb{A}}_0$. Thus, we now assume that $\alpha \neq 0$.

7.2. Let us assume that $o(q) = \infty$ (for the case that $o(q)$ is finite, see 7.6). *There are three kinds of $\mho$-components which involve modules of the form $M(\alpha)$ with $\alpha \in k^*$. There is one component of the form $-\mathbb{N}$, it has $M(q)$ as its source, and there is one component of the form $\mathbb{N}$, it has $M(1)$ as its sink:*

$$\cdots \dashleftarrow M(q^4) \dashleftarrow M(q^3) \dashleftarrow M(q^2) \dashleftarrow M(q) \qquad M(1) \dashleftarrow M(q^{-1}) \dashleftarrow M(q^{-2}) \dashleftarrow \cdots$$

The remaining ones contain the modules $M(\alpha)$ with $\alpha \neq 0$ and $\alpha \notin q^{\mathbb{Z}}$; they are of the form $\mathbb{Z}$:

$$\cdots \dashleftarrow M(q^4\alpha) \dashleftarrow M(q^3\alpha) \dashleftarrow M(q^2\alpha) \dashleftarrow M(q\alpha) \dashleftarrow M(\alpha) \dashleftarrow M(q^{-1}\alpha) \dashleftarrow M(q^{-2}\alpha) \dashleftarrow \cdots$$

The positions of the reflexive modules are shaded.

According to Theorem 1.5, there are the following observations concerning the behavior of the modules $M(\alpha)$ with $\alpha \in k$.

- *The module $M(\alpha)$ is Gorenstein-projective if and only if $\alpha \notin q^{\mathbb{Z}}$.*
- *The module $M(\alpha)$ is not Gorenstein-projective, but semi-Gorenstein-projective if and only if $\alpha = q^t$ for some $t \geq 1$.*
- *The module $M(\alpha)$ is torsionless if and only if $\alpha \neq q$.*

- *The module $M(\alpha)$ is reflexive if and only if $\alpha \notin \{q, q^2\}$.*
- *The module $M(\alpha)$ is not Gorenstein-projective, but ∞-torsionfree if and only if $\alpha = q^t$ for some $t \le 0$.*

It seems worthwhile to know the canonical maps $\phi_X : X \to X^{**}$ for the nonreflexive modules $X = M(q)$ and $X = M(q^2)$. For $M(q)$ we refer to 6.7: there it is shown that $M(q)^{**} = \Omega M(1)$ *and that the image of* $\phi_{M(q)}$ *is* Λm_1.

It remains to look at $X = M(q^2)$. *The module $M(q^2)^{**}$ is the submodule $\Lambda m_q + \Lambda z$ of Λ and $\phi_{M(q^2)}$ is the inclusion map*

$$M(q^2) = \Lambda m_q \to \Lambda m_q + \Lambda z = M(q^2)^{**}.$$

Proof. Since $M(q^2)$ is torsionless, the map $\phi_{M(q^2)}$ is injective. There is the following commutative diagram with exact rows:

$$\begin{tikzcd}
0 \arrow[r] & M(q^2) \arrow[r, "u_q"] \arrow[d, "\phi_{M(q^2)}"] & \Lambda \arrow[r, "\pi_q"] \arrow[d, equal] & M(q) \arrow[r] \arrow[d, "\phi_{M(q)}"] & 0 \\
0 \arrow[r] & M(q^2)^{**} \arrow[r, "u_q^{**}"] & \Lambda \arrow[r, "\pi_q^{**}"] & M(q)^{**} \arrow[r] & \mathrm{Ext}^1(M(q^2)^*, \Lambda_\Lambda)\ .
\end{tikzcd}$$

As we know already, the image of $\phi_{M(q)}$ and therefore of π_q^{**}, is Λm_1. Thus the kernel of π_q^{**} is the submodule $\Lambda m_q + \Lambda z$ of Λ. Therefore $M(q^2)^{**} = \Lambda m_q + \Lambda z$ and $\phi_{M(q^2)}$ is the inclusion map $M(q^2) = \Lambda m_q \to \Lambda m_q + \Lambda z = M(q^2)^{**}$. □

7.3. The $\mho$-components involving right Λ-modules $m_\alpha\Lambda$. *The $\mho$-sequences which involve a right Λ-module of the form $m_\alpha\Lambda$ with $\alpha \in k$ are those exhibited in* 6.6 *as well as*

$$0 \to m_q\Lambda \xrightarrow{\begin{bmatrix} u_q \\ h \end{bmatrix}} \Lambda_\Lambda \oplus \Lambda_\Lambda \to \mho(m_q\Lambda) \to 0,$$

and, for $q \neq 1$,

$$0 \to m_1\Lambda \xrightarrow{\begin{bmatrix} u_1 \\ h' \end{bmatrix}} \Lambda_\Lambda \oplus \Lambda_\Lambda \to \mho(m_1\Lambda) \to 0.$$

Here, $h : m_q\Lambda \to \Lambda_\Lambda$ is defined by $h(m_q) = z$, whereas $h' : m_1\Lambda \to \Lambda_\Lambda$ is defined by $h'(m_1) = zx$.

Proof. It is easy to check that the map $\begin{bmatrix} u_q \\ h \end{bmatrix}$ and, for $q \neq 1$, the map $\begin{bmatrix} u_1 \\ h' \end{bmatrix}$ are minimal left add(Λ_Λ)-approximations. Clearly, the corresponding cokernels are not torsionless.

In addition, we have to show that there is no $\mho$-sequence ending in $m_{q^2}\Lambda$ or in $m_q\Lambda$. But this follows from the fact that the inclusion maps $u'_q : m_q\Lambda = \Omega(m_{q^2}\Lambda) \to P(m_{q^2}\Lambda)$ and $u'_1 : m_1\Lambda = \Omega(m_q\Lambda) \to P(m_q\Lambda)$ are not add(Λ_Λ)-approximations. □

Let $o(q) = \infty$ (the case that $o(q) < \infty$ will be discussed in 7.6). *There are five kinds of* $\mho$*-components involving right* Λ*-modules of the form* $m_\alpha\Lambda$ *with* $\alpha \in k$*, namely a component of the form* $\mathbb{N}$ *with* $m_{q^2}\Lambda$ *as a sink, a component of the form* $-\mathbb{N}$ *with* $\mho(m_1\Lambda)$ *as a source, and a component of the form* $\mathbb{A}_2$ *with sink* $m_q\Lambda$ *and source* $\mho(m_q\Lambda)$*:*

$$\begin{tikzcd}[column sep=small]
\cdots \arrow[r, dashed] & m_{q^3}\Lambda \arrow[r, dashed] & m_{q^2}\Lambda & & m_q\Lambda & & m_1\Lambda \arrow[r, dashed] & m_{q^{-1}}\Lambda \arrow[r, dashed] & m_{q^{-2}}\Lambda \arrow[r, dashed] & m_{q^{-3}}\Lambda \arrow[r, dashed] & \cdots \\
& & & \mho(m_q\Lambda) \arrow[ur, dashed] & & \mho(m_1\Lambda) \arrow[ur, dashed] & & & & &
\end{tikzcd}$$

The $\mho$*-components containing right* Λ*-modules* $m_\alpha\Lambda$ *with* $\alpha \in k \setminus q^{\mathbb{Z}}$ *are of the form* $\mathbb{Z}$*:*

$$\begin{tikzcd}[column sep=small]
\cdots \arrow[r, dashed] & m_{q^3\alpha}\Lambda \arrow[r, dashed] & m_{q^2\alpha}\Lambda \arrow[r, dashed] & m_{q\alpha}\Lambda \arrow[r, dashed] & m_\alpha\Lambda \arrow[r, dashed] & m_{q^{-1}\alpha}\Lambda \arrow[r, dashed] & m_{q^{-2}\alpha}\Lambda \arrow[r, dashed] & m_{q^{-3}\alpha}\Lambda \arrow[r, dashed] & \cdots
\end{tikzcd}$$

In addition, there is the $\mho$*-component consisting of the single right* Λ*-modules* $m_0\Lambda$*, it is of the form* $\tilde{\mathbb{A}}_0$.

For the convenience of the reader, the pictures in 7.1 and 7.2 have been arranged so that the A-duality is respected. Thus, in 7.1, the arrows are drawn from right to left, in 7.2 from left to right. Also we recall from 6.3(8) that the A-dual of $M(q\alpha)$ is $m_\alpha\Lambda$, therefore the position of $m_\alpha\Lambda$ in the pictures 7.2 is the same as the position of $M(q\alpha)$ in 7.1.

7.4. We complete the description of the behavior of the modules $M(\alpha)$ started in 7.2.

- *The module* $M(\alpha)$ *is not Gorenstein-projective, but* $M(\alpha)^*$ *is semi-Gorenstein-projective, if and only if* $\alpha = q^t$ *for some* $t \le 1$.
- *The module* $M(\alpha)$ *is not Gorenstein-projective, but* $M(\alpha)^*$ *is* ∞*-torsionfree, if and only if* $\alpha = q^t$ *for some* $t \ge 3$.

Proof. According to 7.2, the module $M(\alpha)$ is Gorenstein-projective if and only if $\alpha \notin q^{\mathbb{Z}}$. Thus, we can assume that $\alpha = q^t$ for some $t \in \mathbb{Z}$. According to 6.3(8), the module $M(q^t)$ is isomorphic to $m_{q^{t-1}}\Lambda$. The display of the $\mho$-components shows that $m_{q^{t-1}}\Lambda$ is semi-Gorenstein-projective if and only if $t-1 \le 0$, thus if and only if $t \le 1$, see Theorem 1.5. Similarly, we see that $m_{q^{t-1}}\Lambda$ is ∞-torsionfree if and only if $t-1 \ge 2$, thus if and only if $t \ge 3$. $\square$

7.5. We have mentioned in 1.7 that one may use the algebra $\Lambda = \Lambda(q)$ with $o(q) = \infty$ in order to exhibit examples of modules M which satisfy precisely two of the three properties (G1), (G2) and (G3):

(1) $M = M(q)$ *satisfies* (G1), (G2), *but not* (G3).

(2) $M = M(q^3)$ *satisfies* (G1), (G3), *but not* (G2).

(3) $M = M(1)$ *satisfies* (G2), (G3), *but not* (G1).

Proof. For (1): this is the main assertion of Theorem 1.5. For (2): see 7.2 and 7.3. For (3): according to 7.2, $M(1)$ is reflexive, but not Gorenstein-projective. According to 6.3(8), we have $M(1)^* = m_{q^{-1}}\Lambda$, and $m_{q^{-1}}\Lambda$ is semi-Gorenstein-projective, see 7.3. $\square$

Let us look for similar examples for Λ^{op}, thus, for right Λ-modules N.

(1*) *There is* **no** *right* Λ*-module of the form* $N = m_\alpha\Lambda$ *satisfying* (G1)*,* (G2)*, but not* (G3)*.*

(2*) *The right* Λ*-module* $N = m_{q^{-2}}\Lambda$ *satisfies* (G1)*,* (G3)*, but not* (G2)*.*

(3*) *The right* Λ*-module* $N = m_{q^2}\Lambda$ *satisfies* (G2)*,* (G3)*, but not* (G1)*.*

Proof. (2*) There starts an infinite $\mho$-path at $N = m_{q^{-2}}\Lambda$, thus N satisfies (G1). There ends an $\mho$-path of length 2 at N, thus N satisfies (G3). Of course, N^* cannot be semi-Gorenstein-projective, since otherwise N would be Gorenstein-projective.

(3*) Let $N = m_{q^2}\Lambda$. According to 6.5(8), $N = M(q^3)^*$. As we know from 7.1, $M(q^3)$ is reflexive, thus N is reflexive and $N^* = M(q^3)^{**} = M(q^3)$ is semi-Gorenstein-projective.

(1*) Assume that $N = m_\alpha\Lambda$ and N^* are both semi-Gorenstein-projective. Since N cannot be Gorenstein-projective, it is not reflexive. Thus $\alpha \in \{1, q\}$. Since $[m_q\Lambda]$ is the sink of an $\mho$-component, $m_\alpha\Lambda$ is not semi-Gorenstein-projective. Thus $\alpha = 1$. But $(m_1\Lambda)^* = M(q)^{**} = \Omega M(1)$, according to 6.5(8) and Theorem 1.5. As we have mentioned already in the proof 6.7, $\Omega M(1) \simeq \Lambda m_1 \oplus k$, where k is the simple Λ-module. We claim that k is not semi-Gorenstein-projective, thus $\Omega M(1)$ is not semi-Gorenstein-projective.

Lemma. *Let A be a local artin algebra which is not self-injective, and S its simple A-module. Then* $\mathrm{Ext}^i(S, {}_AA) \neq 0$ *for all $i \geq 1$.*

Proof. Let $0 \to {}_AA \to I_0 \to I_1 \to \cdots$ be a minimal injective coresolution. Since ${}_AA$ is not injective, all the modules I_i are nonzero. We have $\mathrm{Ext}^i(S, {}_AA) \cong \mathrm{Hom}(S, I_i)$. □

7.6. Let us look also at the case when $o(q) = n < \infty$.

Left modules $M(\alpha)$ with $\alpha \in k^*$. *There are two kinds of $\mho$-components which involve modules of the form $M(\alpha)$ with $\alpha \in k^*$. There is one $\mho$-component of the form $\mathbb{A}_n$, it has $M(q)$ as its source, and $M(1)$ as its sink:*

$$M(1) \leftarrow M(q^{n-1}) \leftarrow \cdots \leftarrow M(q^3) \leftarrow M(q^2) \leftarrow M(q)$$

The remaining ones (containing the modules $M(\alpha)$ with $\alpha \in k^* \setminus q^{\mathbb{Z}}$) are directed cycles of cardinality n :

$$\begin{tikzcd}[column sep=small] M(\alpha) \arrow[rrrrrr, dashed, bend right=15] & M(q^{n-1}\alpha) \arrow[l, dashed] & \cdots \arrow[l, dashed] & M(q^3\alpha) \arrow[l, dashed] & M(q^2\alpha) \arrow[l, dashed] & & M(q\alpha) \arrow[ll, dashed] \end{tikzcd}$$

All modules in the cycles are reflexive. In the $\mho$-component of form $\mathbb{A}_n$, the modules $M(q)$ and $M(q^2)$ are not reflexive (they coincide for $o(q) = 1$); for $o(q) \geq 3$, there are $n - 2$ additional modules $M(1) = M(q^n)$, $M(q^{n-2})$, $\cdots$, $M(q^4)$, $M(q^3)$ in the $\mho$-component, and these modules are reflexive.

Right modules $m_\alpha\Lambda$ with $\alpha \in k^*$. *There is always the $\mho$-component of form $\mathbb{A}_2$ with $\Omega(m_q\Lambda)$ as its source and $m_q\Lambda$ as its sink. In addition, for $n \geq 2$, there is an $\mho$-component of form $\mathbb{A}_n$ containing the modules $m_{q^i}\Lambda$ with $2 \leq i \leq n$ as well as $\Omega(m_1\Lambda)$; it has $\Omega(m_1\Lambda)$ as its source, and $m_{q^2}\Lambda$ as its sink*:

$$\begin{tikzcd}[column sep=small] m_{q^{-1}}\Lambda \arrow[r, dashed] & m_{q^{-2}}\Lambda \arrow[r, dashed] & \cdots \arrow[r, dashed] & m_{q^2}\Lambda & & m_q\Lambda & & m_1\Lambda \arrow[lllllll, dashed, bend right=12] \\ & & & & \mho(m_q\Lambda) \arrow[ur, dashed] & & \mho(m_1\Lambda) \arrow[ur, dashed] & \end{tikzcd}$$

The remaining $\mho$*-components (containing the right modules* $m_\alpha\Lambda$ *with* $\alpha \in k^* \setminus q^{\mathbb{Z}}$*) are directed cycles of cardinality* n*:*

$$\begin{tikzcd} m_{q^{-1}\alpha}\Lambda \arrow[r, dashed] & m_{q^{-2}\alpha}\Lambda \arrow[r, dashed] & \cdots \arrow[r, dashed] & m_{q^2\alpha}\Lambda \arrow[r, dashed] & m_{q\alpha}\Lambda \arrow[r, dashed] & m_\alpha\Lambda \arrow[lllll, dashed, bend left] \end{tikzcd}$$

Again, the modules in the cycles are reflexive. In the $\mho$-components of form $\mathbb{A}_n$ and $\mathbb{A}_2$, the modules $m_1\Lambda$ and $\mho(m_1\Lambda)$, as well as $m_q\Lambda$ and $\mho(m_q\Lambda)$ are not reflexive; whereas (for $o(q) \geq 3$) the modules $m_{q^i}\Lambda$ with $2 \leq i \leq n-1$ are reflexive.

Proof. First, let us look at left modules. According to 7.1, the $\mho$-sequences presented here are the only ones involving modules of the form $M(\alpha)$. Thus, $[M(q)]$ is a source in the $\mho$-quiver and $[M(1)]$ is a sink. This holds true also for $o(q) = 1$: here $q = 1$ and $[M(1)]$ is both a sink and a source, thus a singleton $\mho$-component (without any arrow). Finally, for any n, the elements $1, q, \ldots, q^{n-1}$ are pairwise different, as are the elements $\alpha, q\alpha, \ldots, q^{n-1}\alpha$ for $\alpha \in k \setminus q^{\mathbb{Z}}$.

For dealing with the right modules, we refer to 7.3. □

7.7. We have shown in 1.5 that any $\mho$-component is a linearly oriented quiver of type $\mathbb{A}_n$ (with $n \geq 1$ vertices), a directed cycle $\tilde{\mathbb{A}}_n$ (with $n+1 \geq 1$ vertices), or of the form $-\mathbb{N}$, or $\mathbb{N}$, or $\mathbb{Z}$. Conversely, 7.2 and 7.6 show that all these cases arise for algebras of the form $\Lambda(q)$.

7.8. A forthcoming paper [RZ1] will be devoted to a detailed study of all the 3-dimensional local Λ-modules for the algebra $\Lambda = \Lambda(q)$. If q has infinite multiplicative order, we will encounter a whole family of 3-dimensional local modules which are semi-Gorenstein-projective, but not torsionless. A local artin algebra A with radical J is said to be *short* if $J^3 = 0$. In particular, the algebras $\Lambda(q)$ are short local algebras. It is shown in [RZ2] that if A is a short local algebra with a module M which is semi-Gorenstein-projective, but not Gorenstein-projective, then $|J^2| = |J/J^2| - 1 \geq 2$. This paper [RZ2], as well as [RZ3], are devoted to the syzygy modules of modules over short local algebras.

8. Remarks

The first remarks draw the attention to the papers [JS] and [CH]. In 8.1, we show that the $\Lambda(q)$-modules $M(q^{-s})$ with $s \geq 0$ and $o(q) = \infty$ satisfy some further conditions which were discussed by Jorgensen and Şega. In 8.2 we show that the algebra $\Lambda(q)$ for $o(q) = \infty$ does not satisfy the so-called Auslander condition of Christensen and Holm.

In 8.3, we show that essential features of $\Lambda(q)$ are related to corresponding ones of its subalgebra $\Lambda'(q)$, which is the quantum exterior algebra. 8.4 presents a two-fold covering of $\Lambda(q)$ which has properties similar to $\Lambda(q)$, but provides for $o(q) = \infty$ examples of semi-Gorenstein-projective modules M which are not Gorenstein-projective, with the additional property that $\mathrm{End}(M) = k$.

8.1. The conditions (TR$_i$) **of Jorgensen and Şega.** As we have mentioned, Jorgensen and Şega have shown in [JS] that there exist semi-Gorenstein-projective modules which are not Gorenstein-projective. Actually, the main result of [JS] is a stronger assertion.

Following [JS], we say that an R-module M satisfies the condition (TR_i) for some $i \geq 1$ provided $\mathrm{Ext}^i(M, R) = 0$, and that M satisfies the condition (TR_i) for some $i \leq -1$ provided $\mathrm{Ext}^{-i}(\mathrm{Tr}\, M, R_R) = 0$. Note that ($\mathrm{TR}_i$) is defined only for $i \neq 0$. Thus, M is semi-Gorenstein-projective if and only if M satisfies (TR_i) for all $i \geq 1$, and M is ∞-torsionfree (i.e., $\mathrm{Tr}\, M$ is semi-Gorenstein-projective) if and only if M satisfies (TR_i) for all $i \leq -1$. Note that M satisfies (TR_i) if and only if $\mathrm{Tr}(M)$ satisfies (TR_{-i}). The main theorem of Jorgensen and Şega asserts that *there exists a local artinian ring R and a family M_s of R-modules, with $M_s = \Omega M_{s+1}$ for $s \geq 1$, such that M_s satisfies* (TR_i) *if and only if $i < s$.*

Such a module M_s satisfies the conditions (G2) and (G3), and satisfies in addition the condition that $\mathrm{Ext}^i(M_s, R) = 0$ if and only if $1 \leq i \leq s-1$. Of course, this is a condition which is much stronger than the negation of (G1).

Let us show that our algebra $\Lambda(q)$ with $o(q) = \infty$ also provides such examples. Of course, in contrast to the algebra R exhibited by Jorgensen and Şega, $\Lambda(q)$ is noncommutative. There is the following general result:

Proposition. *Let R be a local artinian algebra which is not self-injective, with simple R-module S.*

If M is an indecomposable ∞-torsionfree module such that S is a proper direct summand of ΩM, then M satisfies (TR_i) *if and only if $i < 0$.*

If M is an indecomposable module such that M satisfies (TR_i) *if and only if $i < 0$, then for every $s \geq 1$, the module $\mho^{s-1}M$ satisfies* (TR_i) *if and only if $i < s$.*

Proof. First, let M be indecomposable, ∞-torsionfree, with $\Omega M \cong S \oplus X$ for some nonzero module X. Since M is ∞-torsionfree, M satisfies (TR_i) for $i \leq -1$. Since ΩM is decomposable, we have $\mathrm{Ext}^1(M, R) \neq 0$, i.e., M does not satisfy (TR_1). By Lemma 7.5, $\mathrm{Ext}^i(S, R) \neq 0$ for all $i \geq 1$. Thus, for $i \geq 2$ we have $\mathrm{Ext}^i(M, R) \cong \mathrm{Ext}^{i-1}(\Omega M, R) \cong \mathrm{Ext}^{i-1}(S, R) \oplus \mathrm{Ext}^{i-1}(X, R) \neq 0$, which means that M does not satisfy (TR_i).

Next, assume that M is an indecomposable module such that M satisfies (TR_i) if and only if $i \leq -1$. For $s \geq 1$ consider the module $M_s = \mho^{s-1}M$. For $i \leq -1$, M_s satisfies (TR_i): in fact, by Lemma 4.4, $\mathrm{Ext}^{-i}(\mathrm{Tr}(M_s), R) = \mathrm{Ext}^{-i}(\mathrm{Tr}(\mho^{s-1}M), R) \cong \mathrm{Ext}^{-i}(\mathrm{Tr}(\mathrm{Tr}\, \Omega^{s-1}\, \mathrm{Tr}(M)), R) \cong \mathrm{Ext}^{-i}(\Omega^{s-1}\, \mathrm{Tr}(M), R) \cong \mathrm{Ext}^{-i+s-1}(\mathrm{Tr}(M), R) = 0$.

If $1 \leq i \leq s-1$, then $s-i \geq 1$ and $\mathrm{Ext}^i(M_s, R) = \mathrm{Ext}^i(\mho^{s-1}M, R) \cong \mathrm{Ext}^1(\mho^{s-i}M, R) = 0$, since $s-i-1 \geq 0$ and $\mho^{s-i-1}M$ is torsionless.

If $i \geq s$, then $i-s+1 \geq 1$ shows that $\mathrm{Ext}^i(M_s, R) \simeq \mathrm{Ext}^{i-s+1}(M, R) \neq 0$, i.e., M_s does not satisfy (TR_i). □

Application: Let $R = \Lambda = \Lambda(q)$ with $o(q) = \infty$. Then $M = M(1)$ is an indecomposable ∞-torsionfree module and S is a proper direct summand of ΩM, thus the Proposition above shows that for $s \geq 1$, $M_s = \mho^{s-1}M = M(q^{-(s-1)})$ satisfies (TR_i) if and only if $i < s$. □

8.2. The Auslander condition of Christensen and Holm. Christensen and Holm [CH] say that a left-noetherian ring A satisfies the *Auslander condition,* provided that for every finitely generated left A-module M, there is an integer $b(M)$ with the following property: if M' is a finitely generated left

A-module, then the vanishing $\mathrm{Ext}^{\gg 0}(M, M') = 0$ implies that $\mathrm{Ext}^{>b(M)}(M, M') = 0$. We are indebted to Christensen and Holm for having drawn our attention to Theorem C of [CH] which asserts: *If A is a finite-dimensional k-algebra A satisfying the Auslander condition, then A is left weakly Gorenstein* (here, we have taken into account that a finite-dimensional k-algebra has a dualizing complex, see 3.4 in [CH]). This shows that *the algebra $\Lambda(q)$ with $o(q) = \infty$ does not satisfy the Auslander condition.* Actually, this can be seen directly, using the following easy observation.

Proposition. *Assume that A is a finite-dimensional k-algebra which satisfies the Auslander condition. Let N_i with $i \in \mathbb{Z}$ be finite-dimensional right A-modules with $\Omega N_i = N_{i-1}$ for all i. If at least one of the modules N_i is semi-Gorenstein-projective, then all the modules N_i are semi-Gorenstein-projective, thus Gorenstein-projective.*

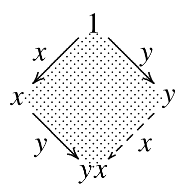
Proof. Note that A satisfies the Auslander condition if and only if for every finite-dimensional right A-module N, there is an integer $c(N)$ such that for every finite-dimensional right A-module N', the vanishing $\mathrm{Ext}^{\gg 0}(N', N) = 0$ implies that $\mathrm{Ext}^{>c(N)}(N', N) = 0$ (here, $c(N) = b(DN)$, where $D = \mathrm{Hom}(, -, k)$ denotes the k-duality).

We assume that N_0 is semi-Gorenstein-projective, whereas N_1 is not semi-Gorenstein-projective. Then we must have $\mathrm{Ext}^1(N_1, A_A) \neq 0$. Since N_0 is semi-Gorenstein-projective, $\mathrm{Ext}^t(N_0, A_A) = 0$ for all $t \geq 1$ and therefore $\mathrm{Ext}^{t+j}(N_j, A_A) = 0$ for all $t \geq 1$ and $j \geq 0$. In particular, we have $\mathrm{Ext}^{\gg 0}(N_j, A_A) = 0$ for all $j \geq 0$. Now we use the Auslander condition with $c = c(A_A)$. Since $\mathrm{Ext}^{\gg 0}(N_{c+1}, A_A) = 0$, we have $\mathrm{Ext}^{c+1}(N_{c+1}, A_A) = 0$. On the other hand, $\mathrm{Ext}^{c+1}(N_{c+1}, A_A) \simeq \mathrm{Ext}^1(N_1, A_A) \neq 0$. This is a contradiction. □

For our algebra $\Lambda(q)$ with $o(q) = \infty$, let $N_i = m_{q^i}\Lambda$ with $i \in \mathbb{Z}$. According to 6.5(6), we have $\Omega N_i = N_{i-1}$. As we know, the right module $N_0 = m_1\Lambda = M(q)^*$ is semi-Gorenstein-projective, but not Gorenstein-projective, see Theorem 1.6. This shows that $\Lambda(q)$ with $o(q) = \infty$ does not satisfy the Auslander condition.

8.3. The quantum exterior algebra $\Lambda' = \Lambda'(q)$ in two variables (see, for example [S]). Let Λ' be the k-algebra generated by x, y with the relations $x^2,\ y^2,\ xy + qyx$. It has a basis $1,\ x,\ y$, and yx. We may use the following picture as an illustration:

Λ' :



If we factor out the socle of Λ', we obtain the 3-dimensional local algebra Λ'' with radical square zero (it is generated by x, y with relations x^2, y^2, xy, yx).

Note that $\Lambda'(q)$ is a subalgebra of $\Lambda(q)$, and that $\Lambda z \Lambda = \Lambda z = \mathrm{span}\{z, zx\}$. The composition $\Lambda' \hookrightarrow \Lambda \twoheadrightarrow \Lambda/\Lambda z \Lambda$ of the canonical maps is an isomorphism of algebras. In this way, the Λ'-modules may be considered as the Λ-modules which are annihilated by z. We should stress that the elements $m_\alpha = x - \alpha y$ (which play a decisive role in our investigation) belong to Λ'.

For $\alpha \in k$, let $M'(\alpha)$ be the Λ'-module with basis v, v', such that $xv = \alpha v'$, $yv = v'$, and $xv' = 0 = yv'$. In addition, we define $M'(\infty)$ as the Λ'-module with basis v, v', such that $xv = v'$, $yv = xv' = yv' = 0$. Here are the corresponding illustrations:

$$M'(\alpha):\quad \begin{tikzcd} v \arrow[d, shift right, dashed, "x"'] \arrow[d, shift left, "y"] \\ v' \end{tikzcd} \qquad\qquad M'(\infty):\quad \begin{tikzcd} v \arrow[d, "x"'] \\ v' \end{tikzcd}$$

The modules $M'(\alpha)$ with $\alpha \in k \cup \{\infty\}$ are pairwise nonisomorphic and indecomposable, and any two-dimensional indecomposable Λ'-module is of this form. In particular, the left ideal $\Lambda' m_\alpha$ is isomorphic to $M'(q\alpha)$, for any $\alpha \in k \cup \{\infty\}$. The essential property of the modules $M'(\alpha)$ is the following: $\Omega_{\Lambda'} M'(\alpha) = M'(q\alpha)$. This follows from the fact that $m_{q\alpha} m_\alpha = 0$ and it is this equality which has been used frequently in sections 6 and 7.

For all $\alpha \in k$, $M(\alpha)$ considered as a Λ'-module, is equal to $M'(\alpha) \oplus k$, where k is the simple Λ'-module. Also, we should stress that rad Λ considered as a left Λ'-module is the direct sum of I and $M'(\infty)$, where I is the indecomposable injective Λ''-module.

8.4. A variation. Let $\tilde{\Lambda}$ be the algebra defined by a quiver with two vertices, say labeled by 1 and 2, with three arrows $1 \to 2$ labeled by x, y, z and with three arrows $2 \to 1$, also labeled by x, y, z, satisfying the "same" relations as Λ (of course, now we have 14 relations: seven concerning paths $1 \to 2 \to 1$ and seven concerning paths $2 \to 1 \to 2$). Whereas Λ is a local algebra, the algebra $\tilde{\Lambda}$ is a connected algebra with two simple modules $S(1)$ and $S(2)$.

For all the considerations in sections 6 and 7, there are corresponding ones for $\tilde{\Lambda}$, but always we have to take into account that now we deal with two simple modules $S(1)$ and $S(2)$: Corresponding to the module $M(\alpha)$, there are two different modules, namely $M^1(\alpha)$ with top $S(1)$ and $M^2(\alpha)$ with top $S(2)$. The modules $M^1(\alpha)$ and $M^2(\alpha)$ have similar properties as $M(\alpha)$, in particular, $M^1(q)$ and $M^2(q)$ are semi-Gorenstein-projective and not Gorenstein-projective provided that $o(q) = \infty$. There is one decisive difference between the Λ-modules and the $\tilde{\Lambda}$-modules: The endomorphism ring of $M^1(\alpha)$ and $M^2(\alpha)$ is equal to k, whereas the endomorphism ring of any $M(\alpha)$ is 3-dimensional.

9. Questions

9.1. We have constructed a module which satisfies the conditions (G1), (G2), but not (G3). As we have mentioned already in the introduction, it is an open problem whether such a module does exist in case we deal with commutative rings.

9.2. One may ask whether or not the finiteness of gp A implies that A is left weakly Gorenstein, There is a weaker question: is A left weakly Gorenstein, in case all the Gorenstein-projective A-modules are projective?

9.3. Following Marczinzik [M1, question 1], one may ask whether a left weakly Gorenstein artin algebra is also right weakly Gorenstein, thus whether the existence of an $\mho$-component of the form $\mathbb{N}$ implies that also an $\mho$-component of the form $-\mathbb{N}$ exists.

Note that if any right weakly Gorenstein algebra is left weakly Gorenstein, then the Gorenstein symmetry conjecture holds true. Namely, we claim: *If* $\operatorname{inj.dim.}{}_AA \le d$ *and* $\operatorname{inj.dim.} A_A > d$ (the Gorenstein symmetry conjecture asserts that this should not happen), *then A is right weakly Gorenstein, but not left weakly Gorenstein.*

Proof. Let Q be an injective cogenerator of $\operatorname{mod} A$. We assume that $\operatorname{inj.dim.}{}_AA$ is at most d. As we have seen in 4.9, A is right weakly Gorenstein and any module of the form $\Omega^d M$ is semi-Gorenstein projective. Now assume that A is also left weakly Gorenstein. Then all the modules $\Omega^d M$ are Gorenstein-projective. In particular, $Q' = \Omega^d Q$ is Gorenstein-projective. A well-known argument shows that if Q' is Gorenstein-projective, then Q' is even projective. [Namely, assume that Q' is Gorenstein-projective. Then there is a Gorenstein-projective module Q'' such that $Q' = P' \oplus \Omega^{d+1}Q''$ with P' projective. Now $\operatorname{Ext}^1(\Omega^d Q'', Q') \simeq \operatorname{Ext}^{d+1}(Q'', Q') \simeq \operatorname{Ext}^1(Q'', Q) = 0$, here the first isomorphism is the usual index shift, whereas the second comes from the fact that Q'' is (semi-)Gorenstein-projective and $Q' = \Omega^d Q$ (for a semi-Gorenstein-projective module N, and any module Z, we have $\operatorname{Ext}^{i+1}(N, \Omega Z) \simeq \operatorname{Ext}^i(N, Z)$ for all $i \ge 1$). But $\operatorname{Ext}^1(\Omega^d Q'', P' \oplus \Omega^{d+1}Q'') = 0$ implies that $\operatorname{Ext}^1(\Omega^d Q'', \Omega^{d+1}Q'') = 0$, thus the canonical exact sequence $0 \to \Omega^{d+1}Q'' \to P(\Omega^d Q'') \to \Omega^d Q'' \to 0$ splits and $\Omega^{d+1}Q''$ has to be projective (even zero). It follows that $Q' = P' \oplus \Omega^{d+1}Q''$ is projective.] Since Q' is projective, the projective dimension of Q is at most d. Using duality, we see that $\operatorname{inj.dim.} A_A \le d$. $\square$

9.4. Assume that there exists a nonreflexive A-module M such that both M and M^* are semi-Gorenstein-projective. Is then the same true for A^{op}? Even for $A = \Lambda(q)$ with $o(q) = \infty$, we do not know the answer. According to 7.5(1*), a right A-module N of the form $N = m_\alpha \Lambda(q)$ is reflexive, if both N and N^* are semi-Gorenstein-projective. But, there could exist some other right A-module N satisfying (G1), (G2) and not (G3).

9.5. The Nunke condition. Does there exist a semi-Gorenstein-projective module $M \neq 0$ with $M^* = 0$? Such a module would be an extreme example of a module satisfying (G1), (G2) and not (G3). Marczinzik has pointed out that this question concerns the Nunke condition [H] for M, which asserts that $\operatorname{Ext}^i(M, A) \neq 0$ for some $i \ge 0$, see [J]. Colby and Fuller [CF] have conjectured that the Nunke condition should hold for any module M; they called this the *strong Nakayama conjecture*. The strong Nakayama conjecture obviously implies the generalized Nakayama conjecture which asserts that *for any simple module S there should exist some $i \ge 0$ such that* $\operatorname{Ext}^i(S, A) \neq 0$. It is known that the Nunke condition is satisfied in case the finitistic dimension conjecture holds true.

Note that if M is indecomposable and semi-Gorenstein-projective, then M^* may be decomposable, as Theorem 1.5 shows: the $\Lambda(q)^{\mathrm{op}}$-module $M(q)^*$ is indecomposable and semi-Gorenstein-projective, but $M(q)^{**}$ is decomposable.

9.6. The conditions (TR$_i$)**.** Following Jorgensen and Şega [JS], one may ask whether an A-module which satisfies (TR$_i$) for all but finitely many values of i, has to be Gorenstein-projective. In general, this is not the case, since there is the following proposition.

Proposition. *If both M and M^* are semi-Gorenstein-projective, then M satisfies the conditions* (TR$_i$) *for all $i \notin \{-1, -2\}$.*

Proof. Let M be semi-Gorenstein-projective. Then M satisfies (TR$_i$) for $i \geq 1$. Since $\mathrm{Ext}^1(M, A) = 0$ for $i = 1, 2$, Lemma 2.5 asserts that there is a projective module Y such that $M^* \simeq \Omega^2 \operatorname{Tr} M \oplus Y$. Assume now that also M^* is semi-Gorenstein-projective. Then for $i \geq 1$, we have $\mathrm{Ext}^{i+2}(\operatorname{Tr} M, A_A) = \mathrm{Ext}^i(\Omega^2 \operatorname{Tr} M, A_A) = \mathrm{Ext}^i(M^*, A_A) = 0$, thus M satisfies also (TR$_i$) for $i \leq -3$. $\square$

Thus, our paper shows that there are (noncommutative) artinian rings with modules M which satisfy (TR$_i$) for all $i \notin \{-1, -2\}$ and which are not Gorenstein-projective. For commutative artinian rings (and this was the setting considered by Jorgensen and Şega) the question is open.

Acknowledgment

We thank Alex Martsinkovsky for providing copies of [Br] and [A1]. We are indebted to Lars Christensen, Henrik Holm, Zhaoyong Huang, Rene Marczinzik, Deja Wu for helpful comments. We also are grateful to two referees for carefully reading the manuscript and making valuable suggestions.

References

[A1] M. Auslander, Anneaux de Gorenstein et torsion en algébre commutative, Sèminaire d'algébre commutative dirigé par P. Samuel (1966-1967), tome 1, notes by: M. Mangeney, C. Peskine, L. Szpiro, Secrétariat Mathématique, Paris, 2–69 (1967).

[A2] M. Auslander, Coherent functors. In: Proc. the Conf. on Categorical Algebra. La Jolla 1965. Springer, 189–231.

[AB] M. Auslander, M. Bridger, Stable module theory, Mem. Amer. Math. Soc. 94., Amer. Math. Soc., Providence, R.I., 1969.

[AR1] M. Auslander, I. Reiten, Applications of contravariantly finite subcategories, Adv. Math. 86 (1991), 111–152.

[AR2] M. Auslander, I. Reiten, Syzygy modules for Noetherian rings, J. Algebra 183 (1996), 167–185.

[ARS] M. Auslander, I. Reiten, S. O. Smalø, Representation theory of Artin algebras, Cambridge Studies in Advanced Math. 36. Cambridge University Press, 1995.

[AM] L. L. Avramov, A. Martsinkovsky, Absolute, relative, and Tate cohomology of modules of finite Gorenstein dimension, Proc. London Math. Soc. 85(3) (2002), 393–440.

[B1] A. Beligiannis, The homological theory of contravariantly finite subcategories: Auslander–Buchweitz contexts, Gorenstein categories and (co-)stabilization, Comm. Algebra 28(10) (2000), 4547–4596.

[B2] A. Beligiannis, Cohen–Macaulay modules, (co)torsion pairs and virtually Gorenstein algebras, J. Algebra 288(1) (2005), 137–211.

[B3] A. Beligiannis, On algebras of finite Cohen–Macaulay type, Adv. Math. 226(2) (2011), 1973–2019.

[Br] M. Bridger, The $\mathrm{Ext}^i_R(M, R)$ and other invariants of M, Brandeis University, Mathematics. Ph.D. (1967).

[Buch] R.-O. Buchweitz, Maximal Cohen–Macaulay modules and Tate cohomology over Gorenstein rings, Unpublished manuscript, Hamburg (1987), 155pp.

[Che] X. W. Chen, Algebras with radical square zero are either self-injective or CM-free, Proc. Amer. Math. Soc. 140(1) (2012), 93–98.

[Chr] L. W. Christensen. Gorenstein dimensions, Lecture Notes in Math. 1747, Springer-Verlag, 2000.

[CH] L. W. Christensen, H. Holm, Algebras that satisfy Auslander's condition on vanishing of cohomology, Math. Z. 265 (2010), 21–40.

[CF] R. R. Colby, K. R. Fuller, A note on the Nakayama conjecture, Tsukuba J. Math. 14 (1990), 343–352.

[EJ1] E. E. Enochs, O. M. G. Jenda, Gorenstein injective and projective modules, Math. Z. 220(4) (1995), 611–633.

[EJ2] E. E. Enochs, O. M. G. Jenda, Relative homological algebra, De Gruyter Exp. Math. 30. Walter De Gruyter Co., 2000.

[H] D. Happel, Homological conjectures in representation theory of finite-dimensional algebras, Unpublished. See: https://www.math.uni-bielefeld.de/~sek/dim2/happel2.pdf (retrieved Aug 6, 2018).

[HH] C. Huang, Z. Y. Huang, Torsionfree dimension of modules and self-injective dimension of rings, Osaka J. Math. 49 (2012), 21–35.

[J] J. P. Jans, Duality in Noetherian rings, Proc. Amer. Math. Soc. 12 (1961), 829–835.

[JS] D. A. Jorgensen, L. M. Şega, Independence of the total reflexivity conditions for modules, Algebras and Representation Theory 9(2) (2006), 217–226.

[K] B. Keller, Chain complexes and stable categories, Manuscripta Math. 67 (1990), 379–417.

[L] T. S. Lam, Lectures on modules and rings, Springer, 1999.

[M1] R. Marczinzik, Gendo-symmetric algebras, dominant dimensions and Gorenstein homological algebra, arXiv:1608.04212.

[M2] R. Marczinzik, On stable modules that are not Gorenstein projective, arXiv:1709.01132v3.

[R] C. M. Ringel, On the representation dimension of artin algebras, Bull. the Institute of Math., Academia Sinica, Vol. 7(1) (2012), 33–70.

[RX] C. M. Ringel, B. L. Xiong, Finite dimensional algebras with Gorenstein-projective nodes. In preparation.

[RZ1] C. M. Ringel, P. Zhang, Gorenstein-projective and semi-Gorenstein-projective modules II, arXiv:1905.04048.

[RZ2] C. M. Ringel, P. Zhang, Gorenstein-projective modules over short local algebras, arXiv:1912.02081.

[RZ3] C. M. Ringel, P. Zhang, Koszul modules (and the Ω-growth of modules) over short local algebras, arXiv:1912.07512.

[RZ4] C. M. Ringel, P. Zhang, On semi-Gorenstein-projective modules. In preparation.

[S] S. O. Smalø, Local limitations of the Ext functor do not exist, Bull. London Math. Soc. 38 (2006), 97–98.

[Y] Y. Yoshino, A functorial approach to modules of G-dimension zero, Illinois J. Math. 49(3) (2005), 345–367.

[ZX] P. Zhang, B. L. Xiong. Separated monic representations II: Frobenius subcategories and RSS equivalences, Trans. Amer. Math. Soc. 372(2) (2019), 981–1021.

Communicated by J. Toby Stafford
Received 2018-08-06 Revised 2019-07-22 Accepted 2019-08-23

ringel@math.uni-bielefeld.de *Fakultät für Mathematik, PO Box 100131, D-33501, Universität Bielefeld, Germany*

pzhang@sjtu.edu.cn *School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai, 200240, P. R. China*

ALGEBRA AND NUMBER THEORY 14:1 (2020)
dx.doi.org/10.2140/ant.2020.14.37

The 16-rank of $\mathbb{Q}(\sqrt{-p})$

Peter Koymans

Recently, a density result for the 16-rank of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$ was established when p varies among the prime numbers, assuming a short character sum conjecture. We prove the same density result unconditionally.

1. Introduction

If K is a quadratic number field with narrow class group $\mathrm{Cl}(K)$, there is an explicit description of $\mathrm{Cl}(K)[2]$ due to Gauss. Since then the class group of quadratic number fields has been extensively studied. If one is interested in the 2-part of the class group, i.e., $\mathrm{Cl}(K)[2^\infty]$, the explicit description of $\mathrm{Cl}(K)[2]$ is often very useful. It is for this reason that our current understanding of the 2-part of the class group is much better than the p-part for odd p.

In 1984, Cohen and Lenstra put forward conjectures regarding the average behavior of the class group $\mathrm{Cl}(K)$ of imaginary and real quadratic fields K. Despite significant effort, there has been relatively little progress in proving these conjectures. Almost all major results are about the 2-part with the most notable exception being the classical result of Davenport and Heilbronn [1971] regarding the distribution of $\mathrm{Cl}(K)[3]$. Very little is known about $\mathrm{Cl}(K)[p]$ for $p > 3$. The nonabelian version of Cohen–Lenstra has recently also attracted great interest; see [Alberts 2016; Alberts and Klys 2016; Klys 2017; Wood 2019].

Gerth [1984] studied the distribution of $2\mathrm{Cl}(K)[4]$, when the number of prime factors of the discriminant of K is fixed. Fouvry and Klüners [2007] computed all the moments of $2\mathrm{Cl}(K)[4]$, when K varies among imaginary or real quadratic fields. In [Fouvry and Klüners 2006], they deduced the probability that the 4-rank of a quadratic field has a given value. Their work was based on earlier ideas of Heath-Brown [1994].

The study of $\mathrm{Cl}(K)[2^\infty]$ has often been conducted through the lens of *governing fields*. Let $k \geq 1$ be an integer and let d be an integer with $d \not\equiv 2 \bmod 4$. For a

MSC2010: primary 11R29; secondary 11N45, 11R45.
Keywords: arithmetic statistics, class groups.

finite abelian group A we define the 2^k-rank of A to be $\mathrm{rk}_{2^k} A := \dim_{\mathbb{F}_2} 2^{k-1}A/2^kA$. Then a governing field $M_{d,k}$ is a normal field extension of $\mathbb{Q}$ such that

$$\mathrm{rk}_{2^k} \mathrm{Cl}(\mathbb{Q}(\sqrt{dp}))$$

is determined by the splitting of p in $M_{d,k}$. Cohn and Lagarias [1983] were the first to define the concept of a governing field, and conjectured that they always exist.

If $k \leq 3$, then governing fields are known to exist for all values of d. In case $k = 2$ this follows from [Rédei 1934]. Stevenhagen dealt with the case $k = 3$ [1988]. The topic was recently revisited by Smith [2016], who found a very explicit description for $M_{d,3}$ for most values of d. He then used this description to prove density results for $4\mathrm{Cl}(K)[8]$ assuming GRH. Not much later Smith [2017] introduced *relative governing fields*, which allowed him to prove the most impressive result that $2\mathrm{Cl}(K)[2^\infty]$ has the expected distribution when K varies among all imaginary quadratic fields.

If we let $P(d, k)$ be the statement that a governing field $M_{d,k}$ exists, then there is currently not a single value of d for which the truth or falsehood of $P(d, 4)$ is known. This has been the most significant obstruction in proving density results for the 16-rank in thin families of the shape $\{\mathbb{Q}(\sqrt{dp})\}_{p \text{ prime}}$.

This barrier was first broken by Milovic [2017a], who dealt with the 16-rank in the family $\{\mathbb{Q}(\sqrt{-2p})\}_{p \equiv -1 \bmod 4}$. Milovic proves his density result with Vinogradov's method, and does not rely on the existence of a governing field. His use of Vinogradov's method was inspired by work of Friedlander et al. [2013], which is based on earlier work of Friedlander and Iwaniec [1998].

Density results for the families $\{\mathbb{Q}(\sqrt{-2p})\}_{p \equiv 1 \bmod 4}$ and $\{\mathbb{Q}(\sqrt{-p})\}_p$ were established by Milovic and the author; see respectively [Koymans and Milovic 2019a; 2019b] with the latter work being conditional on a short character sum conjecture. Both of these works follow the ideas of [Friedlander et al. 2013] closely in their treatment of the sums of type I; see Section 3 for a definition. However, if one applies the method of [Friedlander et al. 2013] to a number field of degree n, one is naturally led to consider character sums of modulus q and length $q^{1/n}$.

In [Koymans and Milovic 2019a] we apply the method from [Friedlander et al. 2013] to a number field of degree 4. This leads to character sums just outside the range of the Burgess bound. Fortunately, the lemmas in Section 3.2 of [Koymans and Milovic 2019a] allow us to reduce the size of the modulus from q to $q^{1/2}$, and this enables us to deal with the sums of type I unconditionally. In [Koymans and Milovic 2019b] we use a criterion for the 16-rank of $\mathbb{Q}(\sqrt{-p})$ due to [Bruin and Hemenway 2013], and this criterion is stated most naturally over $\mathbb{Q}(\zeta_8, \sqrt{1+i})$, which has degree 8. The resulting character sums are far outside the reach of the Burgess bound and we resort to assuming a short character sum conjecture; see [Koymans and Milovic 2019b, p. 8].

In this paper we manage to deal with the 16-rank of $\mathbb{Q}(\sqrt{-p})$ unconditionally by using a criterion of Leonard and Williams [1982], which one can naturally state over $\mathbb{Q}(\zeta_8)$. However, the Leonard–Williams criterion has the significant downside that it is the product of two residue symbols instead of one residue symbol, namely a quadratic and a quartic residue symbol. The resulting sums of type I can still not be treated unconditionally with the method from [Friedlander et al. 2013]. Instead, we use a rather ad hoc argument to deal with the resulting character sum.

Theorem 1.1. *Let $h(-p)$ be the class number of $\mathbb{Q}(\sqrt{-p})$. Then*

$$\lim_{X\to\infty} \frac{|\{p \text{ prime} : p \le X \text{ and } 16 \mid h(-p)\}|}{|\{p \text{ prime} : p \le X\}|} = \frac{1}{16}.$$

Milovic [2017b] has previously shown that there are infinitely many primes p with 16 dividing $h(-p)$. Theorem 1.1 gives an affirmative answer to conjectures in both [Cohn and Lagarias 1984] and [Stevenhagen 1993]. For p a prime number, we define e_p by

$$e_p := \begin{cases} 1 & \text{if } 16 \mid h(-p), \\ -1 & \text{if } 8 \mid h(-p),\ 16 \nmid h(-p), \\ 0 & \text{otherwise.} \end{cases} \tag{1-1}$$

Theorem 1.1 is an immediate consequence of the following theorem.

Theorem 1.2. *We have*

$$\sum_{p \le X} e_p \ll \frac{X}{\exp((\log X)^{0.1})}.$$

It is natural to wonder if the other conditional results in [Koymans and Milovic 2019b] can be proven unconditionally using the methods from this paper. This is likely to be the case, but it would require some effort to obtain suitable algebraic results similar to the Leonard–Williams criterion [1982] used in this paper.

We believe that the ideas introduced by Smith [2017] do not apply to the thin families that we deal with here. Indeed, in Smith's paper a crucial ingredient for both the algebraic and analytic part is the fact that a typical integer N has roughly $\log\log N$ prime divisors and that $\log\log N$ goes to infinity as N goes to infinity.

2. Preliminaries

Quadratic and quartic reciprocity. Let K be a number field with ring of integers O_K. We say that an ideal $\mathfrak{n}$ of O_K is odd if $(\mathfrak{n}, 2) = (1)$. Similarly, we say that an element w of O_K is odd if the ideal generated by w is odd. If $\mathfrak{p}$ is an odd prime ideal of O_K and $\alpha \in O_K$, we define the quadratic residue symbol

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{2,K} := \begin{cases} 1 & \text{if } \alpha \notin \mathfrak{p} \text{ and } \alpha \equiv \beta^2 \bmod \mathfrak{p} \text{ for some } \beta \in O_K, \\ -1 & \text{if } \alpha \notin \mathfrak{p} \text{ and } \alpha \not\equiv \beta^2 \bmod \mathfrak{p} \text{ for all } \beta \in O_K, \\ 0 & \text{if } \alpha \in \mathfrak{p}. \end{cases}$$

Then Euler's criterion states

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{2,K} \equiv \alpha^{\frac{1}{2}(\mathrm{N}(\mathfrak{p})-1)} \bmod \mathfrak{p}.$$

For a general odd ideal $\mathfrak{n}$ of O_K, we define

$$\left(\frac{\alpha}{\mathfrak{n}}\right)_{2,K} := \prod_{\mathfrak{p}^e \| \mathfrak{n}} \left(\left(\frac{\alpha}{\mathfrak{p}}\right)_{2,K}\right)^e.$$

Furthermore, for odd $\beta \in O_K$ we set

$$\left(\frac{\alpha}{\beta}\right)_{2,K} := \left(\frac{\alpha}{(\beta)}\right)_{2,K}.$$

We say that an element $\alpha \in K$ is totally positive if for all embeddings σ of K into $\mathbb{R}$ we have $\sigma(\alpha) > 0$. In particular, all elements of a totally complex number field are totally positive. We will make extensive use of the law of quadratic reciprocity.

Theorem 2.1. *Let $\alpha, \beta \in O_K$ be odd. If α or β is totally positive, we have*

$$\left(\frac{\alpha}{\beta}\right)_{2,K} = \mu(\alpha, \beta)\left(\frac{\beta}{\alpha}\right)_{2,K},$$

where $\mu(\alpha, \beta) \in \{\pm 1\}$ depends only on the congruence classes of α and β modulo 8.

Proof. This follows from Lemma 2.1 of [Friedlander et al. 2013]. □

If $K = \mathbb{Q}$, we shall drop the subscript. In this case the symbol

$$\left(\frac{\cdot}{\cdot}\right)$$

is to be interpreted as the Kronecker symbol, which is an extension of the quadratic residue symbol to allow for even arguments in the bottom. We presume that the reader is familiar with the quadratic reciprocity law for the Kronecker symbol. Now let K be a number field containing $\mathbb{Q}(i)$ still with ring of integers O_K. For $\alpha \in O_K$ and $\mathfrak{p}$ an odd prime ideal of O_K, we define the quartic residue symbol $(\alpha/\mathfrak{p})_{4,K}$ to be the unique element in $\{\pm 1, \pm i, 0\}$ such that

$$\left(\frac{\alpha}{\mathfrak{p}}\right)_{4,K} \equiv \alpha^{\frac{1}{4}(\mathrm{N}(\mathfrak{p})-1)} \bmod \mathfrak{p}.$$

We extend the quartic residue symbol to all odd ideals $\mathfrak{n}$ and then to all odd elements β in the same way as the quadratic residue symbol. Then we have the following theorem.

Theorem 2.2. *Let $\alpha, \beta \in \mathbb{Z}[\zeta_8]$ with β odd. Then for fixed α, the symbol $(\alpha/\beta)_{4,\mathbb{Q}(\zeta_8)}$ depends only on β modulo $16\alpha\mathbb{Z}[\zeta_8]$. Furthermore, if α is also odd, we have*

$$\left(\frac{\alpha}{\beta}\right)_{4,\mathbb{Q}(\zeta_8)} = \mu(\alpha, \beta)\left(\frac{\beta}{\alpha}\right)_{4,\mathbb{Q}(\zeta_8)},$$

where $\mu(\alpha, \beta) \in \{\pm 1, \pm i\}$ depends only on the congruence classes of α and β modulo 16.

Proof. Use Proposition 6.11 of [Lemmermeyer 2000, p. 199]. □

A fundamental domain. Let F be a number field of degree n over $\mathbb{Q}$ and let O_F be its ring of integers. Suppose that F has r real embeddings and s pairs of conjugate complex embeddings so that $r + 2s = n$. Define T to be the torsion subgroup of O_F^*. Then, by Dirichlet's unit theorem, there exists a free abelian group $V \subseteq O_F^*$ of rank $r + s - 1$ with $O_F^* = T \times V$. Fix one choice of such a V.

There is a natural action of V on O_F. The goal of this subsection is to construct a fundamental domain $\mathcal{D}$ for this action. Such a fundamental domain allows us to transform a sum over ideals into a sum over elements. It will be important that the resulting fundamental domain has nice geometrical properties, so that we have good control over the elements we are summing.

Fix an integral basis $\omega_1, \ldots, \omega_n$ for O_F. We view $\omega_1, \ldots, \omega_n$ as an ordered list and write ω for this ordered list. Then we get an isomorphism of $\mathbb{Q}$-vector spaces $i_\omega : \mathbb{Q}^n \to F$, where i_ω is given by $(a_1, \ldots, a_n) \mapsto a_1\omega_1 + \cdots + a_n\omega_n$. For a subset $S \subseteq \mathbb{R}^n$ and an element $\alpha \in F$, we will say that $\alpha \in S$ if $i_\omega^{-1}(\alpha) \in S$. Define for our integral basis ω and a real number $X > 0$

$$B(X, \omega) := \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n : \left|\prod_{i=1}^{n}(x_1\sigma_i(\omega_1) + \cdots + x_n\sigma_i(\omega_n))\right| \le X\right\},$$

where $\sigma_1, \ldots, \sigma_n$ are the embeddings of F into $\mathbb{C}$.

Lemma 2.3. *Let F be a number field with ring of integers O_F and integral basis $\omega = \{\omega_1, \ldots, \omega_n\}$. Choose a splitting $O_F^* = T \times V$, where T is the torsion subgroup of O_F^*. There exists a subset $\mathcal{D} \subseteq \mathbb{R}^n$ such that:*

(i) *For all $\alpha \in O_F \setminus \{0\}$, there exists a unique $v \in V$ such that $v\alpha \in \mathcal{D}$. Furthermore, we have the equality*

$$\{u \in O_F^* : u\alpha \in \mathcal{D}\} = \{tv : t \in T\}.$$

(ii) *$\mathcal{D} \cap B(1, \omega)$ has an $(n-1)$-Lipschitz parametrizable boundary.*

(iii) *There is a constant $C(\omega)$ depending only on ω such that for all $\alpha \in \mathcal{D}$ we have $|a_i| \le C(\omega) \cdot |\mathrm{N}(\alpha)|^{\frac{1}{n}}$, where $a_i \in \mathbb{Z}$ are such that $\alpha = a_1\omega_1 + \cdots + a_n\omega_n$.*

Proof. This is Lemma 3.5 of [Koymans and Milovic 2019a]. □

We will use Lemma 2.3 for $F := \mathbb{Q}(\zeta_8)$; in order to do so we must make some choices. We choose $V := \langle 1+\sqrt{2}\rangle$ and integral basis $\omega := \{1, \zeta_8, \zeta_8^2, \zeta_8^3\}$. The resulting fundamental domain will be called $\mathcal{D}$, and we define $\mathcal{D}(X) := \mathcal{D} \cap B(X, \omega)$.

3. The sieve

Let $\{a_p\}$ be a sequence of complex numbers indexed by the primes and define

$$S(X) := \sum_{p \le X} a_p.$$

To prove our main theorem, we must prove oscillation of $S(X)$ for the specific sequence $\{e_p\}$ defined in (1-1). There are relatively few methods that can deal with such sums. The most common approach is to attach an L-function and then use the zero-free region. This approach requires that our sequence $\{e_p\}$ has good multiplicative properties. It turns out that $\{e_p\}$ is instead twisted multiplicative (see Lemmas 6.1 and 6.3), and this suggests we use Vinogradov's method instead.

Recall that $h(-p)$ denotes the class number of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$. By definition of e_p we have $e_p = 0$ if and only if $8 \nmid h(-p)$. It is well-known that $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ is a *governing field* for the 8-rank of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$, in fact a prime p splits completely in $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ if and only if $8 \mid h(-p)$. This is extremely convenient. Indeed, if we apply Vinogradov's method to our governing field, primes of degree 1 will give the dominant contribution and these primes automatically have $e_p \neq 0$.

Unfortunately, $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ is a field of degree 8, which is simply too large to make our analytic methods work unconditionally. Indeed, using the same approach for the sums of type I as [Friedlander et al. 2013], one ends up with short character sums of modulus q and length roughly $q^{1/8}$, which is far outside the reach of the Burgess bound. However, assuming a short character sum conjecture, one still obtains the desired oscillation and this is the approach taken in [Koymans and Milovic 2019b]. Instead we work over $\mathbb{Q}(\zeta_8)$; fortunately, $\mathbb{Q}(\zeta_8, \sqrt{1+i})$ is an abelian extension of $\mathbb{Q}(\zeta_8)$, which implies that the splitting of a prime $\mathfrak{p}$ of $\mathbb{Q}(\zeta_8)$ in the extension $\mathbb{Q}(\zeta_8, \sqrt{1+i})/\mathbb{Q}(\zeta_8)$ is determined by a congruence condition. Such a congruence condition can easily be incorporated in Vinogradov's method.

We will follow Section 5 of [Friedlander et al. 2013], which adapted Vinogradov's method to number fields. Let F be a number field. Define for a nonzero ideal $\mathfrak{n}$ of O_F

$$\Lambda(\mathfrak{n}) := \begin{cases} \log \mathrm{N}\mathfrak{p} & \text{if } \mathfrak{n} = \mathfrak{p}^l, \\ 0 & \text{otherwise.} \end{cases}$$

and suppose that we want to prove oscillation of

$$S(X) := \sum_{\mathrm{N}\mathfrak{n} \le X} a_{\mathfrak{n}} \Lambda(\mathfrak{n}),$$

where $a_{\mathfrak{n}}$ is of absolute value at most 1. The power of Vinogradov's method lies in the fact that one does not have to deal with $S(X)$ directly. Instead one has to prove cancellations of

$$A(X, \mathfrak{d}) := \sum_{\substack{\mathrm{N}\mathfrak{n} \leq X \\ \mathfrak{d} | \mathfrak{n}}} a_{\mathfrak{n}},$$

which are traditionally called sums of type I or linear sums, and

$$B(M, N) := \sum_{\mathrm{N}\mathfrak{m} \leq M} \sum_{\mathrm{N}\mathfrak{n} \leq N} \alpha_{\mathfrak{m}} \beta_{\mathfrak{n}} a_{\mathfrak{m}\mathfrak{n}},$$

which are traditionally called sums of type II or bilinear sums. It is important to remark that $S(X)$ depends only on $a_{\mathfrak{n}}$ with $\mathfrak{n}$ a prime power, while $A(X, \mathfrak{d})$ and $B(M, N)$ certainly do not. This gives a substantial amount of flexibility, since we may define $a_{\mathfrak{n}}$ on composite ideals $\mathfrak{n}$ in any way we like provided that we can prove oscillation of $A(X, \mathfrak{d})$ and $B(M, N)$. Constructing a suitable sequence $a_{\mathfrak{n}}$ will be the goal of Section 4. We are now ready to state the precise version of Vinogradov's method we are going to use.

Proposition 3.1. *Let F be a number field and let $a_{\mathfrak{n}}$ be a sequence of complex numbers, indexed by the ideals of O_F, with $|a_{\mathfrak{n}}| \leq 1$. If $0 < \theta_1, \theta_2 < 1$ and $\theta_3 > 0$ are such that*

- *we have for all ideals $\mathfrak{d}$ of O_F*

$$A(X, \mathfrak{d}) \ll_{F, a_{\mathfrak{n}}, \theta_1} \frac{X}{\exp((\log X)^{\theta_1})}; \tag{3-1}$$

- *we have for all sequences of complex numbers $\{\alpha_{\mathfrak{m}}\}$ and $\{\beta_{\mathfrak{n}}\}$ of absolute value at most 1*

$$B(M, N) \ll_{F, a_{\mathfrak{n}}, \theta_2} (M + N)^{\theta_2} (MN)^{1-\theta_2} (\log MN)^{\theta_3}, \tag{3-2}$$

then we have for all $c < \theta_1$

$$S(X) \ll_{c, F, a_{\mathfrak{n}}, \theta_1, \theta_2, \theta_3} \frac{X}{\exp((\log X)^{c})}.$$

Proof. This quickly follows from Proposition 5.1 of [Friedlander et al. 2013] with $y := \exp((\log X)^{\frac{1}{2}(c+\theta_1)})$. □

The remainder of this paper is devoted to the three major tasks that are left. We start by constructing a suitable sequence $a_{\mathfrak{n}}$ in Section 4 to which we will apply Proposition 3.1 with $F = \mathbb{Q}(\zeta_8)$. The main result of Section 5 is Proposition 5.1, which proves (3-1) for $\theta_1 = 0.2$. Finally, we prove in Section 6 that (3-2) holds with $\theta_2 = \frac{1}{24}$; this is the content of Proposition 6.6. Once we have proven Propositions 5.1 and 6.6, the proof of Theorem 1.2 is complete.

4. Definition of the sequence

By Gauss genus theory we know that the 2-part of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$ is cyclic, and the 2-part of $\mathrm{Cl}(\mathbb{Q}(\sqrt{-p}))$ is trivial if and only if $p \equiv 3 \mod 4$. Let us recall a criterion for $16 \mid h(-p)$ due to Leonard and Williams [1982]. We have

$$4 \mid h(-p) \iff p \equiv 1 \mod 8.$$

Now suppose that $4 \mid h(-p)$. There exist positive integers g and h satisfying

$$p = 2g^2 - h^2.$$

Then a classical result of Hasse [1969] is

$$8 \mid h(-p) \iff \left(\frac{g}{p}\right) = 1 \text{ and } p \equiv 1 \mod 8$$

or equivalently

$$8 \mid h(-p) \iff \left(\frac{-1}{g}\right) = 1 \text{ and } p \equiv 1 \mod 8.$$

We are now ready to state the result of Leonard and Williams [1982]. If p is a prime number with $8 \mid h(-p)$, we have

$$16 \mid h(-p) \iff \left(\frac{g}{p}\right)_4 \left(\frac{2h}{g}\right) = 1.$$

With this in mind, we are going to define a sequence $\{a_{\mathfrak{n}}\}$, indexed by the integral ideals of $\mathbb{Z}[\zeta_8]$, such that for all unramified prime ideals $\mathfrak{p}$ in $\mathbb{Z}[\zeta_8]$ of norm p

$$a_{\mathfrak{p}} = \begin{cases} 1 & \text{if } 16 \mid h(-p), \\ -1 & \text{if } 8 \mid h(-p),\ 16 \nmid h(-p), \\ 0 & \text{otherwise.} \end{cases} \tag{4-1}$$

The sequence $\{a_{\mathfrak{n}}\}$ will be constructed in such a way that we can prove the two estimates in Propositions 5.1 and 6.6. Before we move on, it will be useful to recall some standard facts about $\mathbb{Z}[\zeta_8]$. The ring $\mathbb{Z}[\zeta_8]$ is a PID with unit group generated by ζ_8 and $\epsilon := 1 + \sqrt{2}$. Odd primes are unramified in $\mathbb{Z}[\zeta_8]$, while 2 is totally ramified. Furthermore, an odd prime p splits completely in $\mathbb{Z}[\zeta_8]$ if and only if $p \equiv 1 \mod 8$ if and only if $4 \mid h(-p)$. We will make extensive use of the following field diagram.

$$
\begin{tikzcd}
& M := \mathbb{Q}(\zeta_8) \arrow[dl, dash, "\langle\sigma\rangle"'] \arrow[d, dash, "\langle\sigma\tau\rangle"] \arrow[dr, dash, "\langle\tau\rangle"] & \\
\mathbb{Q}(i) \arrow[dr, dash] & \mathbb{Q}(i\sqrt{2}) \arrow[d, dash] & \mathbb{Q}(\sqrt{2}) \arrow[dl, dash] \\
& \mathbb{Q} &
\end{tikzcd}
$$

If $\mathfrak{n}$ is not odd, we set $a_{\mathfrak{n}} := 0$. From now on $\mathfrak{n}$ is an odd, integral, nonzero ideal of $\mathbb{Z}[\zeta_8]$ and w is a generator of $\mathfrak{n}$. We can write w as

$$w = a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3$$

for certain $a, b, c, d \in \mathbb{Z}$. Define $u, v \in \mathbb{Z}$ by

$$w\tau(w) = u + v\sqrt{2}.$$

We can explicitly compute u and v using the formulas

$$u = \frac{w\tau(w) + \sigma(w)\sigma\tau(w)}{2} = a^2 + b^2 + c^2 + d^2, \tag{4-2}$$

$$v = \frac{w\tau(w) - \sigma(w)\sigma\tau(w)}{2\sqrt{2}} = ab - ad + bc + cd. \tag{4-3}$$

Since w is odd, it follows that $\mathrm{N}w \equiv 1 \bmod 8$. Then it follows from

$$\mathrm{N}w = u^2 - 2v^2$$

that u is an odd integer and v is an even integer. Set

$$g := u + v, \quad h := u + 2v,$$

so that g is an odd integer and h is an odd integer, not necessarily positive. We claim that g is positive. Indeed

$$\begin{aligned} g &= a^2 + b^2 + c^2 + d^2 + ab - ad + bc + cd \\ &= \tfrac{1}{2}(a+b)^2 + \tfrac{1}{2}(a-d)^2 + \tfrac{1}{2}(b+c)^2 + \tfrac{1}{2}(c+d)^2 > 0. \end{aligned}$$

By construction g and h satisfy

$$\mathrm{N}w = 2g^2 - h^2.$$

We start by showing that the value of

$$\left(\frac{-1}{g}\right) \tag{4-4}$$

does not depend on the choice of generator w of our ideal $\mathfrak{n}$.

Lemma 4.1. *Let $\mathfrak{n}$ be an odd, integral ideal of $\mathbb{Z}[\zeta_8]$. Then the value of* (4-4) *is the same for all generators w of $\mathfrak{n}$.*

Proof. Suppose that we replace w by $\zeta_8 w$. Because $\zeta_8 \tau(\zeta_8) = 1$, it follows that u and v, hence also g, do not change. Suppose instead that we replace w by ϵw. In this case u becomes $3u + 4v$ and v becomes $2u + 3v$, so g becomes $5u + 7v$. Hence our lemma boils down to

$$\left(\frac{-1}{u+v}\right) = \left(\frac{-1}{5u+7v}\right),$$

which holds if and only if

$$u + v \equiv 5u + 7v \mod 4.$$

But recall that v is even by our assumption that w is odd. □

We define for odd $w \in \mathbb{Z}[\zeta_8]$ the following symbol:

$$[w] := \left(\frac{g}{w}\right)_{4,M} \left(\frac{2h}{g}\right),$$

where we remind the reader that M is defined to be $\mathbb{Q}(\zeta_8)$. We express this as

$$[w] = [w]_1 [w]_2, \quad [w]_1 := \left(\frac{g}{w}\right)_{4,M}, \quad [w]_2 := \left(\frac{2h}{g}\right). \tag{4-5}$$

It is easily checked that $[\zeta_8 w] = [w]$. Unfortunately, it is not always true that $[\epsilon w] = [w]$. To get around this, we need the following lemma.

Lemma 4.2. *We have for all odd w*

$$[\epsilon^4 w] = [w].$$

Proof. We have for any odd w

$$[w]_1 = \left(\frac{g}{w}\right)_{4,M} = \left(\frac{u+v}{w}\right)_{4,M} = \left(\frac{\left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)\sigma(w)\sigma\tau(w)}{w}\right)_{4,M}, \tag{4-6}$$

where we use the explicit formulas for u and v, see (4-2) and (4-3), in terms of w.

From this expression it quickly follows that $[\epsilon^2 w]_1 = [w]_1$. We also have

$$\begin{aligned}[w]_2 &= \left(\frac{2h}{g}\right) = \left(\frac{2u+4v}{u+v}\right) = \left(\frac{2}{u+v}\right)\left(\frac{v}{u+v}\right) \\ &= \left(\frac{2}{u+v}\right)\left(\frac{-u}{u+v}\right) = \left(\frac{-2}{u+v}\right)\left(\frac{v}{u}\right)(-1)^{\frac{1}{2}(u-1)\cdot\frac{1}{2}(u+v-1)}. \end{aligned} \tag{4-7}$$

A straightforward computation shows that the u and v associated to $\epsilon^4 w$ are respectively $u_1 := 577u + 816v$ and $v_1 := 408u + 577v$. Then we have

$$\left(\frac{v}{u}\right) = \left(\frac{408u+577v}{577u+816v}\right) = \left(\frac{v_1}{u_1}\right) \tag{4-8}$$

due to Proposition 2 in [Milovic 2017a]. It will be useful to observe that the following congruences hold true:

$$u \equiv u_1 \bmod 8, \quad v \equiv v_1 \bmod 8.$$

This immediately implies

$$\left(\frac{-2}{u+v}\right) = \left(\frac{-2}{u_1+v_1}\right), \tag{4-9}$$

and therefore the lemma. □

With this out of the way, we have all the tools necessary to define $a_{\mathfrak{n}}$. Suppose that $\mathfrak{n}$ is an odd, integral ideal of $\mathbb{Z}[\zeta_8]$ with generator w. Then we define

$$a_{\mathfrak{n}} := \begin{cases} \frac{1}{4}\big([w] + [\epsilon w] + [\epsilon^2 w] + [\epsilon^3 w]\big) & \text{if } w \text{ satisfies (4-4)}, \\ 0 & \text{otherwise}. \end{cases} \tag{4-10}$$

for any generator w of $\mathfrak{n}$. Here we say that w satisfies (4-4) if $(-1/g) = 1$, where g is defined in terms of w as above. Then an application of Lemmas 4.1 and 4.2 shows that (4-10) is indeed well-defined.

Lemma 4.3. *The sequence $a_{\mathfrak{n}}$ satisfies* (4-1) *for all unramified prime ideals $\mathfrak{p}$ of degree* 1 *in $\mathbb{Z}[\zeta_8]$.*

Proof. Let $\mathfrak{p}$ be an unramified prime ideal of degree 1 in $\mathbb{Z}[\zeta_8]$ and let w be a generator of $\mathfrak{p}$. Put $p := \mathrm{N}w$. Lemma 4.1 and the aforementioned result of Hasse imply

$$w \text{ does not satisfy (4-4)} \Longleftrightarrow 8 \nmid h(-p),$$

and $a_{\mathfrak{p}}$ is indeed 0 in this case. Now suppose that w does satisfy (4-4). Recall that

$$[w] = \left(\frac{g}{w}\right)_{4,M}\left(\frac{2h}{g}\right),$$

where g and h are explicit functions of w. We stress that these g and h are not necessarily the same g and h from Leonard and Williams. Indeed, Leonard and

Williams require g and h to be positive, while our h is not necessarily positive. However, since w satisfies (4-4), their criterion remains valid irrespective of the sign of h. Then, the criterion implies

$$[w] = [\epsilon w] = [\epsilon^2 w] = [\epsilon^3 w].$$

Furthermore, the criterion also shows that

$$[w] = 1 \iff 16 \mid h(-p).$$

This completes the proof of our lemma. $\square$

5. Sums of type I

The goal of this section is to bound the sum

$$A(X, \mathfrak{d}) = \sum_{\substack{N\mathfrak{n} \leq X \\ \mathfrak{d} \mid \mathfrak{n}}} a_{\mathfrak{n}} = \sum_{\substack{N\mathfrak{n} \leq X \\ \mathfrak{d} \mid \mathfrak{n}, \mathfrak{n} \text{ odd}}} a_{\mathfrak{n}}.$$

By picking a generator for $\mathfrak{n}$ we obtain

$$A(X, \mathfrak{d}) = \frac{1}{8} \sum_{\substack{w \in \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \text{ odd}}} a_{(w)} = \frac{1}{32} \sum_{\substack{w \in \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \text{ odd}}} \mathbf{1}_{w \text{ sat. (4-4)}} \big([w] + [\epsilon w] + [\epsilon^2 w] + [\epsilon^3 w]\big).$$

We define for $i = 0, \ldots, 3$ and ρ an invertible congruence class modulo 2^{10}

$$A(X, \mathfrak{d}, u_i, \rho) := \sum_{\substack{w \in u_i \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \equiv \rho \bmod 2^{10}}} [w] = \sum_{\substack{w \in u_i \mathcal{D}(X) \\ w \equiv 0 \bmod \mathfrak{d} \\ w \equiv \rho \bmod 2^{10}}} \left(\frac{g}{w}\right)_{4,M} \left(\frac{2h}{g}\right),$$

where $u_i := \epsilon^i$. With this definition in place, we may split $A(X, \mathfrak{d})$ as follows

$$A(X, \mathfrak{d}) = \frac{1}{32} \sum_{i=0}^{3} \sum_{\rho \in (O_M / 2^{10} O_M)^*} \mathbf{1}_{\rho \text{ sat. (4-4)}} A(X, \mathfrak{d}, u_i, \rho),$$

since the truth of (4-4) depends only on w modulo 4. Then it is enough to bound each individual sum $A(X, \mathfrak{d}, u_i, \rho)$. In order to bound this sum, our first step is to carefully rewrite the symbol $[w]$ in a more tractable form. While doing so, we will find some hidden cancellation between $[w]_1$ and $[w]_2$ that is vital for making our results unconditional.

Throughout this section we use the convention that $\mu(\cdot) \in \{\pm 1, \pm i\}$ is a function depending only on the variables between the parentheses; at each occurence $\mu(\cdot)$ may be a different function. Since our cancellation will come from fixing b, c and d

while varying a, factors of the shape $\mu(\rho, b, c, d)$ will present no issues for us. Let us start by rewriting $[w]_2$. It follows from (4-7) that

$$\left(\frac{2h}{g}\right) = \left(\frac{v}{u}\right)\mu(\rho). \tag{5-1}$$

Using the formulas for u and v we get

$$\left(\frac{v}{u}\right) = \left(\frac{ab - ad + bc + cd}{a^2 + b^2 + c^2 + d^2}\right). \tag{5-2}$$

If v is not zero, we can uniquely factor v as

$$v := v_1 v_2 t, \tag{5-3}$$

where v_1 is an odd, positive integer satisfying $\gcd(v_1, b - d) = 1$, v_2 is an odd integer consisting only of primes dividing $b - d$ and t is positive and only divisible by powers of 2. Then we have

$$\left(\frac{ab - ad + bc + cd}{a^2 + b^2 + c^2 + d^2}\right) = \left(\frac{v_1}{a^2 + b^2 + c^2 + d^2}\right)\left(\frac{t v_2}{a^2 + b^2 + c^2 + d^2}\right). \tag{5-4}$$

Let ρ' be the congruence class of v_1 modulo 8. Using the following identity modulo v

$$a^2(b - d)^2 \equiv c^2(b + d)^2 \bmod v$$

and the fact that this identity continues to hold for any divisor of v, in particular for v_1, we rewrite the first factor of (5-4) as follows

$$\begin{aligned}
\left(\frac{v_1}{a^2 + b^2 + c^2 + d^2}\right) &= \mu(\rho, \rho')\left(\frac{a^2 + b^2 + c^2 + d^2}{v_1}\right) \\
&= \mu(\rho, \rho')\left(\frac{(a^2 + b^2 + c^2 + d^2)(b - d)^2}{v_1}\right) \\
&= \mu(\rho, \rho')\left(\frac{a^2(b - d)^2 + (b^2 + c^2 + d^2)(b - d)^2}{v_1}\right) \\
&= \mu(\rho, \rho')\left(\frac{c^2(b + d)^2 + (b^2 + c^2 + d^2)(b - d)^2}{v_1}\right) \\
&= \mu(\rho, \rho')\left(\frac{(b^2 + d^2)(2c^2 + (b - d)^2)}{v_1}\right).
\end{aligned} \tag{5-5}$$

Stringing together (5-1), (5-2), (5-4) and (5-5), we conclude that

$$\left(\frac{2h}{g}\right) = \mu(\rho, \rho')\left(\frac{(b^2 + d^2)(2c^2 + (b - d)^2)}{v_1}\right)\left(\frac{t v_2}{a^2 + b^2 + c^2 + d^2}\right). \tag{5-6}$$

Our next goal is to simplify $[w]_1$. We have, by (4-6) and Theorem 2.2,

$$\left(\frac{g}{w}\right)_{4,M} = \left(\frac{\left(\frac{1}{2}-\frac{1}{2\sqrt{2}}\right)\sigma(w)\sigma\tau(w)}{w}\right)_{4,M} = \mu(\rho)\left(\frac{\sigma(w)\sigma\tau(w)}{w}\right)_{4,M}. \tag{5-7}$$

The quartic residue symbol in (5-7) is the product of two quartic residue symbols. One of them is equal to

$$\begin{aligned}\left(\frac{\sigma\tau(w)}{w}\right)_{4,M} &= \left(\frac{a+d\zeta_8-c\zeta_8^2+b\zeta_8^3}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M} = \left(\frac{-2c\zeta_8^2+(d-b)(\zeta_8-\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M} \\ &= \left(\frac{\zeta_8^2}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M}\left(\frac{-2c+(b-d)(\zeta_8+\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M} \\ &= \mu(\rho)\left(\frac{-2c+(b-d)(\zeta_8+\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M},\end{aligned} \tag{5-8}$$

where the last equality is due to Theorem 2.2. For the remainder of this section we assume that $b-d$ is not zero. We factor $-2c+(b-d)(\zeta_8+\zeta_8^3)$ in the ring $\mathbb{Z}[\sqrt{-2}]$ as

$$-2c+(b-d)(\zeta_8+\zeta_8^3) = \eta^4 e_0 e$$

with η and e_0 consisting only of even prime factors, e_0 not divisible by a nontrivial fourth power and e odd. This factorization is unique up to multiplication by units. Then we have, by Theorem 2.2,

$$\left(\frac{-2c+(b-d)(\zeta_8+\zeta_8^3)}{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}\right)_{4,M} = \mu(\rho,b,c,d)\left(\frac{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}{e}\right)_{4,M}. \tag{5-9}$$

But a simple computation shows

$$a+b\zeta_8+c\zeta_8^2+d\zeta_8^3 \equiv \sigma\tau(a+b\zeta_8+c\zeta_8^2+d\zeta_8^3) \bmod e.$$

Let $\mathfrak{p}$ be a prime in $\mathbb{Z}[\sqrt{-2}]$ that divides e. Then we may replace $a+b\zeta_8+c\zeta_8^2+d\zeta_8^3$ by some element in $\mathbb{Z}[\sqrt{-2}]$ by Lemma 3.4 of [Koymans and Milovic 2019a]. In case $\mathfrak{p}$ splits in M, we apply Lemma 3.2 of [Koymans and Milovic 2019a]. While if $\mathfrak{p}$ remains inert, we see that $\mathfrak{p}$ is of degree 1 and $\mathrm{N}\mathfrak{p} \equiv 3 \bmod 8$. In this case we apply Lemma 3.3 of [Koymans and Milovic 2019a]. Hence in all cases

$$\left(\frac{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}{\mathfrak{p}}\right)_{4,M} = \mathbb{1}_{\gcd(a+b\zeta_8+c\zeta_8^2+d\zeta_8^3,\mathfrak{p})=(1)}.$$

This yields

$$\left(\frac{a+b\zeta_8+c\zeta_8^2+d\zeta_8^3}{e}\right)_{4,M} = \mathbb{1}_{\gcd(w,\sigma\tau(w))=(1)}. \tag{5-10}$$

We deduce from (5-8)–(5-10) that

$$\left(\frac{\sigma\tau(w)}{w}\right)_{4,M} = \mu(\rho, b, c, d)\mathbb{1}_{\gcd(w,\sigma\tau(w))=(1)}. \tag{5-11}$$

We will now study the other quartic residue symbol in (5-7) using very similar methods. We start with the identity

$$\begin{aligned}\left(\frac{\sigma(w)}{w}\right)_{4,M} &= \left(\frac{a - b\zeta_8 + c\zeta_8^2 - d\zeta_8^3}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} = \left(\frac{-2\zeta_8(b + d\zeta_8^2)}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} \\ &= \left(\frac{-2\zeta_8}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M}\left(\frac{b + d\zeta_8^2}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} \\ &= \mu(\rho)\left(\frac{b + d\zeta_8^2}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M},\end{aligned} \tag{5-12}$$

where we use Theorem 2.2 once more. We choose $i := \zeta_8^2$ and factor $b + di$ in the ring $\mathbb{Z}[i]$ as

$$b + di = \eta'^4 e_0' e'$$

with η' and e_0' consisting only of even prime factors, e_0' not divisible by a nontrivial fourth power and e' odd. Such a factorization is unique up to multiplication by units. With this factorization we have due to Theorem 2.2

$$\left(\frac{b + di}{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}\right)_{4,M} = \mu(\rho, b, c, d)\left(\frac{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}{e'}\right)_{4,M}. \tag{5-13}$$

We claim that

$$\left(\frac{a + b\zeta_8 + c\zeta_8^2 + d\zeta_8^3}{e'}\right)_{4,M} = \left(\frac{a + c\zeta_8^2}{e'}\right)_{4,M} = \left(\frac{a + ci}{e'}\right)_{2,\mathbb{Q}(i)}. \tag{5-14}$$

Indeed, let $\mathfrak{p}$ be a prime in $\mathbb{Z}[i]$ that divides e'. If $\mathfrak{p}$ splits in M, Lemma 3.2 of [Koymans and Milovic 2019a] shows that

$$\left(\frac{a + c\zeta_8^2}{\mathfrak{p}}\right)_{4,M} = \left(\frac{a + ci}{\mathfrak{p}}\right)_{2,\mathbb{Q}(i)}.$$

Suppose instead that $\mathfrak{p}$ remains inert. Then $\mathfrak{p}$ is of degree 1 and $\mathrm{N}\mathfrak{p} \equiv 5 \bmod 8$. Now we apply Lemma 3.3 of [Koymans and Milovic 2019a] to obtain

$$\left(\frac{a + c\zeta_8^2}{\mathfrak{p}}\right)_{4,M} = \left(\frac{a + ci}{\mathfrak{p}}\right)_{2,\mathbb{Q}(i)}.$$

This establishes our claim and hence (5-13). Combining (5-12)–(5-14) acquires the validity of

$$\left(\frac{\sigma(w)}{w}\right)_{4,M} = \mu(\rho, b, c, d)\left(\frac{a+ci}{e'}\right)_{2,\mathbb{Q}(i)}. \tag{5-15}$$

Put

$$f(w,\rho) := \mu(\rho, \rho', b, c, d)\mathbb{1}_{\gcd(w,\sigma\tau(w))=(1)}\left(\frac{tv_2}{a^2+b^2+c^2+d^2}\right).$$

Using (5-6), (5-11) and (5-15), we conclude that

$$\left(\frac{g}{w}\right)_{4,M}\left(\frac{2h}{g}\right) = f(w,\rho)\left(\frac{(b^2+d^2)(2c^2+(b-d)^2)}{v_1}\right)\left(\frac{a+ci}{e'}\right)_{2,\mathbb{Q}(i)}. \tag{5-16}$$

Our hidden cancellation will come from comparing the Jacobi symbols

$$\left(\frac{b^2+d^2}{v_1}\right) \quad \text{and} \quad \left(\frac{a+ci}{e'}\right)_{2,\mathbb{Q}(i)}.$$

Our goal is to show that these two Jacobi symbols are equal up to some easily controlled factors. We can uniquely factor

$$b^2+d^2 = z_1 z_2,$$

where z_1 and z_2 are positive integers satisfying

- $(z_1, z_2) = 1$;
- z_1 odd and squarefree;
- if p is odd and divides z_2, then p^2 also divides z_2.

With this factorization we have

$$\left(\frac{b^2+d^2}{v_1}\right) = \left(\frac{z_1}{v_1}\right)\left(\frac{z_2}{v_1}\right) = \mu(\rho', b, c, d)\left(\frac{v_1}{z_1}\right)\left(\frac{z_2}{v_1}\right).$$

In a similar vein we uniquely factor, up to multiplication by units, e' in $\mathbb{Z}[i]$ as

$$e' = \gamma_1\gamma_2$$

with $(\mathrm{N}\gamma_1, \mathrm{N}\gamma_2) = (1)$, $\mathrm{N}\gamma_1$ squarefree and $\mathrm{N}\gamma_2$ squarefull. The point of this factorization is that $\mathrm{N}\gamma_1 = z_1$. This gives

$$\left(\frac{v_1}{z_1}\right) = \left(\frac{v_1}{\gamma_1}\right)_{2,\mathbb{Q}(i)}.$$

We claim that

$$(tv_2, \gamma_1) = (d, \gamma_1) = (1). \tag{5-17}$$

We clearly have $(t, \gamma_1) = (1)$, so we first show that $(v_2, \gamma_1) = (1)$. Let $\mathfrak{p}$ be an odd prime of $\mathbb{Z}[i]$ above p such that $\mathfrak{p} \mid v_2$ and $\mathfrak{p} \mid \gamma_1$. Then we have $p \mid v_2$ and $\mathrm{N}\mathfrak{p} \mid \mathrm{N}\gamma_1$. However, v_2 is composed entirely of primes dividing $b-d$, while $\mathrm{N}\gamma_1$ divides b^2+d^2. We conclude that p divides both b and d. But then p can not divide γ_1 by construction. We can prove in a similar way that $(d, \gamma_1) = (1)$, thus proving the claim.

From (5-17) we acquire the validity of

$$\begin{aligned}\left(\frac{v_1}{z_1}\right) = \left(\frac{v_1}{\gamma_1}\right)_{2,\mathbb{Q}(i)} &= \mu(b, c, d, t)\left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{v}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \\ &= \mu(b, c, d, t)\left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{a+ci}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{-d(1+i)}{\gamma_1}\right)_{2,\mathbb{Q}(i)} \\ &= \mu(b, c, d, t)\left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{a+ci}{\gamma_1}\right)_{2,\mathbb{Q}(i)},\end{aligned}$$

where we use the identity

$$v = ab - ad + bc + cd \equiv -ad(1+i) + cd(1-i) = -d(1+i)(a+ci) \bmod \gamma_1.$$

We conclude that

$$\begin{aligned}&\left(\frac{b^2+d^2}{v_1}\right)\left(\frac{a+ci}{e'}\right)_{2,\mathbb{Q}(i)} \\ &\quad = \mu(\rho, \rho', b, c, d, t)\left(\frac{z_2}{v_1}\right)\left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{a+ci}{\gamma_2}\right)_{2,\mathbb{Q}(i)} \mathbb{1}_{\gcd(a+ci,\gamma_1)=(1)}. \qquad (5\text{-}18)\end{aligned}$$

Put

$$\begin{aligned}g(w, \rho) := \mu(\rho, \rho', b, c, d, t)&\left(\frac{tv_2}{a^2+b^2+c^2+d^2}\right) \\ &\times \left(\frac{z_2}{v_1}\right)\left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{a+ci}{\gamma_2}\right)_{2,\mathbb{Q}(i)} \mathbb{1}_{\gcd(a+ci,\gamma_1)=\gcd(w,\sigma\tau(w))=(1)}.\end{aligned}$$

After combining (5-16) and (5-18), we get

$$\begin{aligned}\left(\frac{g}{w}\right)_{4,M}\left(\frac{2h}{g}\right) &= g(w, \rho)\left(\frac{2c^2+(b-d)^2}{v_1}\right) \\ &= \mu(\rho, \rho', b, c, d, t)g(w, \rho)\left(\frac{v_1}{2c^2+(b-d)^2}\right).\end{aligned}$$

With this formula we have finally rewritten our symbol in a satisfactory manner; we now return to estimating the sum $A(X, \mathfrak{d}, u_i, \rho)$. We recall the factorization $v = v_1 v_2 t$, where v_1 is an odd, positive integer satisfying $\gcd(v_1, b-d) = 1$, v_2 is an odd integer consisting only of primes dividing $b-d$ and t is positive and only divisible by powers of 2. We further recall that ρ' is the congruence class of v_1 modulo 8.

Let 2^α be the closest integer power of 2 to $X^{1/100}$. We fix b, c, d such that $b-d$ has 2-adic valuation at most $\alpha/2$. If a modulo 2^α is given, we claim that v_{odd} is determined modulo 8, where v_{odd} is the odd part of

$$v = a(b-d) + c(b+d), \tag{5-19}$$

with the exception of $\ll X^{1/200}$ congruence classes ρ'' for a modulo 2^α. Note that, for fixed b, c and d, ρ'' determines v modulo 2^α. If $\alpha \geq 3$, v modulo 2^α determines v_{odd} modulo 8 unless v is divisible by $2^{\alpha-3}$. There are only 8 congruence classes modulo 2^α divisible by $2^{\alpha-3}$. Now take such a congruence class, say ρ'''. But there are $\ll X^{1/200}$ congruence classes ρ'' modulo 2^α with

$$\rho''(b-d) + c(b+d) \equiv \rho''' \bmod 2^\alpha$$

by our assumption that the 2-adic valuation of $b-d$ is at most $\alpha/2$, and our claim follows.

Similarly, we know the value of t with the exception of $\ll X^{1/200}$ congruence classes for a modulo 2^α. We remove all such congruence classes from the sum, which gives an error of size at most $X^{199/200}$. From now on we assume that a does not lie in such a congruence class. For the remaining congruence classes modulo 2^α, we observe that ρ' is determined by v_{odd} modulo 8 together with b, c and d. Hence both ρ' and t are determined by a modulo 2^α.

We would also like to treat v_2 as fixed, and we use a similar technique to achieve this. Once more we fix b, c and d. We assume that

$$\gcd(b-d, bc+cd) \leq \exp((\log X)^{0.25}).$$

We can uniquely factor a positive integer n as x_1x_2, where $\gcd(x_1, x_2) = 1$, $x_1 > 0$ is squarefree and $x_2 > 0$ is squarefull. We call x_1 the squarefree part, and x_2 the squarefull part. We further assume that the squarefull part of $b-d$ is of size at most $\exp((\log X)^{0.25})$. We now factor

$$\gcd(b-d, bc+cd) = \prod_{i=1}^{k} p_i^{f_i}.$$

Define $f_i'(p_i)$ to be the smallest integer such that

$$p_i^{f_i'(p_i)} \geq \exp(2(\log X)^{0.25})$$

and define

$$G := \prod_{i=1}^{k} p_i^{f_i'(p_i)}.$$

Clearly, we have that $\gcd(b-d, bc+cd)$ divides G, since the squarefull part of $b-d$ is of size at most $\exp((\log X)^{0.25})$. If a modulo G is given, we claim that v_2 is determined modulo G with the exception of at most

$$\ll \log X \max_{1\leq i\leq k} \frac{G}{p_i^{f_i'(p_i)/2}}$$

congruence classes ρ'' for a modulo G. Take a prime divisor p_i of $b-d$. If p_i does not divide $bc+cd$, then clearly

$$p_i \nmid a(b-d)+bc+cd,$$

so we have found the p_i valuation of $a(b-d)+bc+cd$. Now suppose that p_i also divides $bc+cd$. Then we know the p_i valuation unless

$$a(b-d)+bc+cd \equiv 0 \bmod p_i^{f_i'(p_i)}.$$

However, we know that the p_i valuation of $b-d$ is at most $f_i'(p_i)/2$. Hence there are at most $p_i^{f_i'(p_i)/2}$ congruence classes for a modulo $p_i^{f_i'(p_i)}$ for which

$$a(b-d)+bc+cd \equiv 0 \bmod p_i^{f_i'(p_i)},$$

and we call such a congruence class forbidden. We let G_i be the set of forbidden congruence classes modulo $p_i^{f_i(p_i)}$. Now we discard all congruence classes ρ'' modulo G for which there exists a prime p_i dividing $\gcd(b-d, bc+cd)$ such that the reduction of ρ'' modulo $p_i^{f_i(p_i)}$ lies in G_i. This proves the claim.

Set

$$m := \operatorname{lcm}(G, z_2, \mathrm{N}\gamma_2, 2^\alpha, 2^{10}). \tag{5-20}$$

Then

$$\left(\frac{tv_2}{a^2+b^2+c^2+d^2}\right)\left(\frac{z_2}{v_1}\right)\left(\frac{v_2}{\gamma_1}\right)_{2,\mathbb{Q}(i)}\left(\frac{a+ci}{\gamma_2}\right)_{2,\mathbb{Q}(i)}$$

depends only on a modulo m, b, c and d. If we write $\beta := b\zeta_8 + c\zeta_8^2 + d\zeta_8^3$, we have the estimate

$$A(X, \mathfrak{d}, u_i, \rho) \ll \sum_{\beta} \sum_{f\in\mathbb{Z}/m\mathbb{Z}} \Bigg| \sum_{\substack{a\in\mathbb{Z}\\ a \text{ sat. } (*)}} \left(\frac{v_1}{2c^2+(b-d)^2}\right) \mathbb{1}_{\gcd(a+ci,\gamma_1)=\gcd(a+\beta,\sigma\tau(a+\beta))=(1)} \Bigg|,$$

where $(*)$ are the simultaneous conditions

$$a+\beta \in u_i\mathcal{D}(X), \quad a+\beta \equiv 0 \bmod \mathfrak{d}, \quad a+\beta \equiv \rho \bmod 2^{10}, \quad a \equiv f \bmod m.$$

Recall that the condition $a+\beta \in u_i\mathcal{D}(X)$ implies $a, b, c, d \ll X^{1/4}$; see Lemma 2.3. We will only consider β satisfying the following five properties:

(i) $z_2, \mathrm{N}\gamma_2 \leq X^{1/200}$.

(ii) $\gcd(b-d, bc+cd) \leq \exp((\log X)^{0.25})$.

(iii) The 2-adic valuation of $b-d$ is at most $\alpha/2$.

(iv) The squarefull part of $b-d$ is of size at most $\exp((\log X)^{0.25})$.

(v) The odd, squarefree part of $2c^2+(b-d)^2$ is at least $X^{99/200}$.

We claim that there are at most

$$\ll \frac{X^{3/4}}{\exp((\log X)^{0.2})}$$

elements β that do not satisfy all five conditions. To do so, we shall bound the number of β that fail a given property in the above list. For (iii) and (iv) this is easily verified. For (v), we use that $2c^2+(b-d)^2$ represents a given integer at most $\ll_\epsilon X^{(1/4)+\epsilon}$ times, and this reduces the problem to an easy counting problem. A similar argument disposes with (i). Finally, for (ii), we count the number of β such that

$$\gcd(b-d, b+d) > \exp\big(\tfrac{1}{2}(\log X)^{0.25}\big) \text{ or } \gcd(b-d, c) > \exp\big(\tfrac{1}{2}(\log X)^{0.25}\big).$$

For those β, we bound the inner sum trivially by $\ll X^{\frac{1}{4}}/m$ inducing an error of size

$$\ll \frac{X}{\exp((\log X)^{0.2})}.$$

For the remaining β, we have $G \ll_\epsilon X^\epsilon$ and hence $m \ll_\epsilon X^{(1/50)+\epsilon}$ by (i) and the definition of m; see (5-20). Note that

$$\mathbb{1}_{\gcd(a+\beta, \sigma\tau(a+\beta))=(1)} = \mathbb{1}_{\gcd(a+\beta, \sigma\tau(\beta)-\beta)=(1)}.$$

We use the Möbius function to detect the coprimality conditions, which yields the upper bound

$$A(X, \mathfrak{d}, u_i, \rho) \ll \sum_{\beta} \sum_{f \in \mathbb{Z}/m\mathbb{Z}} \sum_{\mathfrak{d}_1 \mid \gamma_1} \sum_{\mathfrak{d}_2 \mid \sigma\tau(\beta)-\beta} \bigg| \sum_{\substack{a \in \mathbb{Z} \\ a \text{ sat. } (**)}} \bigg(\frac{v_1}{2c^2+(b-d)^2}\bigg)\bigg|,$$

where $(**)$ are the simultaneous conditions

$$a+\beta \in u_i\mathcal{D}(X), \quad a+\beta \equiv 0 \bmod \mathfrak{d}, \quad a+\beta \equiv \rho \bmod 2^{10}, \quad a \equiv f \bmod m,$$
$$a+ci \equiv 0 \bmod \mathfrak{d}_1, \quad a+\beta \equiv 0 \bmod \mathfrak{d}_2.$$

Define m' to be the smallest positive integer that is divisible by $\operatorname{lcm}(\mathfrak{d}, \mathfrak{d}_1, \mathfrak{d}_2)$. Put

$$M := \operatorname{lcm}(m, m').$$

The congruence conditions for a in $(**)$ are equivalent to at most one congruence condition modulo M. We assume that it is equivalent to exactly one congruence condition modulo M, say F, otherwise the inner sum is empty. Then we have

$$A(X,\mathfrak{d},u_i,\rho) \ll \sum_{\beta} \sum_{f\in\mathbb{Z}/m\mathbb{Z}} \sum_{\mathfrak{d}_1\mid\gamma_1} \sum_{\mathfrak{d}_2\mid\sigma\tau(\beta)-\beta} \Biggl| \sum_{\substack{a\in\mathbb{Z} \\ a+\beta\in u_i\mathcal{D}(X) \\ a\equiv F \bmod M}} \left(\frac{v_1}{2c^2+(b-d)^2}\right)\Biggr|. \tag{5-21}$$

We assume that $M \le X^{1/8}$, since otherwise the trivial bound suffices. Furthermore, for fixed β, the condition $a+\beta \in u_i\mathcal{D}(X)$ means that a runs over $\ll 1$ intervals with endpoints depending on β and u_i. Since $a \ll X^{1/4}$, we know that each interval has length $\ll X^{1/4}$. We have the factorization

$$2c^2+(b-d)^2 = q_1q_2,$$

where q_1 is the odd, squarefree part. We know that $q_2 \ll X^{1/200}$, and we split the sum over a in congruence classes modulo q_2. For fixed b, c and d, the condition $a \equiv F \mod M$ implies that v_1 is a linear function of a with linear term not divisible by q_1 by our assumptions $q_1 \ge X^{99/200}$ and $M \le X^{1/8}$. Indeed, v_2 and t are determined by F, so this follows immediately from (5-3). Hence we may employ the Burgess bound [1963] to (5-21) with $r=2$ and $q=q_1 \ll X^{1/2}$ to prove

$$A(X,\mathfrak{d},u_i,\rho) \ll_\epsilon X^{\frac{31}{32}+\frac{1}{50}+\frac{1}{200}+\epsilon} + X^{\frac{199}{200}} + X^{\frac{15}{16}} + \frac{X}{\exp((\log X)^{0.2})},$$

where the second term accounts for the discarded congruence classes for a, the third term accounts for those M with $M > X^{1/8}$ and the fourth term accounts for the discarded β. This establishes the following proposition.

Proposition 5.1. *We have for all ideals $\mathfrak{d}$ of $\mathbb{Z}[\zeta_8]$*

$$A(X,\mathfrak{d}) \ll \frac{X}{\exp((\log X)^{0.2})}.$$

6. Sums of type II

In (4-5) we defined $[w]_1$ and $[w]_2$. We have the useful decomposition

$$[w] = [w]_1[w]_2.$$

In this section we need to carefully study the multiplicative properties of $[w]$, and we do so by studying the multiplicative properties of $[w]_1$ and $[w]_2$. These properties will then be used to prove cancellation in sums of type II. We start by studying $[w]_1$; our treatment is almost identical to [Koymans and Milovic 2019a]. If w is an odd element of $\mathbb{Z}[\zeta_8]$, we have

$$[w]_1 = \left(\frac{\left(\frac{1}{2}-\frac{1}{2\sqrt{2}}\right)\sigma(w)\sigma\tau(w)}{w}\right)_{4,M} = \left(\frac{(2-\sqrt{2})\sigma(w)\sigma\tau(w)}{w}\right)_{4,M}.$$

Define

$$\gamma_1(w,z) := \left(\frac{\sigma(z)}{w}\right)_{2,M}. \tag{6-1}$$

For the remainder of this section, we use the convention that $\delta(w,z)$ is a function depending only on the congruence classes of w and z modulo 2^{10}; at each occurrence $\delta(w,z)$ may be a different function.

Lemma 6.1. *We have for all odd $w, z \in \mathbb{Z}[\zeta_8]$*

$$[wz]_1 = \delta(w,z)[w]_1[z]_1\gamma_1(w,z)\mathbb{1}_{\gcd(w,\sigma\tau(z))=(1)}.$$

Proof. By definition of $[w]_1$ we have

$$\begin{aligned}[wz]_1 &= \left(\frac{(2-\sqrt{2})\sigma(wz)\sigma\tau(wz)}{wz}\right)_{4,M} \\ &= [w]_1[z]_1\left(\frac{\sigma(z)}{w}\right)_{4,M}\left(\frac{\sigma\tau(z)}{w}\right)_{4,M}\left(\frac{\sigma(w)}{z}\right)_{4,M}\left(\frac{\sigma\tau(w)}{z}\right)_{4,M}.\end{aligned}$$

Since σ fixes i and therefore any quartic residue symbol, Theorem 2.2 yields

$$\begin{aligned}\left(\frac{\sigma(z)}{w}\right)_{4,M}\left(\frac{\sigma(w)}{z}\right)_{4,M} &= \delta(w,z)\left(\frac{\sigma(z)}{w}\right)_{4,M}\left(\frac{z}{\sigma(w)}\right)_{4,M} \\ &= \delta(w,z)\left(\frac{\sigma(z)}{w}\right)_{4,M}\sigma\left(\left(\frac{\sigma(z)}{w}\right)_{4,M}\right) \\ &= \delta(w,z)\left(\frac{\sigma(z)}{w}\right)_{2,M}.\end{aligned}$$

If we do the same computation for $\sigma\tau$, we obtain

$$\left(\frac{\sigma\tau(z)}{w}\right)_{4,M}\left(\frac{\sigma\tau(w)}{z}\right)_{4,M} = \delta(w,z)\mathbb{1}_{\gcd(w,\sigma\tau(z))=(1)},$$

since $\sigma\tau$ does not fix i. This proves the lemma. □

In the next lemma we collect the most important properties of $\gamma_1(w,z)$.

Lemma 6.2. *Let $w, z \in \mathbb{Z}[\zeta_8]$ be odd and define $\gamma_1(w,z)$ as in* (6-1).

(i) *$\gamma_1(w,z)$ is essentially symmetric*

$$\gamma_1(w,z) = \delta(w,z)\gamma_1(z,w).$$

(ii) *$\gamma_1(w,z)$ is multiplicative in both arguments*

$$\gamma_1(w,z_1z_2) = \gamma_1(w,z_1)\gamma_1(w,z_2), \quad \gamma_1(w_1w_2,z) = \gamma_1(w_1,z)\gamma_1(w_2,z).$$

Proof. This is straightforward. □

With this lemma we have completed our study of $[w]_1$ and $\gamma_1(w, z)$. We will now focus on $[w]_2$. Recall that

$$[w]_2 = \left(\frac{2h}{g}\right) = \delta(w)\left(\frac{v}{u}\right).$$

The second representation of $[w]_2$ is very convenient, since it allows us to use earlier work of Milovic [2017a]. Define

$$\gamma_2(w, z) := \left(\frac{\sigma(wz)\sigma\tau(wz)}{w\tau(w)}\right)_{2,K}, \tag{6-2}$$

where $K := \mathbb{Q}(\sqrt{2})$.

Lemma 6.3. *The following formula is valid for all odd $w, z \in \mathbb{Z}[\zeta_8]$:*

$$[wz]_2 = \delta(w, z)[w]_2[z]_2\gamma_2(w, z).$$

Proof. Milovic [2017a, p. 1009] defined the symbol

$$[u + v\sqrt{2}]_3 := \left(\frac{v}{u}\right).$$

Then it is easily seen that $[w]_2 = \delta(w)[w\tau(w)]_3$ and that $w\tau(w)$ is totally positive. Now apply Proposition 8 of [Milovic 2017a]. □

To further our study of $\gamma_2(w, z)$, it will be convenient to define a second function $\mathrm{m}(w)$ by the formula

$$\mathrm{m}(w) := \gamma_2(w, 1) = \left(\frac{\sigma(w)\sigma\tau(w)}{w\tau(w)}\right)_{2,K}.$$

It turns out that $\gamma_2(w, z)$ is neither symmetric nor multiplicative. Instead, it is symmetric and multiplicative twisted by the factor m.

Lemma 6.4. *Let $w, z \in \mathbb{Z}[\zeta_8]$ be odd and define $\gamma_2(w, z)$ as in* (6-2).

(i) *$\gamma_2(w, z)$ is twisted symmetric*

$$\gamma_2(w, z)\gamma_2(z, w) = \mathrm{m}(wz).$$

(ii) *$\gamma_2(w, z)$ is twisted multiplicative in z*

$$\gamma_2(w, z_1z_2) = \mathrm{m}(w)\gamma_2(w, z_1)\gamma_2(w, z_2).$$

Proof. The proof is left to the reader. □

With this out of the way we are ready to tackle the sums of type II. Let $\{\alpha_w\}$ and $\{\beta_z\}$ be sequences of complex numbers of absolute value at most 1 and let ρ and μ be invertible congruence classes modulo 2^{10}. We define

$$B_1(M, N, \rho, \mu) := \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \alpha_w \beta_z [wz],$$

where we suppress the dependence on $\{\alpha_w\}$ and $\{\beta_z\}$. Then we have the following proposition.

Proposition 6.5. *There is an absolute constant $\theta_3 > 0$ such that for all sequences of complex numbers $\{\alpha_w\}$ and $\{\beta_z\}$ of absolute value at most* 1, *all invertible congruence classes ρ and μ modulo 2^{10}*

$$B_1(M, N, \rho, \mu) \ll (M^{-1/24} + N^{-1/24}) MN (\log MN)^{\theta_3}.$$

Proof. We start by expanding $[wz]$ using Lemmas 6.1 and 6.3. We may absorb $[w]_1$, $[w]_2$, $[z]_1$ and $[z]_2$ in the coefficients α_w and β_z. Then it suffices to prove for all sequences of complex numbers $\{\alpha_w\}$ and $\{\beta_z\}$ of absolute value at most 1 and all invertible congruence classes ρ and μ modulo 2^{10} the following estimate:

$$B_2(M, N, \rho, \mu) := \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \alpha_w \beta_z \gamma_1(w, z) \gamma_2(w, z) \mathbb{1}_{\gcd(w, \sigma\tau(z)) = (1)}$$
$$\ll (M^{-1/24} + N^{-1/24}) MN (\log MN)^{\theta_3}.$$

Define

$$\gamma_3(w, z) := \left(\frac{\sigma(z)\sigma\tau(z)}{w\tau(w)} \right)_{2,K},$$

so that we have the factorization $\gamma_2(w, z) = \mathrm{m}(w)\gamma_3(w, z)$. Absorbing $\mathrm{m}(w)$ in α_w and using the identity

$$\gamma_3(w, z) \mathbb{1}_{\gcd(w, \sigma\tau(z)) = (1)} = \gamma_3(w, z),$$

we see that it is enough to establish

$$B_3(M, N, \rho, \mu) := \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \alpha_w \beta_z \gamma_1(w, z) \gamma_3(w, z)$$
$$\ll (M^{-1/24} + N^{-1/24}) MN (\log MN)^{\theta_3}.$$

Theorem 2.1 shows that $\gamma_3(w, z)$ is also essentially symmetric, i.e.,

$$\gamma_3(w, z) = \delta(w, z) \gamma_3(z, w).$$

Due to the symmetry of $\gamma_1(w, z)$, see Lemma 6.2 (i), and the symmetry of $\gamma_3(w, z)$, we may further reduce to the case $N \geq M$. We take $k := 12$ and apply Hölder's inequality with $1 = \frac{k-1}{k} + \frac{1}{k}$ to the w variable to obtain

$$|B_3(M, N, \rho, \mu)|^k \leq \Bigg(\sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} |\alpha_w|^{\frac{k}{k-1}} \Bigg)^{k-1} \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \Bigg| \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \beta_z \gamma_1(w, z) \gamma_3(w, z) \Bigg|^k.$$

The first factor is trivially bounded by $\ll M^{k-1}$ with absolute implied constant. Lemma 6.2 (ii) implies that $\gamma_1(w, z)$ is multiplicative in z and Lemma 6.4 (ii) implies that $\gamma_3(w, z)$ is multiplicative in z. Hence $\gamma_1(w, z)\gamma_3(w, z)$ is multiplicative in z. We conclude that

$$|B_3(M, N, \rho, \mu)|^k \ll M^{k-1} \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \epsilon(w) \sum_z \beta'_z \gamma_1(w, z) \gamma_3(w, z), \tag{6-3}$$

where

$$\epsilon(w) := \left(\frac{\left| \sum_{z \in \mathcal{D}(N),\, z \equiv \mu \bmod 2^{10}} \beta_z \gamma_1(w, z) \gamma_3(w, z) \right|}{\sum_{z \in \mathcal{D}(N),\, z \equiv \mu \bmod 2^{10}} \beta_z \gamma_1(w, z) \gamma_3(w, z)} \right)^k$$

and

$$\beta'_z := \sum_{\substack{z = z_1 \cdots z_k \\ z_1, \ldots, z_k \in \mathcal{D}(N) \\ z_1 \equiv \cdots \equiv z_k \equiv \mu \bmod 2^{10}}} \beta_{z_1} \cdots \beta_{z_k}.$$

We will now study the summation condition for z in the inner sum of (6-3) more carefully. By construction, $\mathcal{D}(N)$ contains exactly eight generators of any principal ideal. Furthermore, there are $\ll N^k$ values of z for which $\beta'_z \neq 0$. Hence we obtain the bound

$$\sum_z (\beta'_z)^2 \ll (\log N)^{\theta_3} N^k$$

for some absolute constant θ_3, since k is fixed. An application of the Cauchy–Schwarz inequality over the z variable yields

$$\Bigg(\sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \epsilon(w) \sum_z \beta'_z \gamma_1(w, z) \gamma_3(w, z) \Bigg)^2 = \Bigg(\sum_z \beta'_z \sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \epsilon(w) \gamma_1(w, z) \gamma_3(w, z) \Bigg)^2$$

$$\ll (\log N)^{\theta_3} N^k \times \sum_{\substack{w_1 \in \mathcal{D}(M) \\ w_1 \equiv \rho \bmod 2^{10}}} \sum_{\substack{w_2 \in \mathcal{D}(M) \\ w_2 \equiv \rho \bmod 2^{10}}} \epsilon(w_1)\overline{\epsilon(w_2)} \sum_z \gamma_1(w_1 w_2, z)\gamma_3(w_1 w_2, z), \tag{6-4}$$

because $\gamma_1(w, z)$ and $\gamma_3(w, z)$ are multiplicative in w. Conveniently, inequality (6-4) remains valid if we extend the sum over z to a larger domain. Let $z_1, \ldots, z_k \in \mathcal{D}(N)$ and write

$$z_i = \sum_{j=1}^{4} a_{ij} \zeta_8^j.$$

Then we have $|a_{ij}| \ll N^{1/4}$. Now define

$$\mathcal{B}(C) := \biggl\{ \sum_{j=1}^{4} a_j \zeta_8^j : a_j \in \mathbb{Z}, |a_j| \le C N^{k/4} \biggr\}.$$

Then, if C is sufficiently large, $\beta'_z \neq 0$ implies $z \in \mathcal{B}(C)$. For this choice of C, we extend the range of summation over z in (6-4) to the set $\mathcal{B}(C)$. We split the sum over z in congruence classes ζ modulo $\mathrm{N}(w_1 w_2)$; we claim that for all odd w

$$\sum_{\zeta \bmod \mathrm{N}(w)} \gamma_1(w, \zeta)\gamma_3(w, \zeta) = 0$$

provided that $\mathrm{N}(w)$ is not squarefull. Substituting the definition of $\gamma_1(w, \zeta)$ and $\gamma_3(w, \zeta)$ gives

$$f(w) := \sum_{\zeta \bmod \mathrm{N}(w)} \gamma_1(w, \zeta)\gamma_3(w, \zeta) = \sum_{\zeta \bmod \mathrm{N}(w)} \left(\frac{\sigma(\zeta)\sigma\tau(\zeta)}{w\tau(w)} \right)_{2,K} \left(\frac{\sigma(\zeta)}{w} \right)_{2,M}.$$

Then a calculation shows that for all odd w and w' satisfying $(\mathrm{N}(w), \mathrm{N}(w')) = 1$,

$$f(ww') = f(w)f(w').$$

Hence, to establish the claim, it is enough to prove that $f(w) = 0$ if w is an odd prime of degree 1. To do so, we start with the identity

$$\left(\frac{\sigma(\zeta)\sigma\tau(\zeta)}{w\tau(w)} \right)_{2,K} = \left(\frac{\sigma(\zeta)\sigma\tau(\zeta)}{w} \right)_{2,M}.$$

Here we rely in an essential way that w is an odd prime of degree 1, so we have an isomorphism of finite fields $O_M/w \cong O_K/w\tau(w)$. We use this to give a simple expression for $f(w)$,

$$f(w) = \sum_{\zeta \bmod \mathrm{N}(w)} \left(\frac{\sigma\tau(\zeta)}{w} \right)_{2,M} \mathbb{1}_{(\sigma(\zeta), w) = (1)},$$

which apart from a nonzero factor is

$$\sum_{\zeta \bmod \sigma(w)\sigma\tau(w)} \left(\frac{\sigma\tau(\zeta)}{w}\right)_{2,M} \mathbb{1}_{(\sigma(\zeta),w)=(1)}$$
$$= \sum_{\zeta \bmod \sigma\tau(w)} \left(\frac{\sigma\tau(\zeta)}{w}\right)_{2,M} \sum_{\zeta \bmod \sigma(w)} \mathbb{1}_{(\sigma(\zeta),w)=(1)} = 0.$$

Note that $\sigma(w)$ and $\sigma\tau(w)$ are coprime, so we are allowed to expand the sum over $\sigma(w)\sigma\tau(w)$ as the product of the two sums over $\sigma(w)$ and $\sigma\tau(w)$. With the claim established, we can give an upper bound for the sum over $z \in \mathcal{B}(C)$

$$\sum_{z\in\mathcal{B}(C)} \gamma_1(w_1w_2, z)\gamma_3(w_1w_2, z) \ll \begin{cases} N^k & \text{if } \mathrm{N}(w_1w_2) \text{ is squarefull,} \\ \sum_{i=1}^4 M^{2i} N^{k(1-\frac{1}{4}i)} & \text{otherwise,} \end{cases}$$

where the second bound uses the claim and $\mathrm{N}(w_1w_2) \le M^2$. Because of our choice of k and $N \ge M$, we can simplify the second bound to $M^2 N^{\frac{3}{4}k}$. Equations (6-3), (6-4) and the above bound acquire the validity of

$$|B_3(M, N, \rho, \mu)|^{2k} \ll (\log N)^{\theta_3} M^{2k-2} N^k (M \cdot N^k + M^2 \cdot M^2 N^{\frac{3}{4}k})$$
$$\ll (\log N)^{\theta_3} (M^{2k-1} \cdot N^k + M^{2k+2} \cdot N^{\frac{7}{4}k}).$$

Since the first term above dominates the second term due to our choice of k and $N \ge M$, the proof of the proposition is complete. □

Having dealt with sums of type II for the symbol $[wz]$, we now turn to sums of type II with $a_{\mathfrak{mn}}$. For sequences of complex numbers $\{\alpha_{\mathfrak{m}}\}$ and $\{\beta_{\mathfrak{n}}\}$ of absolute value at most 1 we defined in Section 3 the following sum:

$$B(M, N) = \sum_{\mathrm{N}\mathfrak{m} \le M} \sum_{\mathrm{N}\mathfrak{n} \le N} \alpha_{\mathfrak{m}} \beta_{\mathfrak{n}} a_{\mathfrak{mn}}.$$

Proposition 6.6. *There is an absolute constant $\theta_3 > 0$ such that for all sequences of complex numbers $\{\alpha_{\mathfrak{m}}\}$ and $\{\beta_{\mathfrak{n}}\}$ of absolute value at most* 1,

$$B(M, N) \ll (M^{-1/24} + N^{-1/24}) MN (\log MN)^{\theta_3}.$$

Proof. By picking generators for $\mathfrak{m}$ and $\mathfrak{n}$ we obtain the identity

$$B(M, N) = \sum_{\mathrm{N}\mathfrak{m} \le M} \sum_{\mathrm{N}\mathfrak{n} \le N} \alpha_{\mathfrak{m}} \beta_{\mathfrak{n}} a_{\mathfrak{mn}} = \frac{1}{64} \sum_{w\in\mathcal{D}(M)} \sum_{z\in\mathcal{D}(N)} \alpha_w \beta_z a_{(wz)}.$$

We split the sum $B(M, N)$ in congruence classes modulo 2^{10}. We need only consider invertible congruence classes, since otherwise $a_{wz} = 0$ by definition. Furthermore, condition (4-4) depends only on g modulo 4, which is in turn determined by w modulo 4. Therefore, it suffices to bound the sum

$$\sum_{\substack{w \in \mathcal{D}(M) \\ w \equiv \rho \bmod 2^{10}}} \sum_{\substack{z \in \mathcal{D}(N) \\ z \equiv \mu \bmod 2^{10}}} \alpha_w \beta_z \big([wz] + [\epsilon wz] + [\epsilon^2 wz] + [\epsilon^3 wz]\big),$$

where ρ and μ are invertible congruence classes modulo 2^{10} such that $g \equiv 1 \bmod 4$. From Lemmas 6.1 and 6.3 we deduce that

$$[\epsilon wz] = \delta(w, z)[\epsilon][wz].$$

Now apply Proposition 6.5. □

Acknowledgements

I am very grateful to Djordjo Milovic for his support during this project. I would also like to thank Jan-Hendrik Evertse for proofreading.

References

[Alberts 2016] B. Alberts, "Cohen–Lenstra moments for some nonabelian groups", preprint, 2016. arXiv

[Alberts and Klys 2016] B. Alberts and J. Klys, "The distribution of H_8-extensions of quadratic fields", preprint, 2016. arXiv

[Bruin and Hemenway 2013] N. Bruin and B. Hemenway, "On congruent primes and class numbers of imaginary quadratic fields", *Acta Arith.* **159**:1 (2013), 63–87. MR Zbl

[Burgess 1963] D. A. Burgess, "On character sums and L-series, II", *Proc. Lond. Math. Soc.* (3) **13** (1963), 524–536. MR Zbl

[Cohn and Lagarias 1983] H. Cohn and J. C. Lagarias, "On the existence of fields governing the 2-invariants of the classgroup of $\mathbb{Q}(\sqrt{dp})$ as p varies", *Math. Comp.* **41**:164 (1983), 711–730. MR Zbl

[Cohn and Lagarias 1984] H. Cohn and J. C. Lagarias, "Is there a density for the set of primes p such that the class number of $\mathbb{Q}(\sqrt{-p})$ is divisible by 16?", pp. 257–280 in *Topics in classical number theory, I* (Budapest, 1981), edited by G. Halász, Colloq. Math. Soc. János Bolyai **34**, North-Holland, Amsterdam, 1984. MR Zbl

[Davenport and Heilbronn 1971] H. Davenport and H. Heilbronn, "On the density of discriminants of cubic fields, II", *Proc. Roy. Soc. Lond. Ser. A* **322**:1551 (1971), 405–420. MR Zbl

[Fouvry and Klüners 2006] É. Fouvry and J. Klüners, "Cohen–Lenstra heuristics of quadratic number fields", pp. 40–55 in *Algorithmic number theory* (Berlin, 2006), edited by F. Hess et al., Lect. Notes Comput. Sci. **4076**, Springer, 2006. MR Zbl

[Fouvry and Klüners 2007] É. Fouvry and J. Klüners, "On the 4-rank of class groups of quadratic number fields", *Invent. Math.* **167**:3 (2007), 455–513. MR Zbl

[Friedlander and Iwaniec 1998] J. Friedlander and H. Iwaniec, "The polynomial $X^2 + Y^4$ captures its primes", *Ann. of Math.* (2) **148**:3 (1998), 945–1040. MR Zbl

[Friedlander et al. 2013] J. B. Friedlander, H. Iwaniec, B. Mazur, and K. Rubin, "The spin of prime ideals", *Invent. Math.* **193**:3 (2013), 697–749. Correction in **202**:2 (2015), 923–925. MR Zbl

[Gerth 1984] F. Gerth, III, “The 4-class ranks of quadratic fields”, *Invent. Math.* **77**:3 (1984), 489–515. MR Zbl

[Hasse 1969] H. Hasse, “Über die Klassenzahl des Körpers $P(\sqrt{-2p})$ mit einer Primzahl $p \neq 2$”, *J. Number Theory* **1** (1969), 231–234. MR Zbl

[Heath-Brown 1994] D. R. Heath-Brown, “The size of Selmer groups for the congruent number problem, II”, *Invent. Math.* **118**:2 (1994), 331–370. MR Zbl

[Klys 2017] J. Klys, “Moments of unramified 2-group extensions of quadratic fields”, preprint, 2017. arXiv

[Koymans and Milovic 2019a] P. Koymans and D. Milovic, “On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-2p})$ for primes $p \equiv 1 \mod 4$”, *Int. Math. Res. Not. IMRN* **2019**:23 (2019), 7406–7427. MR Zbl

[Koymans and Milovic 2019b] P. Koymans and D. Z. Milovic, “Spins of prime ideals and the negative Pell equation $x^2 - 2py^2 = -1$”, *Compos. Math.* **155**:1 (2019), 100–125. MR Zbl

[Lemmermeyer 2000] F. Lemmermeyer, *Reciprocity laws: from Euler to Eisenstein*, Springer, 2000. MR Zbl

[Leonard and Williams 1982] P. A. Leonard and K. S. Williams, “On the divisibility of the class numbers of $Q(\sqrt{-p})$ and $Q(\sqrt{-2p})$ by 16”, *Canad. Math. Bull.* **25**:2 (1982), 200–206. MR Zbl

[Milovic 2017a] D. Milovic, “On the 16-rank of class groups of $\mathbb{Q}(\sqrt{-8p})$ for $p \equiv -1 \mod 4$”, *Geom. Funct. Anal.* **27**:4 (2017), 973–1016. MR Zbl

[Milovic 2017b] D. Z. Milovic, “The infinitude of $\mathbb{Q}(\sqrt{-p})$ with class number divisible by 16”, *Acta Arith.* **178**:3 (2017), 201–233. MR Zbl

[Rédei 1934] L. Rédei, “Arithmetischer Beweis des Satzes über die Anzahl der durch vier teilbaren Invarianten der absoluten Klassengruppe im quadratischen Zahlkörper”, *J. Reine Angew. Math.* **171** (1934), 55–60. MR Zbl

[Smith 2016] A. Smith, “Governing fields and statistics for 4-Selmer groups and 8-class groups”, preprint, 2016. arXiv

[Smith 2017] A. Smith, “2^∞-Selmer groups, 2^∞-class groups, and Goldfeld’s conjecture”, preprint, 2017. arXiv

[Stevenhagen 1988] P. Stevenhagen, *Class groups and governing fields*, Ph.D. thesis, University of California, Berkeley, 1988.

[Stevenhagen 1993] P. Stevenhagen, “Divisibility by 2-powers of certain quadratic class numbers”, *J. Number Theory* **43**:1 (1993), 1–19. MR Zbl

[Wood 2019] M. M. Wood, “Nonabelian Cohen–Lenstra moments”, *Duke Math. J.* **168**:3 (2019), 377–427. MR Zbl

Communicated by Peter Sarnak
Received 2018-09-19 Revised 2019-08-14 Accepted 2019-09-12

p.h.koymans@math.leidenuniv.nl
Mathematisch Instituut, Leiden University, Leiden, Netherlands

ALGEBRA AND NUMBER THEORY 14:1 (2020)
dx.doi.org/10.2140/ant.2020.14.67

Supersingular Hecke modules as Galois representations

Elmar Grosse-Klönne

Let F be a local field of mixed characteristic $(0, p)$, let k be a finite extension of its residue field, let $\mathcal{H}$ be the pro-p-Iwahori Hecke k-algebra attached to $\mathrm{GL}_{d+1}(F)$ for some $d \geq 1$. We construct an exact and fully faithful functor from the category of supersingular $\mathcal{H}$-modules to the category of $\mathrm{Gal}(\bar{F}/F)$-representations over k. More generally, for a certain k-algebra $\mathcal{H}^\sharp$ surjecting onto $\mathcal{H}$ we define the notion of $\sharp$-supersingular modules and construct an exact and fully faithful functor from the category of $\sharp$-supersingular $\mathcal{H}^\sharp$-modules to the category of $\mathrm{Gal}(\bar{F}/F)$-representations over k.

Introduction

Let F be a local field of mixed characteristic $(0, p)$, let $\pi \in \mathcal{O}_F$ be a uniformizer, let k be a finite extension of the residue field $\mathbb{F}_q$ of F. Let $d \in \mathbb{N}$. An important line of current research in number theory is concerned with relating smooth representations of $G = \mathrm{GL}_{d+1}(F)$ over k with finite dimensional representations of $\mathrm{Gal}(\overline{F}/F)$ over k.

At present, the smooth representation theory of G is understood only up to identifying, constructing and describing the still elusive *supercuspidal* representations of G, or equivalently, the *supersingular* representations of G. An important role in better understanding this theory is played by the module theory of the pro-p-Iwahori Hecke k-algebra $\mathcal{H}$ attached to G and a pro-p-Iwahori subgroup I_0 in G. There is a notion of supersingularity for $\mathcal{H}$-modules which, in contrast to that of supersingularity for G-representations, is transparent and concrete. The notions are compatible in the following sense: at least after replacing k by an algebraically closed extension field, a smooth admissible irreducible G-representation V is supersingular if and only if its space of I_0-invariants V^{I_0} (which carries a natural

MSC2010: 11F85.
Keywords: pro-p Iwahori Hecke algebra, supersingular module, Galois representation, (φ, Γ)-module.

action by $\mathcal{H}$) is supersingular if and only if V^{I_0} admits a supersingular subquotient; see [Ollivier and Vignéras 2018]. It is true that the functor $V \mapsto V^{I_0}$ from G-representations to $\mathcal{H}$-modules often looses information. But the potential of taking into account also its higher derived functors, which again yield (complexes of) $\mathcal{H}$-modules, has been barely explored so far.

The purpose of the present paper is to explain a method for converting (supersingular) $\mathcal{H}$-modules into $\mathrm{Gal}(\overline{F}/F)$-representations over k.

For $F = \mathbb{Q}_p$ we had constructed in [Grosse-Klönne 2016] an exact functor from finite dimensional $\mathcal{H}$-modules to $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-representations over k. The construction was inspired by Colmez's functor from $\mathrm{GL}_2(\mathbb{Q}_p)$-representations to $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$-representations. It was geometric-combinatorial in that it invoked coefficient systems on the Bruhat Tits building of $\mathrm{GL}_n(\mathbb{Q}_p)$. Unfortunately, we see no way to generalize this geometric-combinatorial method to arbitrary finite extensions of F of $\mathbb{Q}_p$. However, when trying to extract its "algebraic essence", we found that the functor indeed admits a generalization to any F, albeit now taking on an entirely algebraic and concrete shape. But in fact, it is this concreteness which allows us to not only investigate its behavior on irreducible objects, but also to prove that it accurately preserves extension structures. In this way, even for $F = \mathbb{Q}_p$ we significantly improve on our previous work [Grosse-Klönne 2016].

Let $\mathrm{Rep}(\mathrm{Gal}(\overline{F}/F))$ denote the category whose objects are projective limits of finite dimensional $\mathrm{Gal}(\overline{F}/F)$-representations over k. Let $\mathrm{Mod}_{ss}(\mathcal{H})$ denote the category of supersingular $\mathcal{H}$-modules which are inductive limits of their finite dimensional submodules.

Theorem A. *There is an exact and fully faithful functor*

$$\mathrm{Mod}_{ss}(\mathcal{H}) \to \mathrm{Rep}(\mathrm{Gal}(\overline{F}/F)), \quad M \mapsto V(M).$$

We have $\dim_k(M) = \dim_k(V(M))$ *for any* $M \in \mathrm{Mod}_{ss}(\mathcal{H})$.

The radical elimination of the group G (and its building) from our approach allows us to improve Theorem A further as follows. We construct k-algebras $\mathcal{H}^{\sharp\sharp}$ and $\mathcal{H}^{\sharp}$ by looking at a certain small set of distinguished generators of $\mathcal{H}$ and by relaxing resp. omitting some of the usual (braid) relations between them. In this way we get a chain of surjective k-algebra morphisms $\mathcal{H}^{\sharp\sharp} \to \mathcal{H}^{\sharp} \to \mathcal{H}$. There is a again a notion of supersingularity for $\mathcal{H}^{\sharp\sharp}$-modules and for $\mathcal{H}^{\sharp}$-modules (which are inductive limits of their finite dimensional submodules; we assume this for all $\mathcal{H}^{\sharp\sharp}$-, resp. $\mathcal{H}^{\sharp}$-, resp. $\mathcal{H}$-modules appearing in this paper). The simple supersingular modules are the same for $\mathcal{H}^{\sharp\sharp}$, for $\mathcal{H}^{\sharp}$ and for $\mathcal{H}$, but there are more extensions between them in the category of $\mathcal{H}^{\sharp\sharp}$-modules, resp. of $\mathcal{H}^{\sharp}$-modules, than in the category of $\mathcal{H}$-modules. A particular useful category $\mathrm{Mod}_{ss}(\mathcal{H}^{\sharp})$ is formed by what we call $\sharp$-supersingular $\mathcal{H}^{\sharp}$-modules. It contains the category of supersingular $\mathcal{H}$-modules as a full subcategory (but is larger). Now it turns out that the above functor is actually defined on the category of supersingular $\mathcal{H}^{\sharp\sharp}$-modules, and again with $\dim_k(M) = \dim_k(V(M))$ for any M. When restricting to $\mathrm{Mod}_{ss}(\mathcal{H}^{\sharp})$ we furthermore get:

Theorem A$^{\#}$. *There is an exact and fully faithful functor*

$$\mathrm{Mod}_{ss}(\mathcal{H}^{\sharp}) \to \mathrm{Rep}(\mathrm{Gal}(\overline{F}/F)), \quad M \mapsto V(M).$$

We do not know if the k-algebra $\mathcal{H}^\sharp$ admits a group theoretic interpretation, as does the double coset algebra $\mathcal{H} \cong k[I_0\backslash G/I_0]$. However, already from the Galois representation theoretic point of view we think that the additional effort taken in proving Theorem A$^\#$ (rather than just Theorem A) is justified, since in this way we identify an even larger abelian subcategory of $\mathrm{Rep}(\mathrm{Gal}(\overline{F}/F))$ as a (supersingular) module category of a very concretely given k-algebra. In fact, the additional effort is mostly notational.

We define a standard supersingular $\mathcal{H}$-module to be an $\mathcal{H}$-module induced from a supersingular character of a certain subalgebra $\mathcal{H}_{\mathrm{aff}}$ of $\mathcal{H}$ with $[\mathcal{H} : \mathcal{H}_{\mathrm{aff}}] = d + 1$. Each simple supersingular $\mathcal{H}$-module is a subquotient of a standard supersingular $\mathcal{H}$-module. We also define the notion of a $(d+1)$-dimensional standard cyclic $\mathrm{Gal}(\overline{F}/F)$-representation; in particular, each irreducible $\mathrm{Gal}(\overline{F}/F)$-representation of dimension $d+1$ is a $(d+1)$-dimensional standard cyclic $\mathrm{Gal}(\overline{F}/F)$-representation.

Theorem B. *The functor $M \mapsto V(M)$ induces a bijection between standard supersingular $\mathcal{H}$-modules and $(d+1)$-dimensional standard cyclic $\mathrm{Gal}(\overline{F}/F)$-representations. M is irreducible if and only if $V(M)$ is irreducible.*[1]

However, we emphasize that it is rather the much deeper Theorem A (and A$^\#$) which proves that supersingular modules are of a strong inherent arithmetic nature.

In Section 5E we gather some generic statements which come close to describing the image of the functor $M \mapsto V(M)$.

Let us now indicate the main features of the construction of the functor. We fix once and for all a Lubin–Tate group for F. More precisely, as this simplifies many formulae, we work with the Lubin–Tate group associated with the Frobenius power series $\Phi(t) = t^q + \pi t$. On the k-algebra $k[\![t]\!][\varphi]$ with commutation relation $\varphi \cdot t = t^q \cdot \varphi$ we let $\Gamma = \mathcal{O}_F^\times$ act by $\gamma \cdot \varphi = \gamma'\varphi$ and $\gamma \cdot t = [\gamma]_\Phi(t)$, where $[\gamma]_\Phi(t) \in k[\![t]\!]$ describes multiplication with γ with respect to Φ and where $\gamma' \in k^\times$ means the image of $\gamma \in \Gamma$ in $k^\times$. We view a supersingular $\mathcal{H}^{\sharp\sharp}$-module (or $\mathcal{H}^\sharp$-module, or $\mathcal{H}$-module) M as a $k[\![t]\!]$-module by means of $t|_M = 0$. In $k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M$ we then use the $\mathcal{H}$-action on M to define a certain submodule $\nabla(M)$ by giving very explicitly a certain number of generators of it. This is done in such a way that $\Delta(M) = k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M/\nabla(M)$ naturally receives an action by Γ and is a torsion $k[\![t]\!]$-module. A very general construction then allows us to endow $\Delta(M)^* \otimes_{k[\![t]\!]} k((t))$ with the structure of a (φ, Γ)-module over $k((t))$. The notion of a (φ, Γ)-module over $k((t))$ with respect to the chosen Lubin–Tate group Φ is explained in full detail in the book [Schneider 2017], where it is also explained that this category is equivalent with the category of representations of $\mathrm{Gal}(\overline{F}/F)$ over k.

It was pointed out by Cédric Pépin that the syntax of the functor $M \mapsto V(M)$ bears strong resemblance with that of Fontaine's various functors (using "big rings").

One may wonder which of our results remain valid if the coefficient field k is allowed to be a more general field k containing $\mathbb{F}_q$, i.e., not necessarily finite. First, this finiteness is invoked for the equivalence of categories between Galois representations and (φ, Γ)-modules. But it is also invoked in the proofs of

[1] A numerical version (i.e., comparing cardinalities) of Theorem B was known for quite some time, due to work of Ollivier and Vignéras [2005].

Proposition 3.3 (our main result in Section 3 on recovering a supersingular $\mathcal{H}^\sharp$-module from subquotients) and of Theorem 5.10 (on recovering M from $\Delta(M)$).

In Section 2B we list some automorphisms of $\mathcal{H}$ (and of $\mathcal{H}^\sharp$ and $\mathcal{H}^{\sharp\sharp}$). They induce autoequivalences of the category of supersingular $\mathcal{H}$-modules;[2] thus, precomposing them to $M \mapsto V(M)$ we get more functors satisfying Theorems A, A$^\#$ and B.

We end this paper somewhat speculatively by discussing assignments of $\mathrm{Gal}(\overline{F}/F)$-representations to supersingular G-representations. The functor $M \mapsto V(M)$ invites us to search for meaningful assignments of (complexes of) supersingular $\mathcal{H}$-modules to supersingular G-representations Y. First we suggest studying the left derived functor of the functor taking Y to the maximal supersingular $\mathcal{H}$-submodule of Y^{I_0}. This entails working in derived categories and appears to be the most natural approach. Nevertheless, as a variation of this theme we then suggest an *exact* functor from (suitably filtered) G-representations to supersingular $\mathcal{H}$-modules. It builds on a general procedure of turning complexes of $\mathcal{H}$-modules into new $\mathcal{H}$-modules, applied here to complexes arising from E_1-spectral sequences attached to the said left derived functor.

Apparently, the constructions and results of the present paper call for generalizations into various directions. We mention here just the obvious question of what happens if the pro-p-Iwahori Hecke algebra $\mathcal{H}$ attached to $G = \mathrm{GL}_{d+1}(F)$ is replaced by pro-p-Iwahori Hecke algebras $\mathcal{H}$ attached to other p-adic reductive groups G. In extrapolation of what we did here, the general Langlands philosophy suggests searching for a functor from $\mathcal{H}$-modules to Galois representations such that in some way the algebraic k-group with root datum dual to that of G shows up on the Galois side — just as it does here in Theorem B. In a subsequent paper we will propose such a functor. However, in its formal shape it will *not* precisely specialize to the functor discussed here if $G = \mathrm{GL}_{d+1}(F)$,[3] and Theorem A will *not* be a special case of what we will then prove for general G.

Notations. Let $F/\mathbb{Q}_p$ be a finite field extension. Let $\mathbb{F}_q$ be the residue field of F (with q elements). Let π be a uniformizer in $\mathcal{O}_F$. Let k be a finite field extension of $\mathbb{F}_q$.

As explained in [Schneider 2017, Proposition 1.3.4], attached to the Frobenius (or Lubin–Tate) formal power series $\Phi(t) = \pi t + t^q$ is associated a commutative formal group law (the associated Lubin–Tate (formal) group law) $F_\Phi(X, Y)$ over $\mathcal{O}_F$ such that $\Phi(t) \in \mathrm{End}_{\mathcal{O}_F}(F_\Phi(X, Y))$. There is a unique injective homomorphism of rings

$$\mathcal{O}_F \to \mathrm{End}_{\mathcal{O}_F}(F_\Phi(X, Y)), \quad a \mapsto [a]_\Phi(t)$$

such that $\Phi(t) = [\pi]_\Phi(t)$, see [Schneider 2017, Proposition 1.3.6], where we recall that, by definition,

$$\mathrm{End}_{\mathcal{O}_F}(F_\Phi(X, Y)) = \{h \in \mathcal{O}_F[\![t]\!];\ h(0) = 0 \text{ and } h(F_\Phi(X, Y)) = F_\Phi(h(X), h(Y))\}.$$

Lemma 0.1. *Assume that $F \neq \mathbb{Q}_2$. Writing $[a]_\Phi(t) = at + \sum_{i \geq 2} a_i t^i$ (with $a_i \in \mathcal{O}_F$), we have $a_i = 0$ whenever $i - 1 \notin (q-1)\mathbb{N}$. If $a^{q-1} = 1$ we even have $a_i = 0$ for all $i \geq 2$.*

[2] But this is not so evident, if true at all, for the category of $\sharp$-supersingular $\mathcal{H}^\sharp$-modules

[3] But of course, it will be closely related

Proof. As $\Phi(t) = \pi t + t^q$, the power series $[a]_\Phi(t) = at + \sum_{i\geq 2} a_i t^i$ is characterized by the formula

$$\pi[a]_\Phi(t) + ([a]_\Phi(t))^q = [a]_\Phi(\pi t + t^q).$$

If $a^{q-1} = 1$ we see that $[a]_\Phi(t) = at$ satisfies this formula. Given a general a, consider the equalities $[a]_\Phi([b]_\Phi(t)) = [b]_\Phi([a]_\Phi(t))$ for all $b \in \mathcal{O}_F$ with $b^{q-1} = 1$. Since we know $[b]_\Phi(t) = bt$, and since $F \neq \mathbb{Q}_2$ implies the existence of primitive such $b's$ different from 1, we indeed obtain $a_i = 0$ whenever $i - 1 \notin (q-1)\mathbb{N}$. □

1. Lubin–Tate (φ, Γ)-modules

In the first two subsections we transpose some constructions and results from the theory of cyclotomic (φ, Γ)-modules over k (i.e., where $F = \mathbb{Q}_p$ and where the underlying Lubin–Tate group is $\mathbb{G}_m$) to the context of (φ, Γ)-modules over k with respect to the Lubin–Tate group attached to $\Phi(t) = \pi t + t^q$ (with arbitrary F). Namely, we define an exact functor from admissible (torsion) $k[\![t]\!]$-modules with commuting semilinear actions by $\Gamma = \mathcal{O}_F^\times$ and φ to étale (φ, Γ)-modules over k. The former category is closely related to that of ψ-stable lattices in étale (φ, Γ)-modules $\boldsymbol{D}$, and we are lead to transpose some of Colmez's constructions [2010] involving the ψ-stable lattices $\boldsymbol{D}^\natural$ and $\boldsymbol{D}^\sharp$ to our context. One difference is that in our context the ψ-operator on $k((t))$ does not satisfy $\psi(1) = 1$, but this necessitates only minor modifications.

We then identify a category of admissible (torsion) $k[\![t]\!]$-modules with actions by Γ and φ on which the above functor is fully faithful.

1A. (φ, Γ)-*modules and torsion* $k[\![t]\!]$-*modules.* Put $\Phi(t) = \pi t + t^q$. Put $\Gamma = \mathcal{O}_F^\times$. The formula $\gamma \cdot t = [\gamma]_\Phi(t)$ with $\gamma \in \Gamma$ defines an action of Γ by k-algebra automorphisms on $k[\![t]\!]$ and on $k((t))$. Consider the k-algebra

$$\mathfrak{O} = k[\![t]\!][\varphi, \Gamma]$$

with commutation rules given by

$$\gamma\varphi = \varphi\gamma, \quad \gamma t = [\gamma]_\Phi(t)\gamma, \quad \varphi t = t^q\varphi$$

for $\gamma \in \Gamma$. (Here we read $[\gamma]_\Phi(t)\gamma = ([\gamma]_\Phi(t))\gamma$.)[4]

Definition. A ψ-operator on $k[\![t]\!]$ is a k-linear map $\psi : k[\![t]\!] \to k[\![t]\!]$ such that $\psi(\gamma \cdot t) = \gamma \cdot (\psi(t))$ for all $\gamma \in \Gamma$ and such that the following holds true:[5] if we view φ as acting on $k[\![t]\!]$, then

$$\psi(\varphi(a)x) = a\psi(x) \quad \text{for } a, x \in k[\![t]\!]. \tag{1}$$

Lemma 1.1. *There is a surjective ψ-operator on $k[\![t]\!]$ which extends to a surjective k-linear operator $\psi = \psi_{k((t))}$ on $k((t))$ satisfying formula* (1) *analogously.*

[4] As $t^q = \Phi(t) = [\pi]_\Phi(t)$ in $k[\![t]\!]$ one may also think of $\mathfrak{O}$ as $\mathfrak{O} = k[\![t]\!][\mathcal{O}_F - \{0\}]$ with commutation rules $at = [a]_\Phi(t)a$ for all $a \in \mathcal{O}_F - \{0\}$.

[5] We do not require $\psi(1) = 1$.

We may choose $\psi_{k((t))}$ *on* $k((t))$ *such that for* $m \in \mathbb{Z}$ *and* $0 \leq i \leq q-1$ *we have*[6]

$$\psi_{k((t))}(t^{mq+i}) = \begin{cases} \frac{q}{\pi} t^m & i = 0, \\ 0 & 1 \leq i \leq q-2, \\ t^m & i = q-1. \end{cases} \tag{2}$$

Proof. This is explained in [Grosse-Klönne 2019]; it relies on [Schneider and Venjakob 2016, Section 3]. □

In the following we fix the surjective ψ-operator ψ on $k[[t]]$ satisfying formula (2). We extend it to a k-linear operator $\psi = \psi_{k((t))}$ on $k((t))$ as in Lemma 1.1.

Definition. An étale (φ, Γ)-module over $k((t))$ is an $\mathfrak{O} \otimes_{k[[t]]} k((t))$-module $\boldsymbol{D}$ which is finite dimensional over $k((t))$ such that the $k((t))$-linearized structure map

$$\mathrm{id} \otimes \varphi : k((t)) \otimes_{\varphi, k((t))} \boldsymbol{D} \xrightarrow{\cong} \boldsymbol{D}$$

is bijective. We define $\mathrm{Mod}^{\mathrm{et}}(k((t)))$ to be the category of étale (φ, Γ)-module over $k((t))$.

Theorem 1.2 (Fontaine, Kisin–Ren, Schneider). *There is an equivalence between* $\mathrm{Mod}^{\mathrm{et}}(k((t)))$ *and the category of continuous representations of* $\mathrm{Gal}(\overline{F}/F)$ *on finite dimensional* k*-vector spaces.*

Proof. For $F = \mathbb{Q}_p$ and the Frobenius power series $(1+t)^p - 1$ (instead of $\Phi(t) = \pi t + t^q$) this is a theorem of Fontaine, see paragraph 1.2 in [Fontaine 1990]. The analog of the theorem (for an arbitrary Frobenius power series) for a coefficient field of characteristic 0 (hence not k) is due to Kisin and Ren [2009]. A detailed proof of the theorem stated here can be found in Schneider's book [2017]. □

Definition. A torsion $k[[t]]$-module Δ is called admissible if

$$\Delta[t] = \{x \in \Delta; tx = 0\}$$

is a finite dimensional k-vector space.

We remark that admissible $k[[t]]$-modules on which t acts surjectively are precisely the Pontrjagin duals of finitely generated torsion free, and hence free $k[[t]]$-modules.

Definition. $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ is the category of $\mathfrak{O}$-modules which are finitely generated over $k[[t]][\varphi]$ and admissible (in particular: torsion) over $k[[t]]$.

Lemma 1.3. *The categories* $\mathrm{Mod}^{\mathrm{et}}(k((t)))$ *and* $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ *are abelian.*

Proof. An $\mathfrak{O} \otimes_{k[[t]]} k((t))$-module subquotient of an étale (φ, Γ)-module is again an étale (φ, Γ)-module: to see that the étaleness condition (the bijectivity of $\mathrm{id} \otimes \varphi$) is preserved under passing to such subquotients, just notice that it is equivalent with saying that the matrix of φ in an arbitrary $k((t))$-basis is invertible. Thus, $\mathrm{Mod}^{\mathrm{et}}(k((t)))$ is abelian. (Of course, one could also point to Theorem 1.2.)

An $\mathfrak{O}$-module subquotient of an object in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ is again an object in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$: this is shown in [Emerton 2008, Proposition 3.3]. Thus, $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ is abelian. □

[6] Notice that $\frac{q}{\pi} = 0$ (in k) if $F \neq \mathbb{Q}_p$.

Definition. For a k-vector space Δ we write $\Delta^* = \mathrm{Hom}_k(\Delta, k)$ (algebraic dual). For a $k[[t]]$-module Δ we endow Δ^* with a $k[[t]]$-action by putting

$$(S \cdot f)(\delta) = f(S\delta)$$

for $S \in k[[t]]$, $f \in \Delta^*$, $\delta \in \Delta$. If Δ even carries a $k[[t]][\Gamma]$-module structure then also Δ^* receives one, with $\gamma \in \Gamma$ acting as

$$(\gamma \cdot f)(\delta) = f(\gamma^{-1}\delta)$$

for $\gamma \in \Gamma$, $f \in \Delta^*$, $\delta \in \Delta$.

Proposition 1.4. *For $\Delta \in \mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ there is a natural structure of an étale (φ, Γ)-module on $\Delta^* \otimes_{k[[t]]} k((t))$. The contravariant functor*

$$\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O}) \to \mathrm{Mod}^{\mathrm{et}}(k((t))), \quad \Delta \mapsto \Delta^* \otimes_{k[[t]]} k((t)) \tag{3}$$

is exact.

Proof. The map $\mathrm{id} \otimes \varphi : k[[t]] \otimes_{\varphi, k[[t]]} \Delta \to \Delta$ gives rise to the $k[[t]]$-linear map

$$\Delta^* \xrightarrow{(\mathrm{id}\otimes\varphi)^*} (k[[t]] \otimes_{\varphi, k[[t]]} \Delta)^*. \tag{4}$$

On the other hand, we have the $k[[t]]$-linear map

$$\begin{aligned} k[[t]] \otimes_{\varphi, k[[t]]} (\Delta^*) &\to (k[[t]] \otimes_{\varphi, k[[t]]} \Delta)^* \\ a \otimes \ell &\mapsto [b \otimes x \mapsto \ell(\psi(ab)x)]. \end{aligned} \tag{5}$$

It is shown in [Grosse-Klönne 2019] that the respective base extended maps $(4)\otimes_{k[[t]]}k((t))$ and $(5)\otimes_{k[[t]]}k((t))$ are bijective. Composing $(5)\otimes_{k[[t]]}k((t))$ with the inverse of $(4)\otimes_{k[[t]]}k((t))$ thus yields a $k((t))$-linear isomorphism

$$k((t)) \otimes_{\varphi, k((t))} (\Delta^* \otimes_{k[[t]]} k((t))) = k((t)) \otimes_{\varphi, k[[t]]} (\Delta^*) \to \Delta^* \otimes_{k[[t]]} k((t))$$

and hence the desired φ-operator on $\Delta^* \otimes_{k[[t]]} k((t))$. The exactness of $\Delta \mapsto \Delta^* \otimes_{k[[t]]} k((t))$ follows from the exactness of taking duals and of applying $(.) \otimes_{k[[t]]} k((t))$. □

1B. *ψ-stable lattices in (φ, Γ)-modules.*

Lemma 1.5. *Let $\boldsymbol{D} \in \mathrm{Mod}^{\mathrm{et}}(k((t)))$. There is a natural additive operator $\psi : \boldsymbol{D} \to \boldsymbol{D}$ satisfying*

$$\psi(a\varphi(x)) = \psi(a)x \quad \textit{and} \quad \psi(\varphi(a)x) = a\psi(x)$$

for all $a \in k((t))$ and all $x \in \boldsymbol{D}$, and commuting with the action of Γ.

Proof. We define the composed map

$$\psi : \boldsymbol{D} \to k((t)) \otimes_{\varphi, k((t))} \boldsymbol{D} \to \boldsymbol{D}$$

where the first arrow is the inverse of the structure isomorphism $\mathrm{id}\otimes\varphi$ and where the second arrow is given by $a\otimes x\mapsto\psi(a)x$. By construction, it satisfies $\psi(a\varphi(x))=\psi(a)x$. To see $\psi(\varphi(a)x)=a\psi(x)$ observe that by assumption we may write $x=\sum_i a_i\varphi(d_i)$ with $d_i\in\boldsymbol{D}$ and $a_i\in k((t))$. We then compute

$$\psi(\varphi(a)x)=\sum_i\psi(\varphi(a)a_i\varphi(d_i))=\sum_i\psi(\varphi(a)a_i)d_i=a\sum_i\psi(a_i)d_i=a\sum_i\psi(a_i\varphi(d_i))=a\psi(x).$$

Finally, let $\gamma\in\Gamma$. As γ and φ commute on $k[\![t]\!]$, and as Γ acts semilinearly on $\boldsymbol{D}$, the additive map

$$k((t))\otimes_{\varphi,k((t))}\boldsymbol{D}\to k((t))\otimes_{\varphi,k((t))}\boldsymbol{D},\quad a\otimes d\mapsto\gamma(a)\otimes\gamma(b)$$

is the map corresponding to γ on $\boldsymbol{D}$ under the isomorphism $\mathrm{id}\otimes\varphi$, and under $a\otimes x\mapsto\psi(a)x$ it commutes with γ on $\boldsymbol{D}$ since γ and ψ commute on $k((t))$. □

In the following, by a lattice in a $k((t))$-vector space $\boldsymbol{D}$ we mean a free $k[\![t]\!]$-submodule containing a $k((t))$-basis of $\boldsymbol{D}$.

Lemma 1.6. *Let $\boldsymbol{D}\in\mathrm{Mod}^{\mathrm{et}}(k((t)))$ and let D be a lattice in (the $k((t))$-vector space underlying) $\boldsymbol{D}$. Let $\psi:\boldsymbol{D}\to\boldsymbol{D}$ be the operator constructed in Lemma 1.5:*

(a) *$\psi(D)$ is a $k[\![t]\!]$-module.*

(b) *If $\varphi(D)\subset D$ then $D\subset\psi(D)$.*

(c) *If $D\subset k[\![t]\!]\cdot\varphi(D)$ then $\psi(D)\subset D$.*

(d) *If $\psi(D)\subset D$ then $\psi(t^{-1}D)\subset t^{-1}D$, and for each $x\in\boldsymbol{D}$ there is some $n(x)\in\mathbb{N}$ such that for all $n\geq n(x)$ we have $\psi^n(x)\in t^{-1}D$.*

Proof. (a) Use $\psi(\varphi(a)x)=a\psi(x)$ for $a\in k((t))$ and $x\in\boldsymbol{D}$.

(b) Choose $a\in k[\![t]\!]$ with $\psi(a)=1$. For $d\in D$ we have $d=\psi(a\varphi(d))$ which belongs to $\psi(D)$ since $\varphi(D)\subset D$.

(c) Let $d\in D$. By assumption there are $e_i\in D$ and $a_i\in k[\![t]\!]$ with $d=\sum_i a_i\varphi(e_i)$, hence $\psi(d)=\sum_i\psi(a_i)e_i\in D$.

(d) For $i\geq 1$ we have

$$\psi(\varphi^i(t^{-1})D)\subset\varphi^{i-1}(t^{-1})\psi(D)\subset\varphi^{i-1}(t^{-1})D \tag{6}$$

where the second inclusion uses the assumption. From $\varphi(t^{-1})=t^{-q}$ we get

$$\psi(t^{-1}D)\subset\psi(\varphi(t^{-1})D)\subset t^{-1}D.$$

Moreover, if $n(x)\in\mathbb{N}$ is such that $x\in\varphi^n(t^{-1})D$ for $n\geq n(x)$, then iterated application of formula (6) shows

$$\psi^n(x)\in\psi^n(\varphi^n(t^{-1})D)\subset\psi^{n-1}(\varphi^{n-1}(t^{-1})D)\subset\cdots\subset t^{-1}D$$

for $n\geq n(x)$. □

Lemma 1.7. (a) *There are lattices D_0, D_1 in $\boldsymbol{D}$ with*

$$\varphi(D_0) \subset tD_0 \subset D_0 \subset D_1 \subset k[\![t]\!] \cdot \varphi(D_1).$$

(b) *For D_0, D_1 as in (a) and for $n \geq 0$ we have $\psi^n(D_0) \subset \psi^{n+1}(D_0) \subset D_1$.*

Proof. (a) This is a (simplified) subclaim in the proof of Lemma 2.2.10 in [Schneider 2017] (which follows [Colmez 2010, Lemme II 2.3]). One proceeds as follows. Let $d_1, \dots, d_r$ be a $k((t))$-basis of $\boldsymbol{D}$. Then also $\varphi(d_1), \dots, \varphi(d_r)$ is $k((t))$-basis of $\boldsymbol{D}$. We therefore find $\tilde{f}_{ij}, \tilde{g}_{ij} \in k((t))$ with $\varphi(d_j) = \sum_{i=1}^r \tilde{f}_{ij} d_i$ and $d_j = \sum_{i=1}^r \tilde{g}_{ij} \varphi(d_i)$, for any $1 \leq j \leq r$. Choose some $n \geq 0$ with $t^{n(q-1)} \tilde{f}_{ij} \in tk[\![t]\!]$ and $t^{n(q-1)} \tilde{g}_{ij} \in tk[\![t]\!]$ for all i, j. Then $D_0 = \sum_{i=1}^r t^n k[\![t]\!] d_i$ and $D_1 = \sum_{i=1}^r t^{-n} k[\![t]\!] d_i$ work as desired.

(b) Choose $a \in k[\![t]\!]$ with $\psi(a) = 1$. For $x \in D_0$ we have $\psi^n(x) = \psi^{n+1}(a\varphi(x)) \in \psi^{n+1}(D_0)$ since $\varphi(D_0) \subset tD_0$ implies $\varphi(x) \in D_0$ and hence $a\varphi(x) \in D_0$. This shows $\psi^n(D_0) \subset \psi^{n+1}(D_0)$. As $D_0 \subset D_1 \subset k[\![t]\!] \cdot \varphi(D_1)$, an induction using Lemma 1.6(c) shows $\psi^{n+1}(D_0) \subset D_1$. □

Proposition 1.8. *There exists a unique lattice $\boldsymbol{D}^\sharp$ in $\boldsymbol{D}$ with $\psi(\boldsymbol{D}^\sharp) = \boldsymbol{D}^\sharp$ and such that for each $x \in \boldsymbol{D}$ there is some $n \in \mathbb{N}$ with $\psi^n(x) \in \boldsymbol{D}^\sharp$.*

For any lattice D in $\boldsymbol{D}$ we have $\psi^n(D) \subset \boldsymbol{D}^\sharp$ for all $n \gg 0$.

For any lattice D in $\boldsymbol{D}$ with $\psi(D) = D$ we have $t\boldsymbol{D}^\sharp \subset D \subset \boldsymbol{D}^\sharp$.

Proof. Using the previous lemmata, the proof is the same as the one given in [Colmez 2010, Proposition II.4.2]. □

Proposition 1.9. (a) *For any lattice D in $\boldsymbol{D}$ contained in $\boldsymbol{D}^\sharp$ and stable under ψ we have $\psi(D) = D$.*

(b) *The intersection $\boldsymbol{D}^\natural$ of all lattices in $\boldsymbol{D}$ contained in $\boldsymbol{D}^\sharp$ and stable under ψ is itself a lattice, and it satisfies $\psi(\boldsymbol{D}^\natural) = \boldsymbol{D}^\natural$.*

Proof. (See [Colmez 2010, Proposition II.5.11 and Corollaire II.5.12].)

(a) Since $\boldsymbol{D}^\sharp$ as well as D and $\psi(D)$ are lattices in $\boldsymbol{D}^\sharp$, both $\boldsymbol{D}^\sharp/D$ and $\boldsymbol{D}^\sharp/\psi(D)$ are finite dimensional k-vector spaces. ψ induces an isomorphism $\psi(\boldsymbol{D}^\sharp)/D = \boldsymbol{D}^\sharp/\psi(D)$ (as $\psi(D) \subset D$), hence $\psi(D) = D$.

(b) For any D as in (a) we have $t\boldsymbol{D}^\sharp \subset D$ by what we saw in (a) together with Proposition 1.8. This shows $t\boldsymbol{D}^\sharp \subset \boldsymbol{D}^\natural$, hence $\boldsymbol{D}^\natural$ is indeed a lattice, and $\psi(\boldsymbol{D}^\natural) = \boldsymbol{D}^\natural$ follows by applying (a) once more. □

Lemma 1.10. *$\boldsymbol{D}^\natural$ and $\boldsymbol{D}^\sharp$ are stable under the action of Γ.*

Proof. If D is a lattice in $\boldsymbol{D}$, then so is $\gamma \cdot D$ for any $\gamma \in \Gamma$. If in addition $\psi(D) \subset D$, resp. $\psi(D) = D$, then also $\psi(\gamma \cdot D) \subset \gamma \cdot D$, resp. $\psi(\gamma \cdot D) = \gamma \cdot D$. From these observations we immediately get $\gamma \cdot \boldsymbol{D}^\natural = \boldsymbol{D}^\natural$ and $\gamma \cdot \boldsymbol{D}^\sharp = \boldsymbol{D}^\sharp$. □

Proposition 1.11. *The functor* $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O}) \to \mathrm{Mod}^{\mathrm{et}}(k((t)))$ *in Proposition 1.4 sends simple objects to simple objects.*

Proof. Let $\Delta \in \mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ be simple. By construction, ψ on $\Delta^* \otimes_{k[\![t]\!]} k((t))$, when restricted to Δ^*, is the adjoint of φ on Δ. Therefore the simplicity of Δ as an $\mathfrak{O}$-module means that Δ^* admits no nontrivial $k[\![t]\!]$-submodule stable under Γ and ψ. If $\boldsymbol{D}$ is a nonzero (φ, Γ)-submodule of $\Delta^* \otimes_{k[\![t]\!]} k((t))$ then also $\boldsymbol{D}^\natural$ is nonzero and stable under Γ and ψ, see Proposition 1.9 and Lemma 1.10. As $\boldsymbol{D}^{\natural} \subset (\Delta^* \otimes_{k[\![t]\!]} k((t)))^{\natural} \subset \Delta^*$ we get $\boldsymbol{D}^{\natural} = \Delta^*$ (since Δ^* is stable under ψ), as desired. $\square$

Lemma 1.12. *Let $f : \boldsymbol{D}_1 \to \boldsymbol{D}_2$ be a morphism in* $\mathrm{Mod}^{\mathrm{et}}(k((t)))$:

(a) $f(\boldsymbol{D}_1^{\sharp}) \subset \boldsymbol{D}_2^{\sharp}$ *and* $f(\boldsymbol{D}_1^{\natural}) \subset \boldsymbol{D}_2^{\natural}$.

(b) *If $f : \boldsymbol{D}_1 \to \boldsymbol{D}_2$ is injective* (*resp. surjective*), *then so is $f : \boldsymbol{D}_1^{\sharp} \to \boldsymbol{D}_2^{\sharp}$.*

(c) *If $f : \boldsymbol{D}_1 \to \boldsymbol{D}_2$ is injective* (*resp. surjective*), *then so is $f : \boldsymbol{D}_1^{\natural} \to \boldsymbol{D}_2^{\natural}$.*

Proof. (a) $f(\boldsymbol{D}_1^{\sharp})$ is a free $k[\![t]\!]$-submodule of $\boldsymbol{D}_2$ on which ψ acts surjectively. Thus $f(\boldsymbol{D}_1^{\sharp}) + \boldsymbol{D}_2^{\sharp}$ is a lattice satisfying the defining condition for $\boldsymbol{D}_2^{\sharp}$ given in Proposition 1.8, hence $f(\boldsymbol{D}_1^{\sharp}) + \boldsymbol{D}_2^{\sharp} = \boldsymbol{D}_2^{\sharp}$, hence $f(\boldsymbol{D}_1^{\sharp}) \subset \boldsymbol{D}_2^{\sharp}$. Next, let $D = \{x \in \boldsymbol{D}_1^{\natural};\, f(x) \in \boldsymbol{D}_2^{\natural}\}$. It is a lattice in $\boldsymbol{D}_1$ since $\boldsymbol{D}_1^{\natural}$ is a lattice, $f(\boldsymbol{D}_1^{\natural}) \subset f(\boldsymbol{D}_1^{\sharp}) \subset \boldsymbol{D}_2^{\sharp}$ and $\boldsymbol{D}_2^{\sharp}/\boldsymbol{D}_2^{\natural}$ is a finite dimensional k-vector space. It is also stable under ψ, hence contains $\boldsymbol{D}_1^{\natural}$, hence $f(\boldsymbol{D}_1^{\natural}) \subset \boldsymbol{D}_2^{\natural}$.

(b) and (c) If $f : \boldsymbol{D}_1 \to \boldsymbol{D}_2$ is injective then obviously so are $f : \boldsymbol{D}_1^{\sharp} \to \boldsymbol{D}_2^{\sharp}$ and $f : \boldsymbol{D}_1^{\natural} \to \boldsymbol{D}_2^{\natural}$. If $f : \boldsymbol{D}_1 \to \boldsymbol{D}_2$ is surjective then $f(\boldsymbol{D}_1^{\natural})$ is a lattice in $\boldsymbol{D}_2$ stable under ψ, hence contains $\boldsymbol{D}_2^{\natural}$. To see $f(\boldsymbol{D}_1^{\sharp}) = \boldsymbol{D}_2^{\sharp}$ we proceed as in [Colmez 2010, Proposition II.4.6(iii)] Namely, choose a lattice D' in $\boldsymbol{D}_1$ with $f(D') = \boldsymbol{D}_2^{\sharp}$. Put $D = \sum_{n \geq 0} \psi^n(D')$. By construction we have $\psi(D) \subset D$ as well as $f(D) = \boldsymbol{D}_2^{\sharp}$ (since $\psi(\boldsymbol{D}_2^{\sharp}) = \boldsymbol{D}_2^{\sharp}$). Proposition 1.8 shows that D is again a lattice. Let $x \in \boldsymbol{D}_2^{\sharp}$. By Proposition 1.8 we find some $n \in \mathbb{N}$ such that $\psi^n(D) \subset \boldsymbol{D}_1^{\sharp}$. For such an n, choose $x_n \in \boldsymbol{D}_2^{\sharp}$ and $\tilde{x}_n \in D$ with $\psi^n(x_n) = x$ and $f(\tilde{x}_n) = x_n$. Put $u_n = \psi^n(\tilde{x}_n) \in \boldsymbol{D}_1^{\sharp}$. By their construction in Lemma 1.5, the operators ψ on $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ commute with f, thus we may compute

$$f(u_n) = f(\psi^n(\tilde{x}_n)) = \psi^n(f(\tilde{x}_n)) = \psi^n(x_n) = x. \qquad \square$$

Lemma 1.13. *Let* $0 \to \boldsymbol{D}_1 \to \boldsymbol{D}_2 \to \boldsymbol{D}_3 \to 0$ *be an exact sequence in* $\mathrm{Mod}^{\mathrm{et}}(k((t)))$. *For each i let $D_i \subset \boldsymbol{D}_i$ be a lattice with $\psi(D_i) = D_i$, and suppose that the above sequence restricts to an exact sequence*

$$0 \to D_1 \to D_2 \to D_3 \to 0. \tag{7}$$

If we have $D_1 = \boldsymbol{D}_1^{\natural}$ and $D_3 = \boldsymbol{D}_3^{\natural}$, then we also have $D_2 = \boldsymbol{D}_2^{\natural}$. If we have $D_1 = \boldsymbol{D}_1^{\sharp}$ and $D_3 = \boldsymbol{D}_3^{\sharp}$ then we also have $D_2 = \boldsymbol{D}_2^{\sharp}$.

Proof. By Lemma 1.12 the sequence $0 \to \boldsymbol{D}_1^{\natural} \to \boldsymbol{D}_2^{\natural} \to \boldsymbol{D}_3^{\natural} \to 0$ is exact on the left and on the right. Comparing it with the sequence (7) via $\boldsymbol{D}_1^{\natural} = D_1$, $\boldsymbol{D}_2^{\natural} \subset D_2$ and $\boldsymbol{D}_3^{\natural} = D_3$, we immediately get $\boldsymbol{D}_2^{\natural} = D_2$. Next, by Lemma 1.12 the sequence $0 \to \boldsymbol{D}_1^{\sharp} \to \boldsymbol{D}_2^{\sharp} \to \boldsymbol{D}_3^{\sharp} \to 0$ is exact on the left and on the right. We compare it with the sequence (7) via $D_1 = \boldsymbol{D}_1^{\sharp}$, $D_2 \subset \boldsymbol{D}_2^{\sharp}$ and $D_3 = \boldsymbol{D}_3^{\sharp}$. We claim

$$\psi(\boldsymbol{D}_1 \cap \boldsymbol{D}_2^{\sharp}) = \boldsymbol{D}_1 \cap \boldsymbol{D}_2^{\sharp}.$$

Of course, $\psi(\boldsymbol{D}_1 \cap \boldsymbol{D}_2^{\sharp}) \subset \boldsymbol{D}_1 \cap \boldsymbol{D}_2^{\sharp}$ is clear. To see $\boldsymbol{D}_1 \cap \boldsymbol{D}_2^{\sharp} \subset \psi(\boldsymbol{D}_1 \cap \boldsymbol{D}_2^{\sharp})$ take $x \in \boldsymbol{D}_1 \cap \boldsymbol{D}_2^{\sharp}$. Choose $y \in \boldsymbol{D}_2^{\sharp}$ with $\psi(y) = x$. Choose $y' \in D_2$ mapping to the same element in $\boldsymbol{D}_3^{\sharp} = D_3$ as y. We then have

$\psi(y') \in D_2 \cap \boldsymbol{D}_1 = D_1$ and $\psi(y-y') - x \in D_1$, hence there is some $z \in D_1$ with $\psi(z) = \psi(y-y') - x$, hence $x = \psi(y - y' - z) \in \psi(\boldsymbol{D}_1 \cap \boldsymbol{D}_2^\sharp)$ since $y - y' \in \boldsymbol{D}_1 \cap \boldsymbol{D}_2^\sharp$ and $z \in \boldsymbol{D}_1 \cap \boldsymbol{D}_2^\sharp$.

The claim is proven. By the definition of $\boldsymbol{D}_1^\sharp$ it implies $\boldsymbol{D}_1 \cap \boldsymbol{D}_2^\sharp = \boldsymbol{D}_1^\sharp$, hence $\boldsymbol{D}_1 \cap \boldsymbol{D}_2^\sharp = D_1$ since $D_1 = \boldsymbol{D}_1^\sharp$. Thus, $D_2 = \boldsymbol{D}_2^\sharp$. $\square$

Remark. An étale φ-module over $k((t))$ is a $k[[t]][\varphi] \otimes_{k[[t]]} k((t))$-module $\boldsymbol{D}$ which is finite dimensional over $k((t))$ such that the $k((t))$-linearized structure map $\mathrm{id} \otimes \varphi$ is bijective. The above theory of the operator ψ and the lattices $\boldsymbol{D}^\sharp$ and $\boldsymbol{D}^\natural$ works analogously for étale φ-modules $\boldsymbol{D}$ over $k((t))$, i.e., the Γ-action is not really needed.

1C. *Partial full faithfulness of* $\Delta \mapsto \Delta^* \otimes_{k[[t]]} k((t))$.

Lemma 1.14. *Let N be a k-vector space, and suppose that we are given a k-linear automorphism τ of N, a basis $\mathcal{N}$ of N, integers $0 \le k_\nu \le q-1$ and units $\alpha_\nu \in k^\times$ for $\nu \in \mathcal{N}$. View N as a $k[[t]]$-module with $tN = 0$ and let Δ denote the quotient of $k[[t]][\varphi] \otimes_{k[[t]]} N$ by the $k[[t]][\varphi]$-submodule ∇ generated by the elements*

$$1 \otimes \nu + \alpha_\nu t^{k_\nu} \varphi \otimes \tau(\nu)$$

with $\nu \in \mathcal{N}$. We then have:

(a) *$k[[t]][\varphi] \otimes_{k[[t]]} N$ is a torsion $k[[t]]$-module.*

(b) *The map $N \to \Delta[t]$ sending $n \in N$ to the class of $1 \otimes n$ is an isomorphism. In particular, Δ is admissible if N is a finite dimensional k-vector space.*

(c) *The action of φ on Δ is injective.*

Proof. (a) As $\varphi t = t^q \varphi$ in $k[[t]][\varphi]$ we may write any element in $k[[t]][\varphi] \otimes_{k[[t]]} N$ as a finite sum of elements of the form $a\varphi^n \otimes x$ with $a \in k[[t]]$, $n \ge 0$ and $x \in N$. It is therefore enough to show

$$a\varphi^n \otimes x = 0 \quad \text{for each } a \in t^{q^n} k[[t]] \tag{8}$$

where $n \ge 0$ and $x \in N$. We may write $a = a_0 t^{q^n}$ with $a_0 \in k[[t]]$ and compute

$$a\varphi^n \otimes x = a_0 t^{q^n} \varphi^n \otimes x = a_0 \varphi^n t \otimes x = 0.$$

(b) and (c) We may write

$$k[[t]][\varphi] \otimes_{k[[t]]} N \cong \bigoplus_{\nu \in \mathcal{N}} \bigoplus_{i \ge 0} \bigoplus_{0 \le \theta \le q^i - 1} k.t^\theta \varphi^i \otimes \tau(\nu).$$

Indeed, that $k[[t]][\varphi] \otimes_{k[[t]]} N$ is a quotient of the right hand side follows from formula (8). It is in fact an isomorphic quotient since all relations between φ and t in $k[[t]][\varphi]$ can be generated from $\varphi t = t^q \varphi$.

Consider the three k-subvector spaces

$$1\otimes N=\bigoplus_{\nu\in\mathcal{N}}k\otimes\tau(\nu)$$
$$=\bigoplus_{\nu\in\mathcal{N}}k\otimes\nu,$$
$$C=\bigoplus_{\nu\in\mathcal{N}}\bigoplus_{i>0}\bigoplus_{0\leq\theta<q^{i-1}k_\nu}k.t^\theta\varphi^i\otimes\tau(\nu),\tag{9}$$
$$\nabla=\bigoplus_{\nu\in\mathcal{N}}\bigoplus_{i>0}\bigoplus_{\epsilon\geq 0}k.t^\epsilon\varphi^{i-1}(1\otimes\nu+\alpha_\nu t^{k_\nu}\varphi\otimes\tau(\nu)).\tag{10}$$

Using the formula $\varphi t=t^q\varphi$ we see

$$t^\epsilon\varphi^{i-1}(1\otimes\nu+\alpha_\nu t^{k_\nu}\varphi\otimes\tau(\nu))\in k^\times.t^{\epsilon+q^{i-1}k_\nu}\varphi^i\otimes\tau(\nu)+k[\![t]\!]\varphi^{i-1}\otimes\nu.$$

We also see that in the sum (10) all summands with $\epsilon\geq(q-1)q^{i-1}k_\nu-1$ vanish. Equivalently, in the sum (10) only those summands are nonzero for which $\theta=\epsilon+q^{i-1}k_\nu$ satisfies $q^{i-1}k_\nu\leq\theta\leq q^i-1$. Thus we find

$$k[\![t]\!][\varphi]\otimes_{k[\![t]\!]}N\cong 1\otimes N\bigoplus\nabla\bigoplus C.\tag{11}$$

Let C', resp. C'', denote the k-subspace of C spanned by all $t^\theta\varphi^i\otimes\tau(\nu)$ with $\nu\in\mathcal{N}$, $i>1$ and $0\leq\theta<q^{i-1}k_\nu$, resp. by all $t^\theta\varphi\otimes\tau(\nu)$ with $\nu\in\mathcal{N}$ and $0\leq\theta<k_\nu$. Then $\varphi(C)\subset C'$ and $\varphi:C\to C'$ is injective. On the other hand, $\varphi(1\otimes N)\subset C''$ and $\varphi:1\otimes N\to C''$ is injective. Since $C'\cap C''=0$ we conclude that φ acts injectively on Δ. Now consider the composed map

$$C\to k[\![t]\!][\varphi]\otimes_{k[\![t]\!]}N\xrightarrow{t(.)}k[\![t]\!][\varphi]\otimes_{k[\![t]\!]}N\to 1\otimes N\bigoplus C$$

where the first arrow is the inclusion, the last arrow is the projection. This map is bijective, the critical point being the computation

$$t(k.t^{q^{i-1}k_\nu-1}\varphi^i\otimes\tau(\nu))=k.t^{q^{i-1}k_\nu}\varphi^i\otimes\tau(\nu)=k.\varphi^{i-1}t^{k_\nu}\varphi\otimes\tau(\nu)\equiv k.\varphi^{i-1}\otimes\nu$$

modulo ∇ (for $i>0$). It follows that indeed the image of $1\otimes N$ in Δ is the kernel of t acting on Δ. □

Definition. An object $\Delta\in\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ is called standard cyclic if it is generated over $k[\![t]\!][\varphi]$ by $\ker(t|_\Delta)=\Delta[t]$ and if there is a basis of $\Delta[t]$ consisting of Γ-eigenvectors $e_0,\ldots,e_d$ such that

$$t^{k_i}\varphi e_{i-1}=\rho_i e_i\quad\text{for all }0\leq i\leq d$$

(reading $e_{-1}=e_d$), for certain $0\leq k_i\leq q-1$ and $\rho_i\in k^\times$ such that $k_i>0$ for at least one i, as well as $k_i<q-1$ for at least one i.

In the following, we extend any indexing by $0,\ldots,d$ to an indexing by $\mathbb{Z}$ in the obvious way (i.e., $k_i=k_{i+d+1}$, $e_i=e_{i+d+1}$, $\rho_i=\rho_{i+d+1}$, $\eta_i=\eta_{i+d+1}$ for all $i\in\mathbb{Z}$). Let ∇ denote the $k[\![t]\!][\varphi]$-submodule of $k[[t]][\varphi]\otimes_{k[[t]]}\Delta[t]$ generated by the elements $t^{k_i}\varphi\otimes e_{i-1}-1\otimes\rho_i e_i$. The inclusion $\Delta[t]\to\Delta$ extends to a natural $k[[t]][\varphi]$-linear map

$$k[[t]][\varphi]\otimes_{k[[t]]}\Delta[t]/\nabla\to\Delta. \tag{12}$$

Proposition 1.15. *Let $\Delta\in\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ be standard cyclic, with e_i, k_i, ρ_i, ρ_i as above*:

(a) *t acts surjectively on Δ, and there is a distinguished isomorphism of free $k[[t]]$-modules of rank $d+1$*

$$\Delta^*\cong k[[t]]\otimes_k(\Delta[t]^*). \tag{13}$$

The map (12) *is a $k[[t]][\varphi]$-linear isomorphism.*

(b) *If for any $1\le j\le d$ there is some $0\le i\le d$ with $k_i\ne k_{i+j}$, then Δ is irreducible as a $k[[t]][\varphi]$-module.*

(c) *For $0\le i\le d$ let $\eta_i:\Gamma\to k^\times$ be the character with $\gamma\cdot e_i=\eta_i(\gamma)e_i$ for all $\gamma\in\Gamma$. Suppose that for any $1\le j\le d$ which satisfies $k_i=k_{i+j}$ for all $0\le i\le d$ there is some $0\le i\le d$ with $\eta_i\ne\eta_{i+j}$. Then Δ is irreducible as an $\mathfrak{O}$-module.*

(d) *At least after a finite extension of k we have: Δ admits a filtration such that each associated graded piece is an irreducible standard cyclic object in* $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$. *If p does not divide $d+1$ then Δ is even the direct sum of irreducible standard cyclic objects in* $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$.

Proof. (This is very similar to [Grosse-Klönne 2016, Proposition 6.2].)

(a) For $0\le j\le d$ consider

$$w_j=k_j+qk_{j-1}+\cdots+q^jk_0+q^{j+1}k_d+\cdots+q^dk_{j+1}.$$

Repeated substitution of $\varphi t=t^q\varphi$ (recall $\Phi(t)=t^q$ modulo π) shows that $t^{w_j}\varphi^{d+1}e_j\in k^\times e_j$. As $k_i>0$ for at least one i we have $w_j>0$, and hence $e_j\in t\Delta$. As $\Delta[t]$ is generated over k by all e_j it follows that $\Delta[t]\subset t\Delta$. As Δ is generated over $k[[t]][\varphi]$ by $\Delta[t]$, the equation $\varphi t=t^q\varphi$ therefore shows $\Delta\subset t\Delta$, i.e., t acts surjectively on Δ. We deduce that Δ^* is a torsion free, and hence free $k[[t]]$-module of rank $d+1$. As Δ is generated over $k[[t]][\varphi]$ by $\Delta[t]$ the map (12) is surjective. But it is also injective, because Lemma 1.14 tells us that it induces an isomorphisms between the respective kernels of t. We view the bijective map (12) as an identification. The proof of Lemma 1.14 yielded a canonical k-vector space decomposition $\Delta=C\oplus\Delta[t]$ where the k-subvector space C of Δ is generated by the image elements of the elements $t^\theta\varphi^r\otimes e\in k[[t]][\varphi]\otimes_{k[[t]]}\Delta[t]$ which do not belong to $1\otimes\Delta[t]$ (for some $e\in\Delta[t]$, and some $\theta,r\ge 0$). We may thus identify $\Delta[t]^*=\mathrm{Hom}_k(\Delta[t],k)$ with the subspace of $\Delta^*=\mathrm{Hom}_k(\Delta,k)$ consisting of all $f\in\Delta^*$ with $f|_C=0$. The composition of this k-linear embedding $\Delta[t]^*\to\Delta^*$ with the projection $\Delta^*\to(\Delta^*)/t(\Delta^*)$ is a k-linear isomorphism. Therefore, and as Δ^* is free and finitely generated over $k[[t]]$, the $k[[t]]$-linear map $k[[t]]\otimes_k(\Delta[t]^*)\to\Delta^*$ extending the k-linear embedding $\Delta[t]^*\to\Delta^*$ is an isomorphism as stated in formula (13).

(b) Let Z be a nonzero $k[[t]][\varphi]$-submodule of Δ. With Δ also Z is a torsion $k[[t]]$-module, hence $\ker(t|_Z) = Z[t]$ is nonzero. For nonzero elements $z = \sum_{0\leq i\leq d} x_i e_i$ of $Z[t]$ (with $x_i \in k$) put

$$\mathcal{D}(z) = \{0 \leq i \leq d \mid x_i \neq 0\}, \qquad \nu(z) = |\mathcal{D}(z)|,$$
$$\eta(z) = \max\{k_i \mid i \in \mathcal{D}(z)\}, \qquad \Lambda(z) = t^{\eta(z)}\varphi z.$$

Then $\Lambda(z)$ is again a nonzero element of $Z[t]$. We have

$$\mathcal{D}(\Lambda(z)) = \{i+1 \mid \eta(z) = k_i \text{ and } i \in \mathcal{D}(z)\}$$

(we read elements in $\{0 \leq i \leq d\}$ modulo $(d+1)$), in particular $\nu(\Lambda(z)) \leq \nu(z)$. If $\nu(\Lambda(z)) = \nu(z)$ then $\mathcal{D}(\Lambda(z)) = \{i+1 \mid i \in \mathcal{D}(z)\}$ and $k_i = k_{i+j}$ whenever $i, i+j \in \mathcal{D}(z)$. This implies that if we had $\nu(\Lambda^n(z)) = \nu(z) > 1$ for all $n \geq 0$ then there was some $1 \leq j \leq d$ with $k_i = k_{i+j}$ for all $0 \leq i \leq d$. But this would contradict our hypothesis. Thus, for sufficiently large $n \geq 0$ we have $\nu(\Lambda^n(z)) = 1$, i.e., $\Lambda^n(z) \in k^\times e_i$ for some $0 \leq i \leq d$. For such n we then even have $\Lambda^{n+j}(z) \in k^\times e_{i+j}$ for all $j \geq 0$. It follows that Z contains all e_i, hence $Z = \Delta$.

(c) We use the functions ν, Λ already employed in the proof of (b). Let $0 \neq Z \subset \Delta$ be a nonzero $\mathfrak{O}$-submodule. Choose a nonzero $z \in Z[t]$ for which $\nu(z)$ is minimal (for all nonzero $z \in Z[t]$). If $\nu(z) = 1$ then we obtain $Z = \Delta$ as in the proof of (b). Now assume $\nu(z) > 1$. For all $n \geq 0$ we have $\nu(\Lambda^n(z)) \leq \nu(z)$, hence $\nu(\Lambda^n(z)) = \nu(z)$ by the choice of z. Thus, writing $z = \sum_{0\leq i\leq d} x_i e_i$ with $x_i \in k$, we have $x_i \neq 0$ and $x_{i+j} \neq 0$ for some i, j, with j violating the hypothesis in (b). By the hypothesis in (c), replacing i by $i+n$ and z by $\Lambda^n(z)$ we may assume that $\eta_i \neq \eta_{i+j}$. Pick $\gamma \in \Gamma$ with $\eta_i(\gamma) \neq \eta_{i+j}(\gamma)$, and pick $a \in k^\times$ with $ae_i = \gamma \cdot e_i$. Then $az - \gamma \cdot z$ is a nonzero element in $Z[t]$ with $\nu(az - \gamma \cdot z) < \nu(z)$: a contradiction.

(d) Passing to a finite extension of k if necessary we may assume that there is a $(d+1)$-st root of $\prod_{i=0}^d \rho_i$ in k. Thus, rescaling the e_i if necessary we may assume $\rho_i = \rho_j$ for all i, j. We argue by induction on d. If Δ itself is not irreducible then there is, by (c), some $1 \leq j \leq d$ which satisfies $k_i = k_{i+j}$ and $\eta_i = \eta_{i+j}$ for all $0 \leq i \leq d$. The minimal such j is a divisor of $d+1$. Consider the k-subvector space V of $\Delta[t]$ spanned by the vectors $\epsilon_i = e_{ij}$ for $0 \leq i < (d+1)/j$. Then

$$\left(\prod_{i=1}^{j} \rho_i^{-1}\right) t^{k_j}\varphi \cdots t^{k_1}\varphi$$

induces the automorphism f of V with $f(\epsilon_i) = \epsilon_{i+1}$ (where we understand $\epsilon_{(d+1)/j} = \epsilon_0$). Choose (after passing to a finite extension of k if necessary) an f-stable filtration $0 = V_0 \subset V_1 \subset \cdots \subset V_{(d+1)/j} = V$ such that each V_i/V_{i-1} is one dimensional. Then define for $0 \leq s \leq (d+1)/j$ the $\mathfrak{O}$-submodule $\Delta_s = \mathfrak{O}V_0 + \cdots + \mathfrak{O}V_s$ of Δ. It induces on $\Delta[t]$ the filtration

$$\Delta_s[t] = \Delta_{s-1}[t] + V_s + t^{k_1}\varphi V_s + \cdots + t^{k_{j-1}}\varphi \cdots t^{k_1}\varphi V_s.$$

By construction, each Δ_{i+1}/Δ_i is standard cyclic, and the induction hypothesis applies. If p does not divide $(d+1)/j$ then there is even an f-stable direct sum decomposition $V = \oplus_s V_{[s]}$ with one dimensional $V_{[s]}$. Then $\Delta = \oplus_s \Delta_{[s]}$ with $\Delta_{[s]} = \mathfrak{O} V_{[s]}$ is a direct sum decomposition of Δ, and each $\Delta_{[s]}$ is standard cyclic, and the induction hypothesis applies. $\square$

Lemma 1.16. *Let $\Delta \in \mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ be standard cyclic and put $\boldsymbol{D} = \Delta^* \otimes_{k[\![t]\!]} k((t)) \in \mathrm{Mod}^{\mathrm{et}}(k((t)))$, see Proposition 1.4. We have $\boldsymbol{D}^\natural = \Delta^* = \boldsymbol{D}^\sharp$.*

Proof. In Proposition 1.15 we saw that Δ^* is a free $k[\![t]\!]$-module, hence the natural map $\Delta^* \to \boldsymbol{D} = \Delta^* \otimes_{k[\![t]\!]} k((t))$ is injective; we view it as an inclusion.

The φ-operator on Δ is the adjoint of the ψ-operator on $\boldsymbol{D}$, in such a way that $\psi(\Delta^*) = \Delta^*$ since φ acts injectively on Δ. Therefore the definitions of $\boldsymbol{D}^\natural$ and $\boldsymbol{D}^\sharp$ yield $\boldsymbol{D}^\natural \subset \Delta^* \subset \boldsymbol{D}^\sharp$. Since $\boldsymbol{D}^\natural$ is a lattice with $\psi(\boldsymbol{D}^\natural) = \boldsymbol{D}^\natural$ we get $t\Delta^\sharp \subset \boldsymbol{D}^\natural$, together

$$\boldsymbol{D}^\natural \subset \Delta^* \subset \boldsymbol{D}^\sharp \quad \text{and} \quad t\boldsymbol{D}^\sharp \subset \boldsymbol{D}^\natural. \tag{14}$$

Let e_i and k_i be as in the definition of Δ being standard cyclic.

Formula (14) implies $t\Delta^* \subset \boldsymbol{D}^\natural$, hence $t(\Delta^*/\boldsymbol{D}^\natural) = 0$, hence $\Delta^*/\boldsymbol{D}^\natural$ is dual to a subspace W of $\Delta[t]$ stable under φ. To prove $\boldsymbol{D}^\natural = \Delta^*$ it is therefore enough to prove that $\Delta[t]$ does not contain a nonzero subspace W stable under φ. Assume that such a W does exist. A nonzero element $\beta \in W$ may be written as $\beta = \sum_{i=0}^d \alpha_i e_i$ with $\alpha_i \in k$. Let $k = \max\{k_{i+1} \mid \alpha_i \neq 0\}$. Since by assumption $k_i > 0$ for at least one i, replacing β by $\varphi^r \beta$ for some $r \in \mathbb{N}$ if necessary, we may assume $k > 0$. But then $t^k \varphi\beta$ is a nonzero linear combination of the e_i, whereas we also have $t\varphi\beta = 0$ since $\varphi\beta \in W \subset \Delta[t]$: a contradiction.

Formula (14) implies $t\boldsymbol{D}^\sharp \subset \Delta^*$, i.e., $t(\boldsymbol{D}^\sharp/\Delta^*) = 0$. We endow $\boldsymbol{D}^\sharp$ and all its submodules with the t-adic topology. By Pontrjagin duality (as recalled e.g., in [Schneider and Venjakob 2016]) we in particular have $\mathrm{Hom}_k^{\mathrm{cont}}(\Delta^*, k) = \Delta$. Now $t(\boldsymbol{D}^\sharp/\Delta^*) = 0$ means that the kernel W of the natural projection $\mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k) \to \mathrm{Hom}_k^{\mathrm{cont}}(\Delta^*, k) = \Delta$ is contained in $\mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k)[t]$. As t acts injectively on $\boldsymbol{D}^\sharp$, it acts surjectively on $\mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k)$. Hence, if $\Delta^* \neq \boldsymbol{D}^\sharp$ then $W \neq 0$ and there is some $\beta \in \mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k)$ with $0 \neq t\beta \in W$. Now $t\beta \in W$ means that β maps to an element in $\Delta[t]$. Since on the other hand $tW = 0$ (as $W \subset \mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k)[t]$) we may write $\beta = \sum_{i=0}^d \alpha_i \tilde{e}_i$ with $\alpha_i \in k$, where $\tilde{e}_i \in \mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k)$ lifts e_i. We then also have $0 \neq t\tilde{e}_{i_0} \in W$ for some i_0. As φ is injective on W (which follows from the surjectivity of ψ on $\boldsymbol{D}^\sharp$ and hence on $W^* = \boldsymbol{D}^\sharp/\Delta^*$) this gives $t^q\varphi\tilde{e}_{i_0} = \varphi t\tilde{e}_{i_0} \neq 0$ in $\mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k)$. Together with $W \subset \mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^\sharp, k)[t]$ we get $t^{q-1}\varphi e_{i_0} \neq 0$ in Δ. Applying the same argument with $t^{q-1}\varphi\tilde{e}_{i_0}$ instead of $\tilde{e}_i$ (again using that $t^q\varphi\tilde{e}_{i_0} \neq 0$) we see $t^{q-1}\varphi t^{q-1}\varphi e_{i_0} \neq 0$. Next we get $t^{q-1}\varphi t^{q-1}\varphi t^{q-1}\varphi e_{i_0} \neq 0$ etc.. But this means $q - 1 = k_i$ for each i, contradicting the hypothesis. We obtain $\Delta^* = \boldsymbol{D}^\sharp$. $\square$

Definition. Let $\mathrm{Mod}^\clubsuit(\mathfrak{O})$ denote the subcategory of $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ whose objects admit a filtration such that each associated graded piece becomes a standard cyclic object in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ after a suitable field extension of k.

Remark. Proposition 1.15(d) implies that each subquotient in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ of an object in $\mathrm{Mod}^\clubsuit(\mathfrak{O})$ again is an object in $\mathrm{Mod}^\clubsuit(\mathfrak{O})$.

Proposition 1.17. *The restriction of the functor* (3) *to the category* $\mathrm{Mod}^\clubsuit(\mathfrak{O})$ *is exact and fully faithful.*

Proof. We already know that the functor is exact. Next, we claim

$$\boldsymbol{D}^{\natural}=\Delta^*=\boldsymbol{D}^{\sharp}\quad\text{with}\quad \boldsymbol{D}=\Delta^*\otimes_{k[[t]]}k((t)) \tag{15}$$

for $\Delta\in\mathrm{Mod}^{\clubsuit}(\mathfrak{O})$. Indeed, for standard cyclic Δ this is shown in Lemma 1.16. For Δ which become standard cyclic after a field extension k'/k it then follows since the definitions of $(.)^{\natural}$ and $(.)^{\sharp}$ in terms of the k-linear operator ψ imply $\boldsymbol{D}^{\natural}\otimes_k k'=(\boldsymbol{D}\otimes_k k')^{\natural}$ and $\boldsymbol{D}^{\sharp}\otimes_k k'=(\boldsymbol{D}\otimes_k k')^{\sharp}$. For general $\Delta\in\mathrm{Mod}^{\clubsuit}(\mathfrak{O})$ it then follows from Lemma 1.13. We now claim that the reverse functor (on the essential image of the functor under discussion) is given by sending $\boldsymbol{D}$ to the topological dual $(\boldsymbol{D}^{\natural})'$ of $\boldsymbol{D}^{\natural}$ (where we endow $\boldsymbol{D}^{\natural}$ with its t-adic topology). Indeed, for $\boldsymbol{D}$ in this essential image and for $\Delta\in\mathrm{Mod}^{\clubsuit}(\mathfrak{O})$ we have natural isomorphisms

$$((\boldsymbol{D}^{\natural})')^*\otimes_{k[[t]]}k((t))\overset{\text{(i)}}{\cong}\boldsymbol{D}^{\natural}\otimes_{k[[t]]}k((t))\cong\boldsymbol{D},\qquad ((\Delta^*\otimes_{k[[t]]}k((t)))^{\natural})'\overset{\text{(ii)}}{\cong}(\Delta^*)'\overset{\text{(iii)}}{\cong}\Delta,$$

where (i) and (iii) follow from Pontrjagin duality, see e.g., Proposition 5.4 in [Schneider and Venjakob 2016], and where (ii) follows from formula (15). □

1D. *Standard cyclic étale (φ,Γ)-modules.*

Proposition 1.18. *Let $\Delta\in\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ be a standard cyclic object, with $d, e_i, k_i, \rho_i, \eta_i$ as in the definition resp. as in Proposition 1.15. The étale (φ,Γ)-module $\Delta^*\otimes_{k[[t]]}k((t))$ over $k((t))$ admits a $k((t))$-basis $f_0,\dots,f_d$ such that for all $0\le j\le d$ we have*

$$\varphi(f_{j-1})=\rho_{j-1}^{-1}t^{1+k_j-q}f_j \tag{16}$$

(reading $f_{-1}=f_d$), and moreover

$$\gamma\cdot f_j-\eta_j^{-1}(\gamma)f_j\in tk[[t]]f_j\quad\textit{for all }\gamma\in\Gamma. \tag{17}$$

Proof. We use formula (2).

First we assume $F\neq\mathbb{Q}_p$. Put $N=\oplus_{i=0}^d k.e_i$. As explained in the proof of Proposition 1.15, we have a bijective map (12) which we view as an identification. In particular, Lemma 1.14 and its proof apply. In the context of that proof we identify e_i with the class of $1\otimes e_i$ in Δ. By formula (11) we have a k-linear isomorphism $(1\otimes N)\oplus C\cong\Delta$ with C as in formula (9). For $0\le j\le d$ we may therefore define $f_j\in\Delta^*$ by asking $f_j(C)=0$ and $f_j(e_i)=\delta_{ij}$ for $0\le i\le d$. Proposition 1.15 tells us that $f_0,\dots,f_d$ is a $k[[t]]$-basis of Δ^*. For $\theta, r\ge 0$ and any i, j we have $f_j(t^{\theta}\varphi^r\otimes e_i)\neq 0$ if and only if $r\equiv j-i$ modulo $(d+1)\mathbb{Z}$ and $\theta=k_j+qk_{j-1}+\cdots+q^{r-1}k_{j-r+1}$. As before, $\psi\in\mathrm{End}_k(\Delta^*)$ is defined by $(\psi(f))(x)=f(\varphi(x))$ for $x\in\Delta$, $f\in\Delta^*$. We claim

$$\psi(t^{m+k_j+1}f_j)=\rho_{j-1}\psi_{k((t))}(t^m)tf_{j-1} \tag{18}$$

for all j, all $m\ge -k_j-1$. Indeed, for $0\le i\le d$ and $\theta, r\ge 0$ we have

$$(\psi(t^{m+k_j+1}f_j))(t^{\theta}\varphi^r\otimes e_i)=f_j(t^{m+k_j+1}\varphi t^{\theta}\varphi^r\otimes e_i).$$

If $m+1 \notin \mathbb{Z}q$ then this shows $(\psi(t^{m+k_j+1}f_j))(t^\theta\varphi^r \otimes e_i) = 0$ by what we pointed out above. But $m+1 \notin \mathbb{Z}q$ also implies $\psi_{k((t))}(t^m) = 0$. In the case $m+1 = qn$ (some $n \in \mathbb{Z}$) we compute

$$\begin{aligned}(\psi(t^{m+k_j+1}f_j))(t^\theta\varphi^r \otimes e_i) &= f_j(t^{k_j+qn}\varphi t^\theta\varphi^r \otimes e_i)\\ &= f_j(t^{k_j}\varphi t^{n+\theta}\varphi^r \otimes e_i)\\ &= \rho_{j-1}f_{j-1}(t^{n+\theta}\varphi^r \otimes e_i)\\ &= (\rho_{j-1}\psi_{k((t))}(t^m)tf_{j-1})(t^\theta\varphi^r \otimes e_i)\end{aligned}$$

where we used $\psi_{k((t))}(t^m) = t^{n-1}$. We have proven formula (18).

On the other hand, by tracing the construction in Proposition 1.4 we see that $\varphi(tf_{j-1})$ is characterized by satisfying

$$\psi(t^m\varphi(tf_{j-1})) = \psi_{k((t))}(t^m)tf_{j-1} \tag{19}$$

for all m. Comparing formulae (18) and (19) we find $\varphi(tf_{j-1}) = \rho_{j-1}^{-1}t^{k_j+1}f_j$ which is equivalent with formula (16). Next, for $\gamma \in \Gamma$ we compute

$$(\gamma \cdot f_j)(e_i) = f_j(\gamma^{-1} \cdot e_i) = f_j(\eta_i(\gamma^{-1})e_i) = (\eta_i(\gamma^{-1})f_j)(e_i) = (\eta_j(\gamma^{-1})f_j)(e_i).$$

Here the last equation is trivial if $i = j$, whereas if $i \neq j$ then both sides vanish. This shows $(\gamma \cdot f_j - \eta_j(\gamma^{-1})f_j)|_N = 0$, and hence $\gamma \cdot f_j - \eta_j(\gamma^{-1})f_j \in t\Delta^* = tk[\![t]\!]\{f_0, \ldots, f_d\}$. On the other hand, by what we pointed out above, $(\gamma \cdot f_j)(t^\theta\varphi^r \otimes e_i) = f_j([\gamma]_\Phi(t)^\theta\varphi^r \otimes e_i)$ vanishes if $r + i - j \notin (d+1)\mathbb{Z}$, and this shows $\gamma \cdot f_j \in k[\![t]\!]f_j$. We trivially have $\eta_j(\gamma^{-1})f_j \in k[\![t]\!]f_j$, and hence altogether $\gamma \cdot f_j - \eta_j(\gamma^{-1})f_j \in tk[\![t]\!]\{f_0, \ldots, f_d\} \cap k[\![t]\!]f_j = tk[\![t]\!]f_j$, formula (17).

Now we assume $F = \mathbb{Q}_p$. Let us suppose for simplicity that $\pi = q$. For $0 \leq j \leq d$ we may define $f_j \in \Delta^*$ as follows. For $\theta, r \geq 0$ (and any i, j) we require $f_j(t^\theta\varphi^r \otimes e_i) \neq 0$ if and only if $r \equiv j - i$ modulo $(d+1)\mathbb{Z}$ and there are $a_1, \ldots, a_{r-1} \in \{0, 1\}$ such that

$$\theta = k_j + qk_{j-1} + \cdots + q^{r-1}k_{j-r+1} + \sum_{i=1}^{r-1} a_i q^{i-1}(1-q);$$

if this is the case we put

$$f_j(t^\theta\varphi^r \otimes e_i) = \rho_{j-1}\rho_{j-2}\cdots\rho_{j-r}.$$

(As usual, the subindices of the $\rho_?$ are read modulo $(d+1)\mathbb{Z}$.) Again $f_0, \ldots, f_d$ is a $k[\![t]\!]$-basis of Δ^*. Again we claim formula (18). As before we see that both sides vanish if $m \notin \mathbb{Z}q - 1 \cup \mathbb{Z}q$, and coincide if $m \in \mathbb{Z}q - 1$. But the same computation also shows their coincidence if $m = qn$ for some $n \in \mathbb{N}$, as follows:

$$\begin{aligned}(\psi(t^{m+k_j+1}f_j))(t^\theta\varphi^r \otimes e_i) &= f_j(t^{k_j+1}\varphi t^{n+\theta}\varphi^r \otimes e_i)\\ &= \rho_{j-1}f_{j-1}(t^{n+\theta+1}\varphi^r \otimes e_i)\\ &= (\rho_{j-1}\psi_{k((t))}(t^m)tf_{j-1})(t^\theta\varphi^r \otimes e_i)\end{aligned}$$

where we used $\psi_{k((t))}(t^m) = t^n$. With formula (18) being established, the remaining arguments are exactly as before. □

Definition. We say that an object $\boldsymbol{D} \in \mathrm{Mod}^{\mathrm{et}}(k((t)))$ of dimension $d+1$ is standard cyclic if it admits a $k((t))$-basis $f_0, \dots, f_d$ such that there are $\sigma_j \in k^\times$, characters $\alpha_j : \Gamma \to k^\times$ and $m_j \in \{1-q, \dots, -1, 0\}$ for $0 \le j \le d$ satisfying the following conditions:

- $(m_0, \dots, m_d) \notin \{(0, \dots, 0), (1-q, \dots, 1-q)\}$.
- $\varphi(f_{j-1}) = \sigma_j t^{m_j} f_j$ for all j (reading $f_{-1} = f_d$).
- $\gamma \cdot f_j - \alpha_j(\gamma) f_j \in tk[\![t]\!]\{f_0, \dots, f_d\}$ for all $\gamma \in \Gamma$.

Lemma 1.19. (a) *The constant $\prod_{j=0}^d \sigma_j \in k^\times$ as well as, up to cyclic permutation, the ordered tuple $((\alpha_0, m_0), \dots, (\alpha_d, m_d))$, are uniquely determined by the isomorphism class of the (φ, Γ)-module $\boldsymbol{D}$.*

(b) *$\alpha_1, \dots, \alpha_d$ are uniquely determined by α_0 and $m_0, \dots, m_d$.*

Proof. (a) In the following, for elements of $\mathrm{GL}_{d+1}(k((t)))$ we read the (two) respective indices of their entries always modulo $(d+1)\mathbb{Z}$.

The effect of φ on the basis $f_0, \dots, f_d$ is described by $T = (T_{ij})_{0 \le i,j \le d} \in \mathrm{GL}_{d+1}(k((t)))$ with $T_{i,i+1} = \sigma_i t^{m_i}$ for $0 \le i \le d$, but $T_{i,j} = 0$ for $j \ne i+1$.

Let $\sigma'_j \in k^\times$ and $((\alpha'_0, m'_0), \dots, (\alpha'_d, m'_d))$ be another datum as above, let $\boldsymbol{D}'$ be an étale (φ, Γ)-module admitting a $k((t))$-basis $f'_0, \dots, f'_d$ with $\varphi(f'_{j-1}) = \sigma'_j t^{m'_j} f'_j$ and $\gamma \cdot f'_j - \alpha'_j(\gamma) f'_j \in tk[\![t]\!]\{f'_0, \dots, f'_d\}$ for $\gamma \in \Gamma$. Define $T' = (T'_{ij})_{0 \le i,j \le d} \in \mathrm{GL}_{d+1}(k((t)))$ similarly as above.

Suppose that there is an isomorphism of (φ, Γ)-modules $\boldsymbol{D}' \cong \boldsymbol{D}$. With respect to the bases $f_\bullet$ and $f'_\bullet$ it is described by some $A(t) = (a_{i,j}(t))_{0 \le i,j \le d} \in \mathrm{GL}_{d+1}(k((t)))$. In view of $\varphi t = \Phi(t)\varphi$, the compatibility of the isomorphism with the respective φ-actions comes down to the matrix equation

$$T \cdot A(t) = A(\Phi(t)) \cdot T'.$$

For the individual entries this is equivalent with

$$a_{i,j}(t) = \sigma'_j \sigma_i^{-1} t^{m'_j - m_i} a_{i-1,j-1}(\Phi(t))$$

for all i, j. Iteration of this equation yields

$$a_{i,j}(t) = \left(\prod_{\ell=0}^{d} \sigma'_{j-\ell} \sigma_{i-\ell}^{-1} (\Phi^\ell(t))^{m'_{j-\ell} - m_{i-\ell}} \right) a_{i,j}(\Phi^{d+1}(t))$$

for all i, j. (Here $\Phi^\ell(t)$ resp. $\Phi^{d+1}(t)$ means $\Phi(\Phi(\cdots \Phi(t) \cdots))$.) From this we deduce that for fixed i, j either $a_{i,j}$ is a nonzero constant and $\prod_{\ell=0}^d \sigma'_{j-\ell} \sigma_{i-\ell}^{-1} = 1$ and $m'_{j-\ell} = m_{i-\ell}$ for all ℓ, or $a_{i,j} = 0$. But since $A(t)$ is invertible we do find i, j with $a_{i,j} \ne 0$. It already follows that $\prod_{j=0}^d \sigma_j = \prod_{j=0}^d \sigma'_j$ and that $(m'_0, \dots, m'_d)$ coincides with $(m_0, \dots, m_d)$ up to cyclic permutation. But since in addition we just saw that A is a constant matrix, with $a_{i,j} = 0$ if and only if $a_{i-1,j-1} = 0$, we see that the same index permutation takes α'_j to α_j.

(b) This follows from the fact that, in view of the defining formulae, $\boldsymbol{D}$ is generated by f_0 as a φ-module over $k((t))$. □

Proposition 1.20. *The functor $\Delta \mapsto \Delta^* \otimes_{k[[t]]} k((t))$ induces a bijection between the set of standard cyclic objects in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ and the set of standard cyclic objects in $\mathrm{Mod}^{\mathrm{et}}(k((t)))$.*

Proof. $\Delta^* \otimes_{k[[t]]} k((t))$ for a standard cyclic object $\Delta \in \mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ is a standard cyclic object in $\mathrm{Mod}^{\mathrm{et}}(k((t)))$ by Proposition 1.18. With Lemma 1.19(a) we see that the assignment $\Delta \mapsto \Delta^* \otimes_{k[[t]]} k((t))$ is injective on standard cyclic objects in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$. It is also surjective: Proposition 1.18 (together with Lemma 1.19(b)) explicitly says how to convert the parameter data describing a standard cyclic object in $\mathrm{Mod}^{\mathrm{et}}(k((t)))$ into the parameter data describing a standard cyclic object in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$. □

Definition. A $(d+1)$-dimensional standard cyclic $\mathrm{Gal}(\overline{F}/F)$-representation is a $\mathrm{Gal}(\overline{F}/F)$-representation over k which corresponds, under the equivalence of categories in Theorem 1.2, to an object in $\mathrm{Mod}^{\mathrm{et}}(k((t)))$ of dimension $d+1$ which is standard cyclic.

2. Hecke algebras and supersingular modules

2A. ***The pro-p-Iwahori Hecke algebra $\mathcal{H}$.*** We introduce the pro-p-Iwahori Hecke algebra $\mathcal{H}$ of $\mathrm{GL}_{d+1}(F)$ with coefficients in k in a slightly unorthodox way, which however is well suited for our later constructions.

Let $\overline{T}$ be a free $\mathbb{Z}/(q-1)$-module of rank $d+1$. Then $\mathrm{Hom}(\Gamma, \overline{T})$ (with $\Gamma = \mathcal{O}_F^\times$) is also free of rank $d+1$ over $\mathbb{Z}/(q-1)$. We write the group law of $\overline{T}$ multiplicatively, but that of $\mathrm{Hom}(\Gamma, \overline{T})$ we write additively. Let $e^*, \alpha_1^\vee, \dots, \alpha_d^\vee$ be a $\mathbb{Z}/(q-1)$-basis of $\mathrm{Hom}(\Gamma, \overline{T})$. Put $\alpha_0^\vee = -\sum_{i=1}^d \alpha_i^\vee$. We let the symmetric group $\mathfrak{S}_{d+1}$ act on $\mathrm{Hom}(\Gamma, \overline{T})$ as follows. We think of $\mathfrak{S}_{d+1}$ as the permutation group of $\{0, 1, \dots, d\}$, generated by the transposition $s = (01) \in \mathfrak{S}_{d+1}$ and the cycle $\omega \in \mathfrak{S}_{d+1}$ with $\omega(i) = i+1$ for all $0 \le i \le d-1$. We then put

$$\omega \cdot e^* = e^* + \alpha_0^\vee, \quad \omega \cdot \alpha_0^\vee = \alpha_d^\vee \quad \text{and} \quad \omega \cdot \alpha_i^\vee = \alpha_{i-1}^\vee \quad \text{for } 1 \le i \le d.$$

If $d = 1$ we put

$$s \cdot e^* = e^* - \alpha_1^\vee, \quad s \cdot \alpha_i^\vee = -\alpha_i^\vee \quad \text{for } i = 0, 1,$$

but if $d \ge 2$ we put

$$s \cdot e^* = e^* - \alpha_1^\vee, \quad s \cdot \alpha_0^\vee = \alpha_0^\vee + \alpha_1^\vee, \quad s \cdot \alpha_1^\vee = -\alpha_1^\vee, \quad s \cdot \alpha_2^\vee = \alpha_1^\vee + \alpha_2^\vee, \quad s \cdot \alpha_i^\vee = \alpha_i^\vee \quad \text{for } 3 \le i \le d.$$

One easily checks that there is a unique action of $\mathfrak{S}_{d+1}$ on $\overline{T}$ such that for $\gamma \in \Gamma$ and $f \in \mathrm{Hom}(\Gamma, \overline{T})$ we have

$$\omega \cdot (f(\gamma)) = (\omega \cdot f)(\gamma) \quad \text{and} \quad s \cdot (f(\gamma)) = (s \cdot f)(\gamma).$$

Define $\alpha_1^\vee(\mathbb{F}_q^\times)$ to be the image of the composition $\mathbb{F}_q^\times \to \Gamma \xrightarrow{\alpha_1^\vee} \overline{T}$ where the first map is the Teichmüller homomorphism.

Definition. (a) The k-algebra $\mathcal{H}$ is generated by elements $T_\omega^{\pm 1}$, T_s and T_t for $t \in \overline{T}$, subject to the following relations (with $t, t' \in \overline{T}$):

$$T_sT_\omega T_sT_\omega^{-1}T_sT_\omega = T_\omega T_sT_\omega^{-1}T_sT_\omega T_s \qquad \text{if } d > 1, \tag{20}$$

$$T_sT_\omega^{-m}T_sT_\omega^{m} = T_\omega^{-m}T_sT_\omega^{m}T_s \qquad \text{for all } 1 < m < d, \tag{21}$$

$$T_s^2 = T_s\tau_s = \tau_sT_s \qquad \text{with } \tau_s = \sum_{t \in \alpha_1^\vee(\mathbb{F}_q^\times)} T_t, \tag{22}$$

$$T_\omega T_\omega^{-1} = 1 = T_\omega^{-1}T_\omega, , \tag{23}$$

$$T_\omega^{d+1}T_s = T_sT_\omega^{d+1}, \tag{24}$$

$$T_tT_{t'} = T_{t't}, \;\; T_{1_{\overline{T}}} = 1, \tag{25}$$

$$T_tT_\omega = T_\omega T_{\omega\cdot t}, \tag{26}$$

$$T_tT_s = T_sT_{s\cdot t}. \tag{27}$$

Notice that T_ω^{d+1} is central in $\mathcal{H}$.

(b) $\mathcal{H}_{\text{aff}}$ is the k-subalgebra of $\mathcal{H}$ generated by all T_t for $t \in \overline{T}$, by T_ω^{d+1}, T_ω^{-d-1} and by all $T_\omega^mT_sT_\omega^{-m}$ for $m \in \mathbb{Z}$.

(c) $\mathcal{H}^\flat$ is the quotient of $\mathcal{H}$ by the two sided ideal spanned by all elements $T_t - 1$ with $t \in \overline{T}$.

Caution. $\mathcal{H}_{\text{aff}}$ differs from the similarly denoted algebra in [Vignéras 2005]. (The difference is that here we include $(T_\omega^{d+1})^{\mathbb{Z}}$.)

Remark. Let $\overline{T}$ denote the subgroup of $G = \mathrm{GL}_{d+1}(F)$ consisting of diagonal matrices with entries in the image of the Teichmüller homomorphism $\mathbb{F}_q^\times \to \mathcal{O}_F^\times$. For $\gamma \in \Gamma$ let $\overline{\gamma}$ be its image in $\mathbb{F}_q^\times$. In $\overline{T}$ define the elements $e^*(\gamma) = \mathrm{diag}(\overline{\gamma}, 1_d)$ and $\alpha_i^\vee(\gamma) = \mathrm{diag}(1_{i-1}, \overline{\gamma}, \overline{\gamma}^{-1}, 1_{d-i})$ for $1 \le i \le d$. Define the elements $\omega = (\omega_{ij})_{0\le i,j\le d}$ and $s = (s_{ij})_{0\le i,j\le d}$ of G by $\omega_{d0} = \pi$ and $\omega_{i,i+1} = 1$ (for $0 \le i \le d-1$) and $\omega_{ij} = 0$ for all other pairs (i, j), resp. by $s_{10} = s_{01} = s_{ii} = 1$ for $i \ge 2$, and $s_{ij} = 0$ for all other pairs (i, j).

Let I_0 denote the pro-p-Iwahori subgroup of G for which $g = (g_{ij})_{0\le i,j\le d} \in G$ belongs to I_0 if and only if all the following conditions are satisfied: $g_{ij} \in \pi\mathcal{O}_F$ for $i > j$, and $g_{ij} \in \mathcal{O}_F$ for $i < j$, and $g_{ii} \in 1 + \pi\mathcal{O}_F$.

Claim. *The corresponding pro-p-Iwahori Hecke algebra $k[I_0\backslash G/I_0]^{\mathrm{op}} \cong \mathrm{End}_{k[G]}(\mathrm{ind}_{I_0}^G k)^{\mathrm{op}}$ is isomorphic with $\mathcal{H}$, in such a way that the double coset I_0gI_0 for $g \in \overline{T} \cup \{s, \omega\}$ corresponds to the element $T_g \in \mathcal{H}$.*

To prove this claim we use the description of $k[I_0\backslash G/I_0]^{\mathrm{op}}$ worked out by Vignéras [2005] (or rather we use the description of $k[I_0\backslash G/I_0]^{\mathrm{op}}$ which results from the description of $k[I_0\backslash G/I_0]$ given in [loc. cit.]).

Let T denote the maximal torus of diagonal matrices in G, let $N(T)$ be its normalizer in G. Let T_1 (resp. T_0) denote the subgroup of T consisting of diagonal matrices with entries in the kernel of $\mathcal{O}_F^\times \to \mathbb{F}_q^\times$ (resp. in $\mathcal{O}_F^\times$); thus $T_0/T_1 \cong \overline{T}$. For $0 \le i \le d$ define $s_i = \omega^{1-i}s\omega^{i-1}$. The (classes of) $s_0, s_1, \dots, s_d$ are the Coxeter generators of a Coxeter subgroup W_{aff} of $N(T)/T_0$, and $N(T)/T_0$ is generated by W_{aff} together with the element ω. The length function $\ell : W_{\text{aff}} \to \mathbb{Z}_{\ge 0}$ can be extended to a function $\ell : N(T)/T_0 \to \mathbb{Z}_{\ge 0}$

in such a way that $\ell(\omega)=0$. We again denote by ℓ the induced function $W^{(1)}=N(T)/T_1\to\mathbb{Z}_{\geq 0}$. For $w\in W^{(1)}$ and $w'\in N(T)$ lifting w, the double coset $I_0w'I_0$ only depends on w; we denote it by T_w. For $0\leq i\leq d$ let $\overline{T}_i$ be the image of one of the two cocharacters $\mathbb{F}_q^\times\to\overline{T}$ associated with s_i. (Here we identify $\overline{T}$ with the maximal torus of diagonal matrices in $\mathrm{GL}_{d+1}(\mathbb{F}_q)$. If $1\leq i\leq d$ then s_i is the simple reflection associated with the coroot $\alpha_i^\vee$, and $\alpha_i^\vee(\mathbb{F}_q^\times)=\overline{T}_i$.) Now, according to [Vignéras 2005], a k-basis of $k[I_0\backslash G/I_0]^{\mathrm{op}}$ is given by the set of all T_w for $w\in W^{(1)}$, and the multiplication is uniquely determined by the relations

$$T_wT_{w'}=T_{w'w}\qquad\text{for } w,w'\in W^{(1)} \text{ with } \ell(w)+\ell(w')=\ell(ww'),\tag{28}$$

$$T_{s_i}^2=T_{s_i}\tau_i\qquad\text{where } \tau_i=\sum_{t\in\overline{T}_i}T_t \text{ for } 0\leq i\leq d.\tag{29}$$

In the following we repeatedly use that conjugating these relations by powers of T_ω leads to similar relations (since $\ell(\omega)=0$). From formula (28) we first deduce $T_{s_i}=T_\omega^{i-1}T_sT_\omega^{1-i}$ and then that $T_\omega^{\pm1}$ and $T_s=T_{s_1}$ together with the elements T_t for $t\in\overline{T}$ generate $k[I_0\backslash G/I_0]^{\mathrm{op}}$ as a k-algebra. Next, from $s_is_{i-1}s_i=s_{i-1}s_is_{i-1}$ in W_{aff} (for $0\leq i\leq d$; if $i=0$ read $i-1=d$) we get $T_{s_i}T_{s_{i-1}}T_{s_i}=T_{s_{i-1}}T_{s_i}T_{s_{i-1}}$ by applying formula (28) twice, but this comes down to formula (20) (up to conjugation by a power of T_ω). Similarly from $s_is_j=s_js_i$ in W_{aff} for $0\leq i<j-1\leq d-1$ with $i+d>j$ we get $T_{s_i}T_{s_j}=T_{s_j}T_{s_i}$ by applying formula (28) twice, but this comes down to formula (21) (up to conjugation by a power of T_ω). Formula (29) for any i is a T_ω-power conjugate of formula (22). Finally, formulae (23), (24), (25), (26) and (27) are special instances of formula (28). Conversely, it is not hard to see that these, together with formulae (20), (21) and (22) suffice to generate all relations in $k[I_0\backslash G/I_0]^{\mathrm{op}}$. The claim is proven.

We add if I denotes the Iwahori subgroup of G containing I_0, then $\mathcal{H}^\flat$ becomes isomorphic with the Iwahori Hecke algebra $k[I\backslash G/I]^{\mathrm{op}}$.

Definition. A character $\chi:\mathcal{H}_{\mathrm{aff}}\to k$ is called *supersingular* if the following two conditions are both satisfied:

(a) There is an $m\in\mathbb{Z}$ with $\chi(T_\omega^mT_sT_\omega^{-m})=0$.

(b) There is an $m\in\mathbb{Z}$ with either $\chi(T_\omega^mT_sT_\omega^{-m})=-1$ or $\chi(T_\omega^m\tau_sT_\omega^{-m})=0$.[7]

Definition. (a) An $\mathcal{H}$-module M is called *standard supersingular* if it is isomorphic with $\mathcal{H}\otimes_{\mathcal{H}_{\mathrm{aff}},\chi}k.e$, where $\mathcal{H}_{\mathrm{aff}}$ acts on the one dimensional k-vector space $k.e$ through a supersingular character χ.

Equivalently, M is standard supersingular if and only if $M=\bigoplus_{0\leq m\leq d}T_\omega^m(M_1)$ with an $\mathcal{H}_{\mathrm{aff}}$-module M_1 of k-dimension 1 on which $\mathcal{H}_{\mathrm{aff}}$ acts through a supersingular character.[8]

(b) An irreducible $\mathcal{H}$-module is called *supersingular* if it is a subquotient of a standard supersingular $\mathcal{H}$-module.

[7] We have $\chi(T_\omega^m\tau_sT_\omega^{-m})=0$ if and only if $\chi(T_\omega^mT_tT_\omega^{-m})\neq1$ for some $t\in\alpha_1^\vee(\mathbb{F}_q^\times)$, if and only if $\chi(\alpha_{m+1}^\vee(\gamma))\neq1$ for some $\gamma\in\Gamma$.

[8] Then $\mathcal{H}_{\mathrm{aff}}$ acts on each $T_\omega^m(M_1)$ through a supersingular character.

A finite dimensional $\mathcal{H}$-module is called *supersingular* if each of its irreducible subquotients is supersingular.

More generally, an $\mathcal{H}$-module is called supersingular if it is the inductive limit of its finite dimensional $\mathcal{H}$-submodules and if each finite dimensional $\mathcal{H}$-submodule is supersingular.[9]

Remark. For nonzero finite dimensional $\mathcal{H}$-modules, the above definition of supersingularity is equivalent with the one given by Vignéras. This follows from the discussion in Section 6 of [Vignéras 2017]. There is also a notion of supersingularity for $\mathcal{H}$-modules which are not necessarily inductive limits of their finite dimensional submodules. In the present paper however, without further mentioning *all $\mathcal{H}$-modules will be assumed to be inductive limits of their finite dimensional submodules.*

Remark. In the literature on modules over Hecke algebras, the term *standard* module is occasionally used, but this usage is unrelated to our terminology.

2B. *The coverings $\mathcal{H}^{\sharp\sharp}$ and $\mathcal{H}^{\sharp}$ of $\mathcal{H}$.*

Definition. (a) Let $\mathcal{H}^{\sharp}$ denote the k-algebra generated by elements $T_\omega^{\pm 1}$, T_s and T_t for $t \in \overline{T}$, subject to

- the relations (22), (23), (25), (26),
- the relations (27) for $t = \alpha_i^\vee(\gamma)$ (all $0 \le i \le d$, $\gamma \in \Gamma$),
- the relation
$$T_\omega^{d+1} T_s^2 = T_s^2 T_\omega^{d+1}, \tag{30}$$
- the relations
$$T_t T_s^2 = T_s^2 T_t \quad \text{for all } t \in \overline{T}, \tag{31}$$
- the relations
$$T_s^2 T_\omega T_s^2 T_\omega^{-1} T_s^2 T_\omega = T_\omega T_s^2 T_\omega^{-1} T_s^2 T_\omega T_s^2 \qquad \text{if } d > 1, \tag{32}$$
$$T_s^2 T_\omega^{-m} T_s^2 T_\omega^{m} = T_\omega^{-m} T_s^2 T_\omega^{m} T_s^2 \qquad \text{for all } 1 < m < d. \tag{33}$$

(b) Let $\mathcal{H}^{\sharp\sharp}$ denote the k-algebra generated by the elements $T_\omega^{\pm 1}$, T_s and T_t for $t \in \overline{T}$, subject to

- the relations (22), (23), (25), (26),
- the relations (27) for $t = \alpha_i^\vee(\gamma)$ (all $0 \le i \le d$, $\gamma \in \Gamma$),
- the relations (31).

Lemma 2.1. *In $\mathcal{H}$ we have the relations* (30), (31), (32) *and* (33).

[9] It is easy to see that the irreducible subquotients of a supersingular $\mathcal{H}$-module are the irreducible subquotients of its finite dimensional $\mathcal{H}$-submodules.

Proof. It is immediate that the relations (27) and (24) imply the relations (31) and (30), respectively. For $1 < m < d$ and $t \in \alpha_1^\vee(\mathbb{F}_q)$ we have $s\omega^m \cdot t = \omega^m \cdot t$, hence $T_s \sum_{t \in \alpha_1^\vee(\mathbb{F}_q)} T_{\omega^m \cdot t} = \sum_{t \in \alpha_1^\vee(\mathbb{F}_q)} T_{\omega^m \cdot t} T_s$. The same applies with $-m$ instead of m, hence

$$T_s T_\omega^{-m} \tau_s T_\omega^m = T_\omega^{-m} \tau_s T_\omega^m T_s \quad \text{and} \quad T_s T_\omega^m \tau_s T_\omega^{-m} = T_\omega^m \tau_s T_\omega^{-m} T_s.$$

This, together with $T_s^2 = \tau_s T_s = T_s \tau_s$ (formula (22)), justifies (i) and (iii) in

$$T_s^2 T_\omega^{-m} T_s^2 T_\omega^m \overset{\text{(i)}}{=} \tau_s (T_\omega^{-m} \tau_s T_\omega^m) T_s T_\omega^{-m} T_s T_\omega^m \overset{\text{(ii)}}{=} \tau_s (T_\omega^{-m} \tau_s T_\omega^m) T_\omega^{-m} T_s T_\omega^m T_s \overset{\text{(iii)}}{=} T_\omega^{-m} T_s^2 T_\omega^m T_s^2,$$

whereas (ii) is justified by (21). We have shown (33). Finally, to see (32) comes down, using (22), (26) and (27), to comparing

$$T_\omega T_s^2 T_\omega^{-1} T_s^2 T_\omega T_s^2 = \Bigg(\sum_{t_1, t_2, t_3 \in \alpha_1^\vee(\mathbb{F}_q)} T_{\omega^{-1} \cdot t_1} T_{\omega^{-1} s\omega \cdot t_2} T_{\omega^{-1} s\omega s \omega^{-1} \cdot t_3} \Bigg) T_\omega T_s T_\omega^{-1} T_s T_\omega T_s,$$

$$T_s^2 T_\omega^{-1} T_s^2 T_\omega T_s^2 T_\omega = \Bigg(\sum_{t_1, t_2, t_3 \in \alpha_1^\vee(\mathbb{F}_q)} T_{t_1} T_{s\omega^{-1} \cdot t_2} T_{s\omega^{-1} s\omega \cdot t_3} \Bigg) T_s T_\omega T_s T_\omega^{-1} T_s T_\omega.$$

That these are equal follows from (20) and equality of the bracketed terms; for the latter observe $\omega s \omega^{-1} s \omega \cdot t = t$ for any $t \in \alpha_1^\vee(\mathbb{F}_q^\times)$. □

In view of Lemma 2.1 we have natural surjections of k-algebras

$$\mathcal{H}^{\sharp\sharp} \to \mathcal{H}^\sharp \to \mathcal{H} \to \mathcal{H}^\flat.$$

Remark. $\mathcal{H}^{\sharp\sharp}$ (and in particular $\mathcal{H}^\sharp$ and $\mathcal{H}$) is generated as a k-algebra by $T_\omega^{\pm 1}$, T_s and the $T_{e^*(\gamma)}$ for $\gamma \in \Gamma$.

Lemma 2.2. *There are unique k-algebra involutions ι of $\mathcal{H}$, $\mathcal{H}^\sharp$ and $\mathcal{H}^{\sharp\sharp}$ with*

$$\iota(T_\omega) = T_\omega, \quad \iota(T_s) = \tau_s - T_s, \quad \iota(T_t) = T_t \quad \textit{for } t \in \overline{T}.$$

Proof. This is a slightly tedious but straightforward computation. (For $\mathcal{H}$ see [Vignéras 2005, Corollary 2].) □

Remark. Besides ι consider the k-algebra involution β of $\mathcal{H}$, $\mathcal{H}^\sharp$ and $\mathcal{H}^{\sharp\sharp}$ given on generators by

$$\beta(T_\omega) = T_\omega^{-1}, \quad \beta(T_s) = T_s, \quad \beta(T_t) = T_{s\cdot t} \quad \text{for } t \in \overline{T}.$$

Moreover, for any automorphism $\mathfrak{o}$ of Γ there is an associated automorphism $\alpha_{\mathfrak{o}}$ of $\mathcal{H}$, $\mathcal{H}^\sharp$ and $\mathcal{H}^{\sharp\sharp}$ given on generators by

$$\alpha_{\mathfrak{o}}(T_\omega) = T_\omega, \quad \alpha_{\mathfrak{o}}(T_s) = T_s, \quad \alpha_{\mathfrak{o}}(T_{\partial(\gamma)}) = T_{\partial(\mathfrak{o}(\gamma))} \quad \text{for } \gamma \in \Gamma, \partial \in \operatorname{Hom}(\Gamma, \overline{T}).$$

Do ι, β and the $\alpha_{\mathfrak{o}}$ generate the automorphism group of $\mathcal{H}$ (resp. of $\mathcal{H}^\sharp$, resp. of $\mathcal{H}^{\sharp\sharp}$) modulo inner automorphisms?

Lemma 2.3. *Let M be an $\mathcal{H}^{\sharp\sharp}$-module. We have a direct sum decomposition*

$$M = M^{T_s=-\mathrm{id}} \bigoplus M^{T_s^2=0}.$$

Proof. One computes $\tau_s^2 = (q-1)\tau_s = -\tau_s$ and this shows $T_s = -\mathrm{id}$ on $\mathrm{im}(T_s^2)$ as well as $T_s^2 = 0$ on $\mathrm{im}(T_s^2 - \mathrm{id})$. □

Let $[0, q-2]^\Phi$ be the set of tuples $\epsilon = (\epsilon_i)_{0 \le i \le d}$ with $\epsilon_i \in \{0, \ldots, q-2\}$ and $\sum_{0 \le i \le d} \epsilon_i \equiv 0$ modulo $(q-1)$. We often read the indices as elements of $\mathbb{Z}/(d+1)$, thus $\epsilon_i = \epsilon_j$ for $i, j \in \mathbb{Z}$ whenever $i - j \in (d+1)\mathbb{Z}$. We let the symmetric group $\mathfrak{S}_{d+1}$ (generated by s, ω as before) act on $[0, q-2]^\Phi$ as follows:

$$(\omega \cdot \epsilon)_0 = \epsilon_d \quad \text{and} \quad (\omega \cdot \epsilon)_i = \epsilon_{i-1} \quad \text{for } 1 \le i \le d.$$

If $d = 1$ we put

$$(s \cdot \epsilon)_i = -\epsilon_i \quad \text{for } i = 0, 1,$$

but if $d \ge 2$ we put

$$(s \cdot \epsilon)_1 = -\epsilon_1, \quad (s \cdot \epsilon)_0 = \epsilon_0 + \epsilon_1, \quad (s \cdot \epsilon)_2 = \epsilon_1 + \epsilon_2, \quad (s \cdot \epsilon)_i = \epsilon_i \quad \text{for } 3 \le i \le d.$$[10]

Throughout we assume that all eigenvalues of the T_t for $t \in \overline{T}$ acting on an $\mathcal{H}^{\sharp\sharp}$-module belong to k.

Let M be an $\mathcal{H}^{\sharp\sharp}$-module. For $a \in [0, q-2]$ and $\epsilon = (\epsilon_i)_{0 \le i \le d} \in [0, q-2]^\Phi$ and $j \in \{0, 1\}$ put

$$\begin{aligned}
M^\epsilon &= \{x \in M \mid T^{-1}_{\alpha_i^\vee(\gamma)}(x) = \gamma^{\epsilon_i} x \text{ for all } \gamma \in \Gamma, \text{ all } 0 \le i \le d\},\\
M^\epsilon_{\underline{a}} &= \{x \in M^\epsilon \mid T_{e^*(\gamma)}(x) = \gamma^a x \text{ for all } \gamma \in \Gamma\},\\
M^\epsilon_{\underline{a}}[j] &= \{x \in M^\epsilon_{\underline{a}} \mid T_s^2(x) = jx\}.
\end{aligned}$$

The T_t for $t \in \overline{T}$ are of order divisible by $q-1$, hence are diagonalizable on the k-vector space M. Since they commute among each other and with T_s^2, we may simultaneously diagonalize all these operators (see Lemma 2.3 for T_s^2), hence

$$M = \bigoplus_{\epsilon, a, j} M^\epsilon_{\underline{a}}[j]. \tag{34}$$

Lemma 2.4. *For any $\epsilon \in [0, q-2]^\Phi$ and $a \in [0, q-2]$ we have*

$$T_\omega(M^\epsilon_{\underline{a}}) = M^{\omega\cdot\epsilon}_{\underline{a-\epsilon_0}} \quad \textit{and} \quad T_s(M^\epsilon) \subset M^{s\cdot\epsilon}.$$

If M is even an $\mathcal{H}$-module then

$$T_s(M^\epsilon_{\underline{a}}) \subset M^{s\cdot\epsilon}_{\underline{\epsilon_1 + a}}. \tag{35}$$

[10] Here and below we understand $-\epsilon_i$ to mean the representative in $[0, q-2]$ of the class of $-\epsilon_i$ in $\mathbb{Z}/(q-1)$, and similarly for $\epsilon_0 + \epsilon_1$ and $\epsilon_1 + \epsilon_2$.

Proof. $T_\omega(M^\epsilon) = M^{\omega\cdot\epsilon}$ and $T_s(M^\epsilon) \subset M^{s\cdot\epsilon}$ follow from formulas (26) and (27), respectively, for the $t = \alpha_i^\vee(\gamma)$. For the following computation recall that $\omega \cdot e^* = e^* + \alpha_0^\vee$: For $\gamma \in \Gamma$ and $x \in M_{\underline{a}}^\epsilon$ we have

$$T_{e^*(\gamma)} T_\omega(x) = T_\omega T_{(\omega\cdot e^*)(\gamma)}(x) = T_\omega T_{e^*(\gamma)} T_{\alpha_0^\vee(\gamma)}(x) = \gamma^{a-\epsilon_0} T_\omega(x).$$

This shows $T_\omega(M_{\underline{a}}^\epsilon) = M_{\underline{a-\epsilon_0}}^{\omega\cdot\epsilon}$. For formula (35) recall that $s \cdot e^* = e^* - \alpha_1^\vee$ and employ formula (27). $\square$

Any $x \in M$ can be uniquely written as

$$x = \sum_{a\in[0,q-2]} x_{\underline{a}} \quad \text{with } x_{\underline{a}} \in \sum_{\epsilon\in[0,q-2]^\Phi} M_{\underline{a}}^\epsilon.$$

Given $a \in \mathbb{Z}$ and $x \in M$, we write $x_{\underline{a}} = x_{\underline{\tilde{a}}}$ where $\tilde{a} \in [0, q-2]$ is determined by $a - \tilde{a} \in (q-1)\mathbb{Z}$.

Definition. (a) An $\mathcal{H}^\sharp$-module M is called *standard supersingular* if the $\mathcal{H}^\sharp$-action factors through $\mathcal{H}$, making it a standard supersingular $\mathcal{H}$-module.

(b) An irreducible $\mathcal{H}^\sharp$-module is called *supersingular* if it is a subquotient of a standard supersingular $\mathcal{H}^\sharp$-module. An $\mathcal{H}^\sharp$-module M is called *supersingular* if it is the inductive limit of finite dimensional $\mathcal{H}^\sharp$-modules and if each of its irreducible subquotients is supersingular.

(c) An $\mathcal{H}^{\sharp\sharp}$-module M is called *supersingular* if it satisfies the condition analogous to (b).

(d) A supersingular $\mathcal{H}^\sharp$-module is called $\sharp$*-supersingular* if for all $e \in M_{\underline{a}}^\epsilon[0]$ with $\epsilon_1 > 0$ we have

$$(T_s e)_{\underline{c+\epsilon_1+a}} = 0 \quad \text{for all } q-1-\epsilon_1 \le c \le q-2.$$

Lemma 2.5. (a) *An $\mathcal{H}$-module is supersingular if and only if it is supersingular when viewed as an $\mathcal{H}^\sharp$-module. A supersingular $\mathcal{H}$-module is $\sharp$-supersingular when viewed as an $\mathcal{H}^\sharp$-module.*

(b) *The category of supersingular $\mathcal{H}$-modules, the category of supersingular $\mathcal{H}^\sharp$-modules, the category of supersingular $\mathcal{H}^{\sharp\sharp}$-modules and the category of $\sharp$-supersingular $\mathcal{H}^\sharp$-modules are abelian.*

Proof. Statement (a) follows from formula (35). Statement (b) is clear from the definitions. $\square$

3. Reconstruction of supersingular $\mathcal{H}^\sharp$-modules

Given an $\mathcal{H}^\sharp$-module M together with a submodule M_0 such that M/M_0 is supersingular, we address the problem of reconstructing the $\mathcal{H}^\sharp$-module M from the $\mathcal{H}^\sharp$-modules M_0 and M/M_0 together with an additional set of data (intended to be sparse). Our proposed solution (Proposition 3.3) critically relies on the braid relations (32) and (33).

Lemma 3.1. *Let $B_0, \ldots, B_n$ be linear operators on a k-vector space M such that*

$$\begin{aligned} B_j^2 &= B_j && \textit{for all } 0 \le j \le n, \\ B_j B_{j'} B_j &= B_{j'} B_j B_{j'} && \textit{for all } 0 \le j', j \le n, \\ B_j B_{j'} &= B_{j'} B_j && \textit{for all } 0 \le j' < j \le n \textit{ with } j - j' \ge 2. \end{aligned}$$

Put $\beta = B_n \cdots B_1 B_0$ and let $x \in M$ with $\beta^m x = x$ for some $m \geq 1$. Then we have $B_j x = x$ for each $0 \leq j \leq n$.

Proof. We first claim

$$\beta B_{j+1} = B_j \beta \quad \text{for all } 0 \leq j < n. \tag{36}$$

Indeed,

$$\begin{aligned}\beta B_{j+1} &= B_n \cdots B_{j+2} B_{j+1} B_j B_{j-1} \cdots B_1 B_0 B_{j+1}\\ &= B_n \cdots B_{j+2} B_{j+1} B_j B_{j+1} B_{j-1} \cdots B_1 B_0\\ &= B_n \cdots B_{j+2} B_j B_{j+1} B_j B_{j-1} \cdots B_1 B_0\\ &= B_j \beta.\end{aligned}$$

Choose $\nu \geq 1$ with $m\nu \geq n$. For $0 \leq j \leq n$ we then compute

$$x \overset{\text{(i)}}{=} \beta^{m\nu} x = \beta^{n-j} \beta^{m\nu - n + j} x \overset{\text{(ii)}}{=} \beta^{n-j} B_n \beta^{m\nu-n+j} x \overset{\text{(iii)}}{=} B_j \beta^{n-j} \beta^{m\nu-n+j} x = B_j \beta^{m\nu} x \overset{\text{(iv)}}{=} B_j x,$$

where (i) and (iv) follow from the hypothesis $\beta^m x = x$, where (ii) follows from $B_n \beta = \beta$ and where (iii) follows from repeated application of formula (36). □

Proposition 3.2. *Let M be an $\mathcal{H}^\sharp$-module, let $M_0 \subset M$ be an $\mathcal{H}^\sharp$-submodule such that M/M_0 is supersingular. Let $\bar{x} \in (M/M_0)^\epsilon$ (some $\epsilon \in [0, q-2]^\Phi$) be such that $\bar{x}\{i\} = T_\omega^{i+1} \bar{x}$ is an eigenvector under T_s, for each $i \in \mathbb{Z}$. For liftings $x \in M$ of $\bar{x}$ put $x\{i\} = T_\omega^{i+1} x$:*

(a) *If the $\mathcal{H}^\sharp$-action on M factors through $\mathcal{H}$ then we may choose $x \in M^\epsilon$ such that for each i with $T_s(\bar{x}\{i\}) = 0$ and $(\omega^{i+1} \cdot \epsilon)_1 = 0$ we have $T_s(x\{i\}) = 0$.*

(b) *If the $\mathcal{H}^\sharp$-action on M factors through $\mathcal{H}$ then we may choose $x \in M^\epsilon$ such that for each i with $T_s(\bar{x}\{i\}) = -\bar{x}\{i\}$ we have $T_s(x\{i\}) = -x\{i\}$.*

(c) *We may choose $x \in M^\epsilon$ such that for each i with $T_s^2(\bar{x}\{i\}) = 0$ we have $T_s^2(x\{i\}) = 0$.*

(d) *We may choose $x \in M^\epsilon$ such that for each i with $T_s^2(\bar{x}\{i\}) = \bar{x}\{i\}$ we have $T_s^2(x\{i\}) = x\{i\}$.*

Proof. (a) Let $i_1 < \cdots < i_r$ be the increasing enumeration of the set of all $0 \leq i \leq d$ with $T_s T_\omega^{i+1}(\bar{x}) = 0$ and $(\omega^{i+1} \cdot \epsilon)_1 = 0$. Replacing M by its $\mathcal{H}^\sharp$-submodule generated by x and M_0 we may assume that M/M_0 is a subquotient of a standard supersingular $\mathcal{H}$-module, attached to a supersingular character $\chi : \mathcal{H}_{\text{aff}} \to k$. If we had $T_s T_\omega^{i+1}(\bar{x}) = 0$ and $(\omega^{i+1} \cdot \epsilon)_1 = 0$ for all $0 \leq i \leq d$ then this would mean $\chi(T_\omega^m T_s T_\omega^{-m}) = 0$ and $\chi(T_\omega^m \tau_s T_\omega^{-m}) \neq 0$ for all $m \in \mathbb{Z}$, in contradiction with the supersingularity of χ. Hence there is some $0 \leq i \leq d$ not occurring among $\{i_1, \ldots, i_r\}$. Thus, after a cyclic index shift, we may assume $i_r < d$.

Start with an arbitrary lift $x \in M^\epsilon$ of $\bar{x}$.

We claim that for any j with $0 \leq j \leq r$, after modifying x if necessary, we can achieve $T_s(x\{i_s\}) = 0$ for all s with $1 \leq s \leq j$. For $j = r$ this is the desired statement.

Let us illustrate the argument in the case $d = 1$ first. (This will logically not be needed for the general case. Notice e.g., that the subarguments (2) and (3) below are required only if $d > 1$.) Then we have $r = 1$ and $i_1 = 0$, and the claim for $j = 1$ states that there is some $\tilde{x} \in M^\epsilon$ lifting $\bar{x}$ with $T_s T_\omega(\tilde{x}) = 0$.

But indeed, $\tilde{x} = x + T_\omega^{-1} T_s T_\omega x$ works: First, $\tilde{x}$ lifts $\bar{x}$ because of $T_s T_\omega \bar{x} = 0$. Next, $\tilde{x}$ belongs to M^ϵ because of $T_\omega^{-1} T_s T_\omega x \in M^\epsilon$ (which follows from $x \in M^\epsilon$ and the assumption $(\omega^{i_1+1} \cdot \epsilon)_1 = 0$). Finally, $T_s T_\omega(\tilde{x}) = 0$, because $T_\omega \tilde{x} \in M^{\omega \cdot \epsilon} = M^{\omega^{i_1+1} \cdot \epsilon}$ and $(\omega^{i_1+1} \cdot \epsilon)_1 = 0$ imply $(T_s + T_s^2) T_\omega \tilde{x} = 0$.

Now let us consider the case of a general d. Induction on j. For $j = 0$ there is nothing to do. Now fix $1 \le j \le r$ and assume that x satisfies the condition for $j-1$, i.e., assume $T_s(x\{i_s\}) = 0$ for all s with $1 \le s \le j-1$. For $-1 \le i \le d$ and $0 \le m < j$ define inductively

$$x\{i\}_0 = x\{i\} = T_\omega^{i+1} x,$$
$$x\{i\}_{m+1} = T_\omega^{i - i_{j-m}} T_s(x\{i_{j-m}\}_m).$$

We establish several subclaims.

(1) $x\{i\}_m \in M^{\omega^{i+1} \cdot \epsilon}$.

For $m = 0$ there is nothing to do. Next, if the claim is true for an arbitrary m, then we have in particular $x\{i_{j-m}\}_m \in M^{\omega^{i_{j-m}+1} \cdot \epsilon}$. By assumption we know $(\omega^{i_{j-m}+1} \cdot \epsilon)_1 = 0$, which implies $T_s(M^{\omega^{i_{j-m}+1} \cdot \epsilon}) \subset M^{\omega^{i_{j-m}+1} \cdot \epsilon}$. Thus, we get $x\{i_{j-m}\}_{m+1} = T_s(x\{i_{j-m}\}_m) \in M^{\omega^{i_{j-m}+1} \cdot \epsilon}$. From this we get $x\{i\}_{m+1} = T_s(x\{i\}_m) \in M^{\omega^{i+1} \cdot \epsilon}$ for general i by applying powers of T_ω to $x\{i_{j-m}\}_{m+1}$.

(2) $T_s(x\{i_s\}_m) = 0$ for all $1 \le s \le j$ and all $0 \le m < j - s$.

We induct on m. For $m = 0$ this is true by induction hypothesis (on j). Now let $0 < m < j - s$ and assume that we know the claim for $m - 1$ instead of m. In particular we then know $T_s(x\{i_s\}_{m-1}) = 0$. We deduce

$$\begin{aligned} T_s(x\{i_s\}_m) &= T_s T_\omega^{i_s - i_{j-m+1}} T_s T_\omega^{i_{j-m+1} - i_s} T_\omega^{i_s - i_{j-m+1}} (x\{i_{j-m+1}\}_{m-1}) \\ &= T_s T_\omega^{i_s - i_{j-m+1}} T_s T_\omega^{i_{j-m+1} - i_s} (x\{i_s\}_{m-1}) \\ &= T_\omega^{i_s - i_{j-m+1}} T_s T_\omega^{i_{j-m+1} - i_s} T_s(x\{i_s\}_{m-1}) \\ &= 0 \end{aligned}$$

where we use the braid relation (21) (which applies since $|i_s - i_{j-m+1}| > 1$ and $i_r < d$). The induction on m is complete.

(3) $T_s(x\{i_s\}_m) = 0$ for all $1 \le s \le j$ and all $j - s + 1 < m \le j$.

We induct on $m+s-j$. The induction begins with $m+s-j = 2$. By (2) we know $T_s(x\{i_{j-m+1}\}_{m-2}) = 0$. Thus, if $i_{j-m+1} + 1 < i_{j-m+2}$, the same argument as in (2) shows $T_s(x\{i_{j-m+1}\}_{m-1}) = 0$ and hence $x\{i\}_m = 0$ for all i, and there is nothing more to do. If however $i_{j-m+1} + 1 = i_{j-m+2}$ we compute

$$\begin{aligned} T_s(x\{i_{j-m+2}\}_m) &= T_s T_\omega T_s T_\omega^{-1} T_s T_\omega (x\{i_{j-m+1}\}_{m-2}) \\ &= T_\omega T_s T_\omega^{-1} T_s T_\omega T_s (x\{i_{j-m+1}\}_{m-2}) \\ &= 0 \end{aligned}$$

where we use the braid relation (20). This settles the case $m + s - j = 2$. For $m + s - j > 2$ we now argue exactly as in (2) again: $T_s(x\{i_s\}_m) = 0$ implies $T_s(x\{i_s\}_{m+1}) = 0$. The induction is complete.

(4) $T_s(x\{i_{j-m}\}_m + x\{i_{j-m}\}_{m+1}) = 0$ for all $0 \le m < j$.

Indeed, by (1) and our assumption $(\omega^{i_{j-m}+1}\cdot\epsilon)_1=0$ we know that $x\{i_{j-m}\}_m$ is fixed under $T_{\alpha_1^\vee(\Gamma)}$ and hence is killed by $T_s^2+T_s$, as follows from the quadratic relation (22). As $x\{i_{j-m}\}_{m+1}=T_s(x\{i_{j-m}\}_m)$ this gives the claim.

(5) $\tilde{x}=\sum_{0\leq m\leq j}x\{-1\}_m$ lifts $\bar{x}$.

Indeed, we have $T_s(x\{i_j\})\in M_0$ by our defining assumption on i_j. It follows that $x\{-1\}_m\in M_0$ for all $m\geq 1$, hence $x-\tilde{x}\in M_0$.

(6) From (1) we deduce $\tilde{x}\{i\}\in M^{\omega^{i+1}\cdot\epsilon}$. Writing

$$\tilde{x}\{i_s\}=\left(\sum_{0\leq m<j-s}x\{i_s\}_m\right)+(x\{i_s\}_{j-s}+x\{i_s\}_{j-s+1})+\left(\sum_{j-s+1<m\leq j}x\{i_s\}_m\right)$$

we see that (2), (3) and (4) imply $T_s(\tilde{x}\{i_s\})=0$ for all s with $1\leq s\leq j$.

The induction on j is complete; we may substitute $\tilde{x}$ for the old x.

(b) Composing the given $\mathcal{H}$-module structure on M with the involution ι of Lemma 2.2 we get a new $\mathcal{H}$-module structure on M. Applying statement (a) to this new $\mathcal{H}$-module and then translating back via ι, we get statement (b). Notice that here, in contrast to the setting in (a), we *automatically* have $(\omega^{i+1}\cdot\epsilon)_1=0$ for each i with $T_s(\bar{x}\{i\})=-\bar{x}\{i\}$.

(c) Statement (c) is proved in the same way as statement (a), with the following minor modifications: each occurrence of T_s must be replaced by T_s^2, and in the definition of $x\{i\}_{m+1}$ the alternating sign $(-1)^{m+1}$ must be included, i.e.,

$$x\{i\}_{m+1}=(-1)^{m+1}T_\omega^{i-i_{j-m}}T_s^2(x\{i_{j-m}\}_m) \tag{37}$$

In particular, we then have $x\{i_{j-m}\}_{m+1}=-T_s^2(x\{i_{j-m}\}_m)$. In (2) and (3), the appeal to the braid relations (20), (21) must be replaced by an appeal to the braid relations (32), (33). In (4), the appeal to $T_s^2+T_s=0$ on vectors fixed under $T_{\alpha_1^\vee(\Gamma)}$ must be replaced by an appeal to $T_s^4-T_s^2=0$ (it is here where the alternating sign in the defining formula (37) is needed). Notice that here, in contrast to the setting in (a), we do not need to impose $(\omega^{i+1}\cdot\epsilon)_1=0$ for each i with $T_s^2(\bar{x}\{i\})=0$. (On the one hand, because of $T_s^2(M^\epsilon)\subset M^\epsilon$ for *any* ϵ the argument analogous to the one in (a)(1) carries over; on the other hand, because of $T_s^4-T_s^2=0$ on all of M the argument analogous to the one in (a)(4) carries over.)

(d) Composing the given $\mathcal{H}^\sharp$-module structure on M with the involution ι of Lemma 2.2 we get a new $\mathcal{H}^\sharp$-module structure on M. Applying statement (c) to this new $\mathcal{H}^\sharp$-module and then translating back via ι, we get statement (d). □

Proposition 3.3. *Let M be an $\mathcal{H}^\sharp$-module, let $M_0\subset M$ be an $\mathcal{H}^\sharp$-submodule such that M/M_0 is supersingular. The action of $\mathcal{H}^\sharp$ on M is uniquely determined by the following combined data*:

(a) *The action of $\mathcal{H}^\sharp$ on M_0 and on M/M_0.*

(b) *The action of $T_{e^*(\Gamma)}$ and of T_sT_ω on M.*

(c) *The restriction of* T_ω *to* $(T_sT_\omega)^{-1}(M_0)$, *i.e., the map*

$$\{x \in M \mid T_sT_\omega(x) \in M_0\} \xrightarrow{T_\omega} M.$$

(d) *The subspace* $\sum_{\substack{\epsilon\in[0,q-2]^\Phi \\ \epsilon_1=0}} M^\epsilon$ *of* M.

Proof. The k-algebra $\mathcal{H}^\sharp$ is generated by $T_{e^*(\Gamma)}$, by T_s and by $T_\omega^{\pm1}$. Therefore we only need to see that the action of T_s and T_ω on M can be reconstructed from the given data (a), (b), (c), (d). Exhausting M/M_0 step by step we may assume that M/M_0 is an irreducible supersingular $\mathcal{H}^\sharp$-module.

We first show that T_s is uniquely determined. For this we make constant use of Lemma 2.3 (and the decomposition (34)). As $T_s|_{M_0}$ is given to us, it is enough to show that for any nonzero $\bar{x}$ in M/M_0 with either $T_s(\bar{x}) = -\bar{x}$ or $T_s(\bar{x}) = 0$ we find some lifting $x \in M$ such that $T_s(x)$ can be reconstructed. Consider first the case $T_s(\bar{x}) = -\bar{x}$. By the quadratic relation (22) (see Lemma 2.3) we then have $\bar{x} \in \sum_{\substack{\epsilon\in[0,q-2]^\Phi \\ \epsilon_1=0}} (M/M_0)^\epsilon$, and using the datum (d) as well as our knowledge of the subspace T_sM (since $T_sM = T_sT_\omega M$ this is given to us in view of datum (b)), we lift $\bar{x}$ to some $x \in T_sM \cap \sum_{\substack{\epsilon\in[0,q-2]^\Phi \\ \epsilon_1=0}} M^\epsilon$ (use the decomposition (34)). For such x we have $T_s(x) = -x$. Now consider the case $T_s(\bar{x}) = 0$. An arbitrary lifting $x \in M$ of $\bar{x}$ then satisfies $T_s(x) \in M_0$, and $T_s(x)$ is determined by the given data as $T_s(x) = (T_sT_\omega)T_\omega^{-1}(x)$ (notice that the datum (c) is equivalent with the datum $T_s^{-1}(M_0) \xrightarrow{T_\omega^{-1}} M$).

To show that T_ω is uniquely determined, suppose that besides $T_\omega \in \mathrm{Aut}_k(M)$ there is another candidate $\tilde{T}_\omega \in \mathrm{Aut}_k(M)$ extending the data (a), (b), (c), (d) to another $\mathcal{H}^\sharp$-action on M.

We find and choose some nonzero $\bar{x} \in M/M_0$ such that $T_\omega^j(\bar{x})$ is an eigenvector under T_s, for each $j \in \mathbb{Z}$. For any $x \in M$ lifting $\bar{x}$ we have

$$T_\omega = \tilde{T}_\omega \quad \text{on } M_0 + k.T_\omega^{j-1}(x) \text{ if } T_sT_\omega^j(\bar{x}) = 0 \tag{38}$$

as both $\tilde{T}_\omega$ and T_ω respect the datum (c).

Let $i_0 < \cdots < i_n$ be the increasing enumeration of the set

$$\{0 \le i \le d \mid T_s^2T_\omega^i\bar{x} = T_\omega^i\bar{x}\}.$$

As M/M_0 is a subquotient of a standard supersingular $\mathcal{H}$-module, this set is not the full set $\{0 \le i \le d\}$. Applying a suitable power of T_ω and reindexing we may assume that 0 does not belong to this set, i.e., that $i_0 > 0$.

Choose a lifting $x \in M$ of $\bar{x}$ such that for each $i \in \{i_0, \ldots, i_n\} + \mathbb{Z}(d+1)$ we have $T_s^2T_\omega^ix = T_\omega^ix$. This is possible by Proposition 3.2. Put $z_0 = x$. For $i \ge 1$ put

$$z_i = \begin{cases} \tilde{T}_\omega z_{i-1} & i \notin \{i_0, \ldots, i_n\} + \mathbb{Z}(d+1), \\ T_s^2\tilde{T}_\omega z_{i-1} & i \in \{i_0, \ldots, i_n\} + \mathbb{Z}(d+1). \end{cases}$$

We claim

$$z_i = T_\omega^i x \tag{39}$$

for each $i \geq 0$. Induction on i. The case $i = 0$ is trivial. For $i \geq 1$ with $i \notin \{i_0, \dots, i_n\} + \mathbb{Z}(d+1)$ we compute

$$z_i = \tilde{T}_\omega z_{i-1} \stackrel{\text{(i)}}{=} T_\omega z_{i-1} \stackrel{\text{(ii)}}{=} T_\omega^i x$$

where in (i) we use statement (38) and in (ii) we use the induction hypothesis. For $i \geq 1$ with $i \in \{i_0, \dots, i_n\} + \mathbb{Z}(d+1)$ we compute

$$z_i = T_s^2 \tilde{T}_\omega z_{i-1} \stackrel{\text{(i)}}{=} T_s^2 T_\omega z_{i-1} \stackrel{\text{(ii)}}{=} T_\omega^i x$$

where in (i) we use the assumption $T_s T_\omega = T_s \tilde{T}_\omega$, and in (ii) we use the induction hypothesis $T_\omega z_{i-1} = T_\omega^i x$ and the assumption on x. The induction is complete. Put

$$B_{i_j} = \tilde{T}_\omega^{-i_j} T_s^2 \tilde{T}_\omega^{i_j}.$$

The relation (30) implies $B_{i_j} = \tilde{T}_\omega^{-i_j+(d+1)\nu} T_s^2 \tilde{T}_\omega^{i_j-(d+1)\nu}$ for each $\nu \in \mathbb{Z}$. Thus

$$(B_{i_n} \cdots B_{i_1} B_{i_0})^m x \stackrel{\text{(i)}}{=} \tilde{T}_\omega^{-m(d+1)} z_{m(d+1)} \stackrel{\text{(ii)}}{=} \tilde{T}_\omega^{-m(d+1)} T_\omega^{m(d+1)} x$$

for $m \geq 0$, where (i) follows from the definition of $z_{m(d+1)}$, whereas (ii) follows from formula (39). Choosing m large enough we may assume $T_\omega^{m(d+1)} x = x$ and $\tilde{T}_\omega^{m(d+1)} x = x$ (as T_ω and $\tilde{T}_\omega$ are automorphisms of a finite vector space); then

$$(B_{i_n} \cdots B_{i_1} B_{i_0})^m x = x.$$

The braid relations (32), (33) show that the B_{i_j} satisfy the hypotheses of Lemma 3.1 (in particular, the commutation $B_{i_0} B_{i_n} = B_{i_n} B_{i_0}$ if $n > 1$ follows from $i_0 > 0$). This Lemma now tells us $B_{i_j} \cdots B_{i_1} B_{i_0} x = x$ for each $0 \leq j \leq n$. But by the definition of the z_i this means

$$z_i = \tilde{T}_\omega^i x \tag{40}$$

for each $0 \leq i \leq d+1$. When compared with formula (39) this yields $T_\omega = \tilde{T}_\omega$ since M is generated as a k-vector space by M_0 together with the $T_\omega^i x$ (or: the $\tilde{T}_\omega^i x$) for $0 \leq i \leq d$. □

Remarks. The above proof of Proposition 3.3 shows the following:

(i) The subspace in (d) could be replaced by the subspace $\{x \in M \mid T_s^2(x) = x\}$.

(ii) If the $\mathcal{H}^\sharp$-action factors through an $\mathcal{H}$-action, then the datum (d) can be entirely left out (T_ω can then be reconstructed without a priori knowledge of T_s).

4. The functor

Here we define a functor $M \mapsto \Delta(M)$ from supersingular $\mathcal{H}^{\sharp\sharp}$-modules to torsion $k[[t]]$-modules with φ and Γ actions, as outlined in the introduction. Its entire content is encapsulated in the explicit formula for the elements $h(e)$ introduced below.

Let M be an $\mathcal{H}^{\sharp\sharp}$-module. View M as a $k[\![t]\!]$-module with $t = 0$ on M. Let Γ act on M by

$$\gamma \cdot x = T_{e^*(\gamma)}^{-1}(x)$$

for $\gamma \in \Gamma$, making M a $k[\![t]\!][\Gamma]$-module. We have an isomorphism of $k[\![t]\!][\varphi]$-modules

$$\mathfrak{O} \otimes_{k[\![t]\!][\Gamma]} M \cong k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M$$

and hence an action of $\mathfrak{O}$ on $k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M$.

For $e \in M_{\underline{a}}^{\epsilon}[j]$ (any $\epsilon \in [0, q-2]^{\Phi}$, any $a \in [0, q-2]$, any $j \in \{0, 1\}$) define the element

$$h(e) = \begin{cases} t^{\epsilon_1}\varphi \otimes T_{\omega}^{-1}(e) + 1 \otimes e + \sum_{c=0}^{q-2} t^c \varphi \otimes T_{\omega}^{-1}((T_s e)_{\underline{c+\epsilon_1+a}}) & j = 0, \\ t^{q-1}\varphi \otimes T_{\omega}^{-1}(e) + 1 \otimes e & j = 1 \end{cases}$$

of $k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M$. Define $\nabla(M)$ to be the $k[\![t]\!][\varphi]$-submodule of $k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M$ generated by the elements $h(e)$ for all $e \in M_{\underline{a}}^{\epsilon}[j]$ (all ϵ, a, j). Define

$$\Delta(M) = \frac{k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M}{\nabla(M)}.$$

Remark. If M is even an $\mathcal{H}$-module, then in view of formula (35) the definition of $h(e)$ simplifies to become

$$h(e) = \begin{cases} t^{\epsilon_1}\varphi \otimes T_{\omega}^{-1}(e) + 1 \otimes e + \varphi \otimes T_{\omega}^{-1}(T_s e) & j = 0, \\ t^{q-1}\varphi \otimes T_{\omega}^{-1}(e) + 1 \otimes e & j = 1. \end{cases}$$

In this case it is not necessary to split up M into eigenspaces under the action of $T_{e^*(\Gamma)}$, and the *notation* of many of the subsequent computations simplifies (no underlined subscripts are needed). However, they hardly simplify in mathematical complexity, not even if we restrict to $\mathcal{H}^{\flat}$-modules only (in which case always $\epsilon_1 = 0$ and $T_{e^*(\gamma)} = 1$).

Lemma 4.1. *Let $e \in M_{\underline{a}}^{\epsilon}[j]$. The integer*

$$k_e = \begin{cases} \epsilon_1 & j = 0, \\ q-1 & j = 1, \end{cases}$$

satisfies $k_e \equiv \epsilon_1$ modulo $(q-1)$.

Proof. $j = 1$ means $T_s^2(e) = e$, hence the claim follows from the relation (22). □

Lemma 4.2. *For $e \in M_{\underline{a}}^{\epsilon}[j]$ we have $\gamma \cdot h(e) = h(T_{e^*(\gamma)}^{-1}(e))$ for all $\gamma \in \Gamma$. In particular, $\nabla(M)$ is stable under the action of Γ, hence is an $\mathfrak{O}$-submodule of $k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} M$. Hence $\Delta(M)$ is even an $\mathfrak{O}$-module.*

Proof. First notice that $T_{e^*(\gamma)}^{-1}(e) \in M_{\underline{a}}^{\epsilon}[j]$. In particular, $h(T_{e^*(\gamma)}^{-1}(e))$ is well defined. For $\gamma \in \Gamma$ we find

$$\gamma \cdot (1 \otimes e) = 1 \otimes \gamma \cdot e = 1 \otimes T_{e^*(\gamma)}^{-1}(e). \tag{41}$$

Next, we compute

$$\gamma \cdot (t^{k_e}\varphi \otimes T_{\omega}^{-1}(e)) \overset{\text{(i)}}{=} \gamma^{k_e} t^{k_e}\varphi \otimes \gamma \cdot T_{\omega}^{-1}(e) \overset{\text{(ii)}}{=} t^{k_e}\varphi \otimes T_{\omega}^{-1} T_{e^*(\gamma)}^{-1}(e). \tag{42}$$

In (i) we used $\gamma t = [\gamma]_\Phi(t)\gamma$ and $[\gamma]_\Phi(t) \equiv \gamma t$ modulo $t^q k[\![t]\!]$ (Lemma 0.1) and the fact that, since $\pi = 0$ in k, we have $t^q\varphi \otimes M = \Phi(t)\varphi \otimes M = \varphi t \otimes M = 0$. To see (ii) observe

$$\begin{aligned}\gamma \cdot T_\omega^{-1}(e) &= T_{e^*(\gamma)}^{-1} T_\omega^{-1}(e)\\ &= T_\omega^{-1} T_{(\omega^{-1}\cdot e^*)(\gamma)}^{-1}(e)\\ &= T_\omega^{-1} T_{(e^*-\alpha_1^\vee)(\gamma)}^{-1}(e)\\ &= T_\omega^{-1} T_{\alpha_1^\vee(\gamma)} T_{e^*(\gamma)}^{-1}(e)\\ &= \gamma^{-k_e} T_\omega^{-1} T_{e^*(\gamma)}^{-1}(e)\end{aligned}$$

where in the last step we use Lemma 4.1. Combining formulae (41) and (42) we are done in the case $j = 1$. In the case $j = 0$ we in addition need the formula

$$\gamma \cdot \sum_{c=0}^{q-2} t^c\varphi \otimes T_\omega^{-1}((T_s e)_{\underline{c+\epsilon_1+a}}) = \sum_{c=0}^{q-2} t^c\varphi \otimes T_\omega^{-1}((T_s T_{e^*(\gamma)} e)_{\underline{c+\epsilon_1+a}}). \tag{43}$$

Let us prove this (for $e \in M_{\underline{a}}^{\epsilon}[0]$). For $f \in \mathbb{Z}$ and $\gamma \in \Gamma$ we compute

$$\begin{aligned}T_{(\omega^{-1}\cdot e^*)(\gamma)}((T_s e)_{\underline{f}}) &\overset{\text{(i)}}{=} T_{e^*(\gamma)} T_{\alpha_1^\vee(\gamma^{-1})}((T_s e)_{\underline{f}})\\ &\overset{\text{(ii)}}{=} \gamma^{f-\epsilon_1}(T_s e)_{\underline{f}}\\ &= \gamma^{f-\epsilon_1-a}(T_s(\gamma^a e))_{\underline{f}}\\ &= \gamma^{f-\epsilon_1-a}(T_s T_{e^*(\gamma)} e)_{\underline{f}}.\end{aligned} \tag{44}$$

In (i) recall that $\omega^{-1}\cdot e^* = e^* - \alpha_1^\vee$, in (ii) notice that $(T_s e)_{\underline{f}} \in M^{s\cdot\epsilon}$ and $(s\cdot\epsilon)_1 = -\epsilon_1$. For $c \in [0, q-2]$ we deduce

$$\begin{aligned}\gamma \cdot (t^c\varphi \otimes T_\omega^{-1}((T_s e)_{\underline{c+\epsilon_1+a}})) &= \gamma^c t^c \varphi \otimes \gamma \cdot (T_\omega^{-1}((T_s e)_{\underline{c+\epsilon_1+a}}))\\ &= \gamma^c t^c\varphi \otimes T_{e^*(\gamma)}^{-1} T_\omega^{-1}((T_s e)_{\underline{c+\epsilon_1+a}})\\ &= \gamma^c t^c \varphi \otimes T_\omega^{-1} T_{(\omega^{-1}\cdot e^*)(\gamma)}^{-1}((T_s e)_{\underline{c+\epsilon_1+a}})\\ &= t^c\varphi \otimes T_\omega^{-1}((T_s T_{e^*(\gamma)} e)_{\underline{c+\epsilon_1+a}})\end{aligned}$$

where in the last equality we inserted formula (44). □

Proposition 4.3. (a) *If M is supersingular and finite dimensional, then we have: $\Delta(M)$ is a torsion $k[\![t]\!]$-module, generated by M as a $k[\![t]\!][\varphi]$-module, and φ acts injectively on it. The dual $\Delta(M)^* = \mathrm{Hom}_k(\Delta(M), k)$ is a free $k[\![t]\!]$-module of rank $\dim_k(M)$. The map $M \to \Delta(M)$ which sends $m \in M$ to the class of $1 \otimes m$ induces a bijection*

$$M \cong \Delta(M)[t]. \tag{45}$$

(b) *$\Delta(M)$ belongs to* $\mathrm{Mod}^{\clubsuit}(\mathfrak{O})$.

(c) *The assignment $M \mapsto \Delta(M)$ is an exact functor from the category of supersingular $\mathcal{H}^{\sharp\sharp}$-modules to* $\mathrm{Mod}^{ad}(\mathfrak{O})$.

Proof. (a) Notice first that it is enough to prove these claims after a finite base extensions of k.

Assume first that M is irreducible. It can then be realized as a subquotient of a standard supersingular $\mathcal{H}$-module N — in fact, it can even be realized as a submodule or as a quotient of such an N. Observing the decomposition (34) for N, we see that there are a k-basis $e_0, \ldots, e_d$ of N as well as $0 \le k_{e_j} \le q-1$ for $0 \le j \le d$, not all of them $= 0$ and not all of them $= q-1$, such that $\nabla(N)$ is generated by the elements $h(e_j) = t^{k_{e_j}} \varphi \otimes T_\omega^{-1}(e_j) + 1 \otimes e_j$. It follows that $\Delta(N)$ is standard cyclic. Now it is easy to see that $\Delta(M)$ is a subquotient of $\Delta(N)$. Thus, by Proposition 1.15(d), it is standard cyclic as well, at least after a finite extension of k. Therefore all our claims follow from Lemma 1.14 and Proposition 1.15(a).

Now let M be arbitrary (supersingular, finite dimensional). Choose a separated and exhausting descending filtration of M by $\mathcal{H}^{\sharp\sharp}$-submodules $F^\mu M$ with irreducible subquotients $F^{\mu-1}M/F^\mu M$. Since on any standard supersingular $\mathcal{H}$-module (and hence on any of its subquotients, and hence on any irreducible $\mathcal{H}^{\sharp\sharp}$-module) we have $T_s = -T_s^2$ and hence $\ker(T_s^2) = \ker(T_s)$, the filtration satisfies

$$T_s(F^{\mu-1}M \cap \ker(T_s^2)) \subset F^\mu M \tag{46}$$

for each $\mu \in \mathbb{Z}$. Putting

$$\boldsymbol{F}^\mu = k[\![t]\!][\varphi] \otimes_{k[\![t]\!]} F^\mu M$$

we claim

$$\nabla(F^\mu M) = \nabla(M) \cap \boldsymbol{F}^\mu. \tag{47}$$

Arguing by induction, we may assume that this is known with $\mu - 1$ instead of μ. Let $\mathcal{E}$ be a family of elements $e \in (F^{\mu-1}M)^{\epsilon_e}_{\underline{a_e}}[j_e]$ (for suitable $\epsilon_e \in [0, q-2]^\Phi$ and $a_e \in [0, q-2]$ and $j_e \in \{0, 1\}$ depending on e) which induces a k-basis of $F^{\mu-1}M/F^\mu M$. We consider an expression

$$\sum_{j_1, j_2 \in \mathbb{Z}_{\ge 0}, e \in \mathcal{E}} c_{j_1, j_2, e} t^{j_2} \varphi^{j_1} h(e) \tag{48}$$

with $c_{j_1, j_2, e} \in k$. Assuming that the expression (48) belongs to $\boldsymbol{F}^\mu$ we need to see that it even belongs to $\nabla(F^\mu M)$.

Suppose that this is false. We may then define

$$j_1 = \min\{j \ge 0 \mid c_{j, j_2, e} t^{j_2} \varphi^j h(e) \notin \nabla(F^\mu M) \text{ for some } j_2 \ge 0, \text{ some } e \in \mathcal{E}\}.$$

Claim. *We find some j_2 and some e with $c_{j_1, j_2, e} t^{j_2} \varphi^{j_1} h(e) \in \boldsymbol{F}^\mu - \nabla(F^\mu M)$.*

For $e \in \mathcal{E}$ the expression

$$1 \otimes e + t^{k_e} \varphi \otimes T_\omega^{-1}(e) \tag{49}$$

is congruent to $h(e)$ modulo $\boldsymbol{F}^\mu$, in view of $e \in F^{\mu-1}M$ and formula (46). Therefore, modulo $\boldsymbol{F}^\mu$ the expression (48) reads

$$\sum_{j_1, j_2, e} c_{j_1, j_2, e} t^{j_2} \varphi^{j_1} \otimes e + c_{j_1, j_2, e} t^{j_2} \varphi^{j_1} t^{k_e} \varphi \otimes T_\omega^{-1}(e).$$

Notice that $\varphi^{j_1}t^{k_e}\varphi\in k[\![t]\!]\varphi^{j_1+1}$. The claim now follows in view of

$$\frac{\boldsymbol{F}^{\mu-1}}{\boldsymbol{F}^{\mu}}=\bigoplus_{j\geq 0}k[\![t]\!]\varphi^j\otimes_{k[\![t]\!]}\frac{F^{\mu-1}M}{F^{\mu}M}. \tag{50}$$

The claim proven, we may argue by induction on the number of summands in the expression (48) which do not belong to $\nabla(F^{\mu}M)$. We may thus assume from the start that the expression (48) consists of a single summand $t^{j_2}\varphi^{j_1}h(e)$, and that moreover $e\notin F^{\mu}M$ for this e. The aim is then to deduce $t^{j_2}\varphi^{j_1}h(e)\in\nabla(F^{\mu}M)$, which contradicts our above assumption.

Let us write $\epsilon=\epsilon_e$ and $a=a_e$. The vanishing of $t^{j_2}\varphi^{j_1}h(e)$ modulo $\boldsymbol{F}^{\mu}$ means, by the decomposition (50) again, that

$$t^{j_2}\varphi^{j_1}\otimes e\overset{\text{(i)}}{=}0\overset{\text{(ii)}}{=}t^{j_2}\varphi^{j_1}t^{k_e}\varphi\otimes T_{\omega}^{-1}(e)$$

(i.e., absolute vanishing, not just modulo $\boldsymbol{F}^{\mu}$). If $T_s^2(e)=e$ then this shows $t^{j_2}\varphi^{j_1}h(e)=0$. Now suppose $T_s^2(e)=0$ (and hence $k_e<q-1$). The definition of $h(e)$ together with the vanishings (i) and (ii) shows

$$t^{j_2}\varphi^{j_1}h(e)=t^{j_2}\varphi^{j_1}\sum_{c=0}^{q-2}t^c\varphi\otimes T_{\omega}^{-1}((T_se)_{\underline{c+\epsilon_1+a}}).$$

Since the vanishing (ii) also forces $t^{j_2}\varphi^{j_1}t^{k_e}\varphi\in k[\![t]\!]\varphi^{j_1+1}t$, there is some i and some $j_2'\geq 0$ with

$$t^{j_2}\varphi^{j_1}=t^{j_2'}\varphi^{j_1}t^i\quad\text{and}\quad i\geq q-k_e.$$

If $k_e=0$ (and hence $i\geq q$) then again the conclusion is $t^{j_2}\varphi^{j_1}h(e)=0$. It remains to discuss the case where $0<k_e<q-1$. In this case, $(T_se)_{\underline{c+\epsilon_1+a}}\in M^{s\cdot\epsilon}$ and $(s\cdot\epsilon)_1=-\epsilon_1$ implies $q-1-k_e=k_{(T_se)_{\underline{c+\epsilon_1+a}}}$ for each c. We thus see

$$\begin{aligned}t^{q-k_e+c}\varphi\otimes T_{\omega}^{-1}((T_se)_{\underline{c+\epsilon_1+a}})&=t^{1+c}(t^{k_{(T_se)_{\underline{c+\epsilon_1+a}}}}\varphi\otimes T_{\omega}^{-1}((T_se)_{\underline{c+\epsilon_1+a}})+1\otimes(T_se)_{\underline{c+\epsilon_1+a}})\\&=t^{1+c}h((T_se)_{\underline{c+\epsilon_1+a}})-\sum_{c'=0}^{q-2}t^{1+c+c'}\varphi\otimes T_{\omega}^{-1}((T_s((T_se)_{\underline{c+\epsilon_1+a}}))_{\underline{c'+c+a}})\end{aligned}$$

by the definition of $h((T_se)_{\underline{c+\epsilon_1+a}})$, again since $(T_se)_{\underline{c+\epsilon_1+a}}\in M^{s\cdot\epsilon}$ and $(s\cdot\epsilon)_1=-\epsilon_1$. For $0\leq f\leq q-2$ we have

$$\sum_{\substack{0\leq c,c'\leq q-2\\ c+c'=f}}(T_s((T_se)_{\underline{c+\epsilon_1+a}}))_{\underline{f+a}}=\sum_{0\leq c\leq q-2}(T_s((T_se)_{\underline{c+\epsilon_1+a}}))_{\underline{f+a}}=0$$

as follows from $T_s^2(e)=0$. This shows

$$\sum_{c,c'=0}^{q-2}t^{1+c+c'}\varphi\otimes T_{\omega}^{-1}((T_s((T_se)_{\underline{c+\epsilon_1+a}}))_{\underline{c'+c+a}})=0.$$

Since e belongs to $F^{\mu-1}M$, formula (46) shows $h((T_se)_{\underline{c+\epsilon_1+a}})\in\nabla(F^{\mu}M)$. Together we obtain $t^{q-k_e+c}\varphi\otimes T_{\omega}^{-1}((T_se)_{\underline{c+\epsilon_1+a}})\in\nabla(F^{\mu}M)$, hence $t^{i+c}\varphi\otimes T_{\omega}^{-1}((T_se)_{\underline{c+\epsilon_1+a}})\in\nabla(F^{\mu}M)$ for $0\leq c\leq q-2$.

This gives

$$t^{j_2}\varphi^{j_1}h(e) = \sum_{c=0}^{q-2} t^{j_2'}\varphi^{j_1}t^{i+c}\varphi \otimes T_\omega^{-1}((T_s e)_{\underline{c+\epsilon_1+a}}) \in \nabla(F^\mu M),$$

as desired.

Formula (47) is proven. It allows us to deduce all our claims for M from the corresponding claims for the $F^{\mu-1}M/F^\mu M$; but for them they have already been established above.

(b) For each irreducible supersingular $\mathcal{H}$-module M, extending k if necessary, $\Delta(M)$ admits a filtration such that each associated graded piece is a standard cyclic object in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$, as pointed out above. Since the functor Δ is exact (see statement (c)) it therefore takes finite dimensional supersingular $\mathcal{H}^{\sharp\sharp}$-modules to objects in $\mathrm{Mod}^{\clubsuit}(\mathfrak{O})$.

(c) It is clear that $M \mapsto \Delta(M)$ is a (covariant) right exact functor. To see left exactness, let $M_1 \to M_2$ be injective. Since the kernel of $\Delta(M_1) \to \Delta(M_2)$ is a torsion $k[\![t]\!]$-module it has, if nonzero, a nonzero vector killed by t. By formula (45) it must belong to (the image of) M_1, contradicting the injectivity of $M_1 \to M_2$. □

5. Standard objects and full faithfulness

5A. ***The bijection between standard supersingular Hecke modules and standard cyclic Galois representations.*** Let M be a standard supersingular $\mathcal{H}$-module, arising from the supersingular character $\chi : \mathcal{H}_{\mathrm{aff}} \to k$. There is some $e_0 \in M$ such that, putting $e_j = T_\omega^{-j}e_0$, we have $M = \bigoplus_{j=0}^d k.e_j$ and $\mathcal{H}_{\mathrm{aff}}$ acts on $k.e_0$ by χ. Denote by $\eta_j : \Gamma \to k^\times$ the character through which $T_{e^*(.)}^{-1}$ acts on $k.e_j$, i.e., $T_{e^*(\gamma)}^{-1}(e_j) = \eta_j(\gamma)e_j$ for $\gamma \in \Gamma$.

Lemma 5.1. (a) *There are $0 \le k_{e_j} \le q-1$ for $0 \le j \le d$, not all of them $=0$ and not all of them $=q-1$, such that*

$$t^{k_{e_j}}\varphi \otimes T_\omega^{-1}(e_j) = -1 \otimes e_j \tag{51}$$

in $\Delta(M)$ for all $0 \le j \le d$.

(b) *If for any $1 \le j \le d$ there is some $0 \le i \le d$ with $k_{e_i} \ne k_{e_{i+j}}$, then $\Delta(M)$ is irreducible as a $k[\![t]\!][\varphi]$-module.*

(c) *Suppose that for any $1 \le j \le d$ which satisfies $k_{e_i} = k_{e_{i+j}}$ for all $0 \le i \le m$ there is some $0 \le i \le d$ with $\eta_i \ne \eta_{i+j}$. Then $\Delta(M)$ is irreducible as an $\mathfrak{O}$-module.*

Proof. For M as above, $\nabla(M)$ is generated by elements of the form $h(e) = t^{k_e}\varphi \otimes T_\omega^{-1}(e) + 1 \otimes e$. They give rise to formula (51), hence statement (a). For statements (b) and (c) apply Proposition 1.15; in (c) notice that $\gamma \cdot (1 \otimes e_j) = \eta_j(\gamma) \otimes e_j$ for $\gamma \in \Gamma$. □

Lemma 5.2. (a) *Conjugating χ by powers of T_ω means cyclically permuting the ordered tuple $((\eta_0, k_{e_0}), \dots, (\eta_d, k_{e_d}))$ associated with χ as above. Knowing the conjugacy class of χ (under powers of T_ω) is equivalent with knowing the tuple $((\eta_0, k_{e_0}), \dots, (\eta_d, k_{e_d}))$ up to cyclic permutations, together with $\chi(T_\omega^{d+1})$.*

(b) *(Vignéras) Two standard supersingular $\mathcal{H}$-modules are isomorphic if and only if the element $T_\omega^{d+1} \in \mathcal{H}$ acts on them by the same constant in $k^\times$ and if they arise from two supersingular characters $\mathcal{H}_{\mathrm{aff}} \to k$ which are conjugate under some power of T_ω.*

(c) *(Vignéras) A standard supersingular $\mathcal{H}$-module M arising from χ is simple if and only if the orbit of χ under conjugation by powers of T_ω has cardinality $d+1$.*

Proof. Statement (a) is clear. For (b) and (c) see [Vignéras 2005, Proposition 3 and Theorem 5]. □

Proposition 5.3. *The functor $M \mapsto \Delta(M)$ induces a bijection between the set of isomorphism classes of standard supersingular $\mathcal{H}$-modules and the set of standard cyclic objects in $\mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ of k-dimension $d+1$. If the standard supersingular $\mathcal{H}$-module M is simple, then $\Delta(M) \in \mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$ is simple.*

Proof. This follows from Lemmas 5.1 and 5.2. □

Theorem 5.4. (1) *The functor $M \mapsto \Delta(M)^* \otimes_{k[\![t]\!]} k((t))$ induces a bijection between the set of isomorphism classes of standard supersingular $\mathcal{H}$-modules and the set of isomorphism classes of standard cyclic étale (φ, Γ)-modules of dimension $d+1$.*

(2) *The functor $M \mapsto \Delta(M)^* \otimes_{k[\![t]\!]} k((t))$ induces a bijection between the set of isomorphism classes of simple supersingular $\mathcal{H}$-modules of k-dimension $d+1$ and the set of isomorphism classes of simple étale (φ, Γ)-modules of dimension $d+1$.*

Proof. Statement (a) follows from Propositions 1.20 and 5.3. Statement (b) follows from statement (a) and the full faithfulness of the functor $M \mapsto \Delta(M)^* \otimes_{k[\![t]\!]} k((t))$ on supersingular $\mathcal{H}$-modules, see Theorem 5.11 below. (To see that if M is simple then so is $\Delta(M)^* \otimes_{k[\![t]\!]} k((t))$ one may alternatively use Proposition 5.3 together with Proposition 1.11.) □

Remark. We may rewrite (51) as

$$t^{k_{e_j}} \varphi \otimes e_{j+1} = -1 \otimes e_j \quad \text{for } 0 \le j \le d-1$$
$$t^{k_{e_d}} \varphi \otimes \chi(T_\omega^{-d-1}) e_0 = -1 \otimes e_d$$

where we used $T_\omega^{-1}(e_d) = T_\omega^{-d-1}(e_0) = \chi(T_\omega^{-d-1}) e_0$. Thus $(-1)^{d+1} \chi(T_\omega^{-d-1}) \in k^\times$ is the constant referred to in Lemma 1.19.

Corollary 5.5. *The functor $M \mapsto \Delta(M)^* \otimes_{k[\![t]\!]} k((t))$, composed with the functor of Theorem 1.2, induces a bijection between the set of isomorphism classes of standard supersingular $\mathcal{H}$-modules of k-dimension $d+1$ and the set of isomorphism classes of $(d+1)$-dimensional standard cyclic $\mathrm{Gal}(\overline{F}/F)$-representations.*

Proof. Theorem 5.4. □

Remark. (a) Combining Corollary 5.5 and Theorem 5.4 one can derive the following "numerical Langlands correspondence": the set of (absolutely) simple $(d+1)$-dimensional $\mathcal{H}$-modules with fixed scalar action by T_ω^{d+1} has the same cardinality as the set of (absolutely) irreducible $(d+1)$-dimensional $\mathrm{Gal}(\overline{F}/F)$-representations with fixed determinant of Frobenius. This numerical Langlands correspondence was proven already in [Vignéras 2005, Theorem 5].

(b) There is an alternative and arguably more natural definition of supersingularity for $\mathcal{H}$-modules. Its agreement with the one given in Section 2A, and hence the "numerical Langlands correspondence" with respect to this alternative definition of supersingularity, was proven in [Ollivier 2010].

5B. ***Reconstruction of an initial segment of M from $\mathbf{\Delta}(M)$.*** Let $[0, q-1]^\Phi$ be the set of tuples $\mu = (\mu_i)_{0\le i\le d}$ with $\mu_i \in \{0, \dots, q-1\}$ and $\sum_{0\le i\le d} \mu_i \equiv 0$ modulo $(q-1)$. We often read the indices as elements of $\mathbb{Z}/(d+1)$, thus $\mu_i = \mu_j$ for $i, j \in \mathbb{Z}$ whenever $i - j \in (d+1)\mathbb{Z}$.

Let Δ be an $\mathfrak{O}$-module. For $\mu \in [0, q-1]^\Phi$ let $\mathcal{F}\Delta[t]^\mu$ be the k-subvector space of $\Delta[t] = \{x \in \Delta \mid tx = 0\}$ generated by all $x \in \Delta[t]$ satisfying $t^{\mu_i}\varphi \cdots t^{\mu_1}\varphi t^{\mu_0}\varphi x \in \Delta[t]$ for all $0 \le i \le d$, as well as $t^{\mu_d}\varphi \cdots t^{\mu_1}\varphi t^{\mu_0}\varphi x \in k^\times x$.

Put $\mathcal{F}\Delta[t] = \sum_{\mu\in[0,q-1]^\Phi} \mathcal{F}\Delta[t]^\mu$ (sum in $\Delta[t]$).

Lemma 5.6. *$\mathcal{F}\Delta[t] = \bigoplus_{\mu\in[0,q-1]^\Phi} \mathcal{F}\Delta[t]^\mu$, i.e., the sum is direct.*

Proof. Consider the lexicographic enumeration $\mu(1), \mu(2), \mu(3), \dots$ of $[0, q-1]^\Phi$ such that for each pair $r' > r$ there is some $0 \le i_0 \le d$ with $\mu_i(r) \ge \mu_i(r')$ for all $i < i_0$, and $\mu_{i_0}(r) > \mu_{i_0}(r')$. Let $\sum_{r\ge1} x_r = 0$ with $x_r \in \mathcal{F}\Delta[t]^{\mu(r)}$. We prove $x_r = 0$ for all r by induction on r. So, fix r and assume $x_{r'} = 0$ for all $r' < r$, hence $\sum_{r'\ge r} x_{r'} = \sum_{r\ge1} x_r - \sum_{r'<r} x_r = 0$. For $r' > r$ we have $t^{\mu_d(r)}\varphi \cdots t^{\mu_0(r)}\varphi(x_{r'}) = 0$. Therefore

$$\begin{aligned} 0 &= t^{\mu_d(r)}\varphi \cdots t^{\mu_0(r)}\varphi\Big(\sum_{r'\ge r} x_{r'}\Big) \\ &= \sum_{r'\ge r} t^{\mu_d(r)}\varphi \cdots t^{\mu_0(r)}\varphi x_{r'} \\ &= t^{\mu_d(r)}\varphi \cdots t^{\mu_0(r)}\varphi x_r \in k^\times x_r \end{aligned}$$

and hence $x_r = 0$. □

We define k-linear endomorphisms T_ω, T_s and $T_{e^*(\gamma)}$ (for $\gamma \in \Gamma$) of $\mathcal{F}\Delta[t]$ as follows. In view of Lemma 5.6 it is enough to define their values on $x \in \mathcal{F}\Delta[t]^\mu$; we put

$$T_\omega(x) = -t^{\mu_0}\varphi x, \quad T_{e^*(\gamma)}(x) = \gamma^{-1} \cdot x, \quad T_s(x) = \begin{cases} -x & \mu_d = q-11 \\ 0 & \mu_d < q-1. \end{cases}$$

Here $\gamma^{-1} \cdot x$ is understood with respect to the Γ-action induced by the $\mathfrak{O}$-module structure on $\Delta(M)$.

Definition. For an $\mathcal{H}^{\sharp\sharp}$-module M and $\mu \in [0, q-1]^{\Phi}$ let $\mathcal{F}M^{\mu}$ denote the k-subvector space of M consisting of $x \in M$ satisfying the following conditions for all $0 \le i \le d$:

$$T^{-1}_{\alpha_1^{\vee}(\gamma)}(T^i_{\omega}(x)) = \gamma^{\mu_{i-1}} T^i_{\omega}(x) \quad \text{for all } \gamma \in \Gamma, \tag{52}$$

$$T_s(T^i_{\omega}(x)) = \begin{cases} -T^i_{\omega}(x) & \mu_{i-1} = q-1, \\ 0 & \mu_{i-1} < q-1. \end{cases} \tag{53}$$

Let $\mathcal{F}M$ denote the subspace of M generated by the $\mathcal{F}M^{\mu}$ for all $\mu \in [0, q-1]^{\Phi}$.

For $\mu \in [0, q-1]^{\Phi}$ let $\epsilon_{\mu} \in [0, q-2]^{\Phi}$ be the unique element with

$$(\epsilon_{\mu})_{-i} \equiv \mu_i \mod (q-1). \tag{54}$$

for all i.

Lemma 5.7. (a) *We have $\mathcal{F}M^{\mu} \subset M^{\epsilon_{\mu}}$.*

(b) *$\mathcal{F}M$ is an $\mathcal{H}^{\sharp\sharp}$-submodule of M.*

(c) *$\mathcal{F}M$ contains each $\mathcal{H}^{\sharp\sharp}$-submodule of M which is a subquotient of a standard supersingular $\mathcal{H}^{\sharp\sharp}$-module.*

(d) *Suppose that M is supersingular. Viewing the isomorphism $\Delta(M)[t] \cong M$ (Proposition 4.3) as an identity, we have $\mathcal{F}M^{\mu} \subset \mathcal{F}\Delta(M)[t]^{\mu}$ for each $\mu \in [0, q-1]^{\Phi}$, and in particular*

$$\mathcal{F}M \subset \mathcal{F}\Delta(M)[t]. \tag{55}$$

The operators T_{ω}, T_s and $T_{e^*(\gamma)}$ acting on $\mathcal{F}\Delta(M)[t]$ as defined above restrict to the operators T_{ω}, T_s, $T_{e^*(\gamma)} \in \mathcal{H}^{\sharp\sharp}$ acting on $\mathcal{F}M$.

Proof. (a) Let $\mu \in [0, q-1]^{\Phi}$. For $x \in \mathcal{F}M^{\mu}$, any $\gamma \in \Gamma$ and any i we compute

$$T^{-1}_{\alpha^{\vee}_{1-i}(\gamma)}(x) = T^{-1}_{(\omega^i \cdot \alpha_1^{\vee})(\gamma)}(x) = T^{-i}_{\omega} T^{-1}_{\alpha_1^{\vee}(\gamma)} T^i_{\omega}(x) = \gamma^{\mu_{i-1}} x = \gamma^{(\epsilon_{\mu})_{1-i}} x,$$

i.e., $x \in M^{\epsilon_{\mu}}$.

(b) Let $\mu \in [0, q-1]^{\Phi}$ and define $\mu' \in [0, q-1]^{\Phi}$ by $\mu'_i = \mu_{i+1}$ for all i. For $x \in \mathcal{F}M^{\mu}$, any $\gamma \in \Gamma$ and any i we compute

$$T^{-1}_{\alpha_1^{\vee}(\gamma)}(T^i_{\omega}(T_{\omega}(x))) = T^{-1}_{\alpha_1^{\vee}(\gamma)}(T^{i+1}_{\omega}(x)) = \gamma^{\mu_i} T^{i+1}_{\omega}(x) = \gamma^{\mu_i} T^i_{\omega}(T_{\omega}(x)).$$

We also find $T_s(T^i_{\omega}(T_{\omega}(x))) = T_s(T^{i+1}_{\omega}(x)) = -T^{i+1}_{\omega}(x) = -T^i_{\omega}(T_{\omega}(x))$ if $\mu_i = q-1$, but $T_s(T^i_{\omega}(T_{\omega}(x))) = T_s(T^{i+1}_{\omega}(x)) = 0$ if $\mu_i < q-1$. Together this shows $T_{\omega}(x) \in \mathcal{F}M^{\mu'}$, i.e., $T_{\omega}(\mathcal{F}M^{\mu}) \subset \mathcal{F}M^{\mu'}$. It is immediate from the definitions that $T_s(\mathcal{F}M^{\mu}) \subset \mathcal{F}M^{\mu}$. For $x \in \mathcal{F}M^{\mu}$, any $\gamma, \gamma' \in \Gamma$ and any i we compute

$$T^{-1}_{\alpha_1^{\vee}(\gamma)} T^i_{\omega}(T_{e^*(\gamma')}(x)) = T^{-1}_{\alpha_1^{\vee}(\gamma)} T_{(\omega^{-i} \cdot e^*)(\gamma')} T^i_{\omega}(x) = T_{(\omega^{-i} \cdot e^*)(\gamma')} \gamma^{\mu_{i-1}} T^i_{\omega}(x) = \gamma^{\mu_{i-1}} T^i_{\omega}(T_{e^*(\gamma')}(x)).$$

If $\mu_{i-1} = q-1$ we also compute

$$\begin{aligned} T_s T_\omega^i (T_{e^*(\gamma')}(x)) &= T_{(s\cdot\omega^{-i}\cdot e^*)(\gamma')} T_s T_\omega^i(x) \\ &= -T_{(s\cdot\omega^{-i}\cdot e^*)(\gamma')} T_\omega^i(x) \\ &= -T_\omega^i (T_{(\omega^i\cdot s\cdot\omega^{-i}\cdot e^*)(\gamma')}(x)) \\ &= -T_\omega^i (T_{e^*(\gamma')}(x)). \end{aligned}$$

Here, in the last equation we use $\omega^i \cdot s \cdot \omega^{-i} \cdot e^* = e^*$ for $2 \le i \le d$; for $i = 1$ we use $\omega \cdot s \cdot \omega^{-1} \cdot e^* - e^* = \alpha_0^\vee$ and $T_{\alpha_0^\vee(\gamma')}(x) = T_\omega^{-1} T_{\alpha_1^\vee(\gamma')} T_\omega(x) = \gamma^{-\mu_0} x = x$ (as $\mu_0 = q-1$); for $i = 0$ we use $s \cdot e^* - e^* = -\alpha_1^\vee$ and $T_{-\alpha_1^\vee(\gamma')}(x) = \gamma^{\mu_{-1}} x = x$ (as $\mu_{-1} = q-1$). If however $\mu_{i-1} < q-1$ then $T_s T_\omega^i (T_{e^*(\gamma')}(x)) = T_{(s\cdot\omega^{-i}\cdot e^*)(\gamma')} T_s T_\omega^i(x) = 0$. Together this shows $T_{e^*(\gamma')}(x) \in \mathcal{F}M^\mu$, i.e., $T_{e^*(\gamma')}(\mathcal{F}M^\mu) \subset \mathcal{F}M^\mu$.

(c) On a standard supersingular $\mathcal{H}^{\sharp\sharp}$-module, and hence on its subquotients, the actions of T_ω, T_s and $T_{\alpha_1^\vee(\gamma)}$ satisfy formulae (52) and (53), for suitable μ's.

(d) Let $\mu \in [0, q-1]^\Phi$ and define $\mu' \in [0, q-1]^\Phi$ by $\mu'_i = \mu_{i+1}$ for all i. Let $x \in \mathcal{F}M^\mu$. The proof of (b) shows $T_\omega(x) \in \sum_a M_{\underline{a}}^{\epsilon_{\mu'}}[0]$ if $\mu'_d = \mu_0 < q-1$, resp. $T_\omega(x) \in \sum_a M_{\underline{a}}^{\epsilon_{\mu'}}[1]$ if $\mu'_d = \mu_0 = q-1$. In either case, the definition of $\Delta(M)$ then says $T_\omega(x) = -t^{\mu_0}\varphi x$. This shows $\mathcal{F}M^\mu \subset \mathcal{F}\Delta(M)[t]^\mu$ and that the action of T_ω on $\mathcal{F}M$ is indeed as stated. For the actions of T_s and $T_{e^*(\gamma)}$ this is clear anyway. □

Remark. The inclusion (55) is in fact an equality.

5C. *Reconstruction of $\sharp$-supersingular $\mathcal{H}^\sharp$-modules M from $\Delta(M)$.*

Lemma 5.8. *Let M be an irreducible supersingular $\mathcal{H}$-module. Let $\mu \in [0, q-1]^\Phi$, $x \in M$ and $u_{i,c} \in M^{\omega^{-1}s\omega^{i+1}\cdot\epsilon_\mu}$ for $i \ge 0$ and $0 \le c \le q-2$ (with ϵ_μ given by formula (54)). Assume $u_{i,c} = 0$ if*

(i) $\mu_i = 0$, *or*

(ii) $\mu_i = q-1$ *and* $c > 0$, *or*

(iii) $\mu_i < q-1$ *and* $c \ge q-1-\mu_i$.

Assume that, if we put $x\{-1\} = x$, then

$$x\{i\} = t^{\mu_i}\varphi(x\{i-1\}) - \sum_{c=0}^{q-2} t^c \varphi u_{i,c}$$

belongs to $M \cong \Delta(M)[t]$ for each $i \ge 0$. Finally, assume that $x\{D\} = x$ for some $D > 0$ with $D+1 \in \mathbb{Z}(d+1)$. Then there is some $x' \in M$ with $x - x' \in M^{\epsilon_\mu}$ and such that

$$x'\{i\} = t^{\mu_i}\varphi(\cdots(t^{\mu_1}\varphi(t^{\mu_0}\varphi x'))\cdots)$$

belongs to M for each i, and $x'\{D\} = x'$. Moreover, if x is an eigenvector for $T_{e^*(\Gamma)}$, then x' can be chosen to be an eigenvector for $T_{e^*(\Gamma)}$, with the same eigenvalues.

Proof. It is easy to see that all the irreducible subquotients of a standard supersingular $\mathcal{H}$-module are isomorphic. In particular, an irreducible supersingular $\mathcal{H}$-module is isomorphic with a submodule of a standard supersingular $\mathcal{H}$-module. Therefore we may assume that M itself is a (not necessarily irreducible) standard supersingular $\mathcal{H}$-module. We then have a direct sum decomposition $M = \oplus_{j=0}^{d} M^{[j]}$ with $\dim_k(M^{[j]}) = 1$ and integers $0 \le k_j \le q-1$ such that

$$T_\omega(M^{[j+1]}) = t^{k_j}\varphi(M^{[j+1]}) = M^{[j]} \tag{56}$$

(always reading j modulo $(d+1)$). More precisely, we have $M^{[j]} \subset M^{\epsilon_j}$ for certain $\epsilon_j \in [0, q-2]^\Phi$, and choosing the above k_j minimally, we have $k_j \equiv (\omega \cdot \epsilon_{j+1})_1$ modulo $(q-1)$. It follows that

$$k[t]\varphi M = \bigoplus_{j=0}^{d} k[t]\varphi M^{[j]} = \bigoplus_{j=0}^{d} \bigoplus_{c=0}^{k_j} t^c \varphi M^{[j+1]}. \tag{57}$$

For $m \in M$ write $m = \sum_j m^{[j]}$ with $m^{[j]} \in M^{[j]}$. By formulae (56), (57), the defining formula for $x\{i\}$ splits up into the formulae

$$x\{i\}^{[j]} = t^{\mu_i}\varphi(x\{i-1\}^{[j+1]}) - \sum_{c=0}^{q-2} t^c \varphi(u_{i,c}^{[j+1]}) \tag{58}$$

for all j. We use them to show

$$t^c\varphi(u_{i,c}^{[j+1]}) = 0 \quad \text{ if } c - \mu_i \notin (q-1)\mathbb{Z}. \tag{59}$$

If $\mu_i \in \{0, q-1\}$ then formula (59) follows from our assumptions on the $u_{i,c}$. Now assume $\mu_i \notin \{0, q-1\}$ and $u_{i,c}^{[j+1]} \neq 0$ for some c. The assumption $u_{i,c} \in M^{\omega^{-1}s\omega^{i+1}\cdot\epsilon_\mu}$ implies $T_\omega(u_{i,c}^{[j+1]}) \in M^{s\omega^{i+1}\cdot\epsilon_\mu}$, and since

$$q-1-\mu_i = q-1-\epsilon_{-i} = (s\omega^{i+1}\cdot\epsilon_\mu)_1 \quad \text{ if } \mu_i \notin \{0, q-1\}$$

we get $T_\omega(u_{i,c}^{[j+1]}) = -t^{q-1-\mu_i}\varphi(u_{i,c}^{[j+1]})$, i.e., $k_j = q-1-\mu_i$. Now $\sum_{c=0}^{k_j} t^c\varphi M^{[j+1]}$ is a *direct* sum of one dimensional k-vector spaces, with $x\{i\}^{[j]} \in t^{k_j}\varphi M^{[j+1]}$, $t^{\mu_i}\varphi(x\{i-1\}^{[j+1]}) \in t^{\mu_i}\varphi M^{[j+1]}$ and $t^c\varphi(u_{i,c}^{[j+1]}) \in t^c\varphi M^{[j+1]}$ for all c. Since by assumption $u_{i,c} = 0$ for $c \ge q-1-\mu_i = k_j$, formula (58) shows $t^c\varphi(u_{i,c}^{[j+1]}) = 0$ whenever $c \neq \mu_i$.

Formula (59) is proven. Arguing once more with formulae (56), (57) and (58) shows

$$[t^{\mu_i}\varphi(x\{i-1\}^{[j+1]}) = 0 \text{ or } \varphi(u_{i,0}^{[j+1]}) = 0] \quad \text{if } \mu_i = q-1. \tag{60}$$

In the following, by $u_{i,q-1}$ we mean $u_{i,0}$. If $t^{\mu_i}\varphi(u_{i,\mu_i}^{[j+1]}) \neq 0$ we may write

$$t^{\mu_i}\varphi(x\{i-1\}^{[j+1]}) - t^{\mu_i}\varphi(u_{i,\mu_i}^{[j+1]}) = \rho_{i,j} t^{\mu_i}\varphi(u_{i,\mu_i}^{[j+1]})$$

for some $\rho_{i,j} \in k$, since $t^{\mu_i}\varphi(x\{i-1\}^{[j+1]})$ and $t^{\mu_i}\varphi(u_{i,\mu_i}^{[j+1]})$ belong to the same one-dimensional k-vector space. The upshot of formulae (59) and (60) is then that formula (58) simplifies to become either

$$x\{i\}^{[j]} = t^{\mu_i}\varphi(x\{i-1\}^{[j+1]}) \tag{61}$$

or

$$x\{i\}^{[j]} = \rho_{i,j} t^{\mu_i} \varphi(u_{i,\mu_i}^{[j+1]}) \tag{62}$$

for some $\rho_{i,j} \in k$. Departing from $x^{[j]} = x\{D\}^{[j]}$ we repeatedly substitute formula (61); if this is possible $D+1$ many times we end up with

$$x^{[j]} = x\{D\}^{[j]} = t^{\mu_D} \varphi(\cdots (t^{\mu_1} \varphi(t^{\mu_0} \varphi(x^{[j]}))) \cdots),$$

and in this case we put $n(j) = 0$. Otherwise, after $D+1-n(j)$ many substitutions of formula (61), for some $1 \le n(j) \le D+1$, we end the procedure by substituting formula (62) (once) and obtain

$$x^{[j]} = x\{D\}^{[j]} = \rho_j t^{\mu_D} \varphi(\cdots (t^{\mu_{n(j)}} \varphi(t^{\mu_{n(j)-1}} \varphi(u_{n(j)-1,\mu_{n(j)-1}}^{[j+1-n(j)]}))) \cdots)$$

with $t^{\mu_{n(j)-1}} \varphi u_{n(j)-1,\mu_{n(j)-1}}^{[j+1-n(j)]} \neq 0$, for some $\rho_j \in k$.

We study this second case $n(j) > 0$ further. By construction,

$$w_j\{-1\} = t^{\mu_{n(j)-1}} \varphi(u_{n(j)-1,\mu_{n(j)-1}}^{[j+1-n(j)]})$$

is nonzero and belongs to M. On the other hand, $u_{n(j)-1,\mu_{n(j)-1}} \in M^{\omega^{-1} s \omega^{n(j)} \cdot \epsilon_\mu}$ implies $T_\omega(u_{n(j)-1,\mu_{n(j)-1}}^{[j+1-n(j)]}) \in M^{s\omega^{n(j)} \cdot \epsilon_\mu}$ and hence

$$t^{(s\omega^{n(j)} \cdot \epsilon_\mu)_1} \varphi(u_{n(j)-1,\mu_{n(j)-1}}^{[j+1-n(j)]}) = -T_\omega(u_{n(j)-1,\mu_{n(j)-1}}^{[j+1-n(j)]}) \in M^{s\omega^{n(j)} \cdot \epsilon_\mu}.$$

Together this means $\mu_{n(j)-1} \equiv (s\omega^{n(j)} \cdot \epsilon_\mu)_1$ modulo $(q-1)$ and $w_j\{-1\} \in M^{s\omega^{n(j)} \cdot \epsilon_\mu}$. But we also have $\mu_{n(j)-1} \equiv (\omega^{n(j)} \cdot \epsilon_\mu)_1$. Combining we see $\mu_{n(j)-1} \equiv -\mu_{n(j)-1}$ modulo $(q-1)$. Hence, we either have $\mu_{n(j)-1} = 0$ or $\mu_{n(j)-1} = (q-1)/2$ or $\mu_{n(j)-1} = q-1$. In view of the assumed vanishings of the $u_{i,c}$ (and of $u_{n(j)-1,\mu_{n(j)-1}}^{[j+1-n(j)]} \neq 0$) this leaves $\mu_{n(j)-1} = q-1$ as the only possibility. It follows that

$$s\omega^{n(j)} \cdot \epsilon_\mu = \omega^{n(j)} \cdot \epsilon_\mu$$

and hence $w_j\{-1\} \in M^{\omega^{n(j)} \cdot \epsilon_\mu}$. Next, again by construction we know that

$$w_j\{s\} = t^{\mu_{n(j)+s}} \varphi(w_j\{s-1\})$$

belongs to M, for $0 \le s \le D - n(j)$. By what we learned about $w_j\{-1\}$ this implies $w_j\{s\} = (-1)^{s+1} T_\omega^{s+1} w_j\{-1\} \in M^{\omega^{n(j)+s+1} \cdot \epsilon_\mu}$ by an induction on s (and we also see $\mu_{n(j)+s} \in \{k_0, \ldots, k_d\}$ with the k_ℓ from formula (56)). For $s = D - n(j)$ we obtain $x^{[j]} = x\{D\}^{[j]} \in M^{\epsilon_\mu}$.

We now put $x' = \sum_{n(j)=0} x^{[j]}$. □

Lemma 5.9. *Let M be an irreducible supersingular $\mathcal{H}$-module. Let $\mu \in [0, q-1]^\Phi$ and $x \in M$ such that*

$$x\{i\} = t^{\mu_i} \varphi(\cdots (t^{\mu_1} \varphi(t^{\mu_0} \varphi x)) \cdots)$$

belongs to $M \cong \Delta(M)[t]$ for each $i \ge 0$, and such that $x\{D\} = x$ for some $D > 0$ with $D+1 \in \mathbb{Z}(d+1)$. Then $x \in M^{\epsilon_\mu}$ and $x\{i\} = (-T_\omega)^i x$ for each i.

Proof. This follows from the formulae (56) and (57) in the proof of Lemma 5.8. The argument is very similar to the one given in the proof of Lemma 5.6. □

Theorem 5.10. *Let M be a $\natural$-supersingular $\mathcal{H}^{\natural}$-module. Via the isomorphism $M \cong \Delta(M)[t]$, the action of $\mathcal{H}^{\natural}$ on M can be recovered from the action of $\mathfrak{O}$ on $\Delta(M)$.*

Proof. We may assume $\dim_k(M) < \infty$. Define inductively the filtration $(F^i M)_{i\geq 0}$ of M by $\mathcal{H}^{\natural}$-submodules as follows: $F^0 M = 0$, and $F^{i+1} M$ is the preimage of $\mathcal{F}(M/F^i M)$ under the projection $M \to M/F^i M$. The $\mathcal{H}^{\natural}$-action on the graded pieces can be recovered in view of Lemma 5.7. Exhausting M step by step it is therefore enough to consider the following setting: The action of $\mathcal{H}^{\natural}$ has already been recovered on an $\mathcal{H}^{\natural}$-submodule M_0 of M and on the quotient M/M_0, and the latter is irreducible.

We reconstruct the action of $T_{e^*(\Gamma)}$ on M by means of

$$T_{e^*(\gamma)}(x) = \gamma^{-1} \cdot x \quad \text{for } \gamma \in \Gamma$$

as is tautological from our definitions. Next we are going to reconstruct the decomposition

$$M = \bigoplus_{\substack{\epsilon\in[0,q-2]^{\Phi},\\ a\in[0,q-2]}} M_{\underline{a}}^{\epsilon}. \tag{63}$$

Let $D > 0$ be such that $D+1 \in \mathbb{Z}(d+1)$ and $f^{D+1} = \mathrm{id}$ for each k-vector space automorphism f of M. (Such a D does exist. Indeed, M is finite, hence $\mathrm{Aut}_k(M)$ is finite, hence there is some $n \in \mathbb{N}$ with $f^n = \mathrm{id}$ for each $f \in \mathrm{Aut}_k(M)$. Now take $D = (d+1)n - 1$.) For $\epsilon \in [0, q-2]^{\Phi}$ and $a \in [0, q-2]$ define $M_{\underline{a}}^{[\epsilon]}$ to be the k-subspace of M generated by all $x \in M$ with $\gamma \cdot x = \gamma^a x$ (all $\gamma \in \Gamma$) and satisfying the following condition: there is some $\mu \in [0, q-1]^{\Phi}$ (depending on x) with $\epsilon_\mu = \epsilon$, and there are $u_{i,c} \in M_0^{\omega^{-1}s\omega^{i+1}\cdot\epsilon}$ for $i \geq 0$ and $0 \leq c \leq q-2$ with the following properties: Firstly, $u_{i,c} = 0$ if

(i) $\mu_i = 0$, or

(ii) $\mu_i = q-1$ and $c > 0$, or

(iii) $\mu_i < q-1$ and $c \geq q-1-\mu_i$.

Secondly, putting $x\{-1\} = x$ and

$$x\{i\} = t^{\mu_i}\varphi(x\{i-1\}) - \sum_c t^c \varphi u_{i,c}, \tag{64}$$

we have $x\{i\} \in M \cong \Delta(M)[t]$ for any i, as well as $x\{D\} = x$.

It will be enough to prove $M_{\underline{a}}^{\epsilon} = M_{\underline{a}}^{[\epsilon]}$. We first show

$$M_{\underline{a}}^{\epsilon} \subset M_{\underline{a}}^{[\epsilon]}. \tag{65}$$

We start with $\bar{x} \in \mathcal{F}(M/M_0)^{\mu} \cap (M/M_0)_{\underline{a}}^{\epsilon}$ for some μ with $\epsilon_\mu = \epsilon$. By Proposition 3.2 we may lift it to some $x \in M^{\epsilon}$ such that for each i with $T_s T_\omega^{i+1}\bar{x} = 0$ we have $T_s^2 T_\omega^{i+1} x = 0$. As T_ω maps simultaneous eigenspaces for the T_t (with $t \in \overline{T}$) again to such simultaneous eigenspaces, and as T_s^2 commutes with the T_t, we may assume $x \in M_{\underline{a}}^{\epsilon}$. Putting

$$x\{i\} = (-T_\omega)^{i+1} x$$

for $-1 \leq i \leq D$, repeated application of Lemma 2.4 shows $x\{i\} \in M_{\underline{a_{\epsilon,i}}}^{\omega^{i+1}\cdot\epsilon}$ with

$$a_{\epsilon,-1} = a, \quad a_{\epsilon,0} = a - \epsilon_0 \quad \text{and} \quad a_{\epsilon,i} = a - \epsilon_0 - \epsilon_{d-i+1} - \cdots - \epsilon_d$$

for $i \leq d$, and then $a_{\epsilon,i} = a_{\epsilon,i'}$ for $i - i' \in \mathbb{Z}(d+1)$.

If $0 \leq \mu_i < q-1$ put

$$u_{i,c} = T_\omega^{-1}((T_s(x\{i\}))_{\underline{c+\mu_i+a_{\epsilon,i}}}).$$

As $\bar{x} \in \mathcal{F}(M/M_0)^\mu$ and $\mu_i < q-1$ we have $u_{i,c} \in M_0$, and as $x\{i\} \in M^{\omega^{i+1}\cdot\epsilon}$ we have $u_{i,c} \in M^{\omega^{-1}s\omega^{i+1}\cdot\epsilon}$, together $u_{i,c} \in M_0^{\omega^{-1}s\omega^{i+1}\cdot\epsilon}$. From $\mu_i < q-1$ we furthermore deduce $k_{x\{i\}} = (\omega^{i+1}\cdot\epsilon)_1 = \mu_i$, and since $T_s^2 x\{i\} = 0$ we then see

$$t^{\mu_i}\varphi(x\{i-1\}) - x\{i\} - \sum_c t^c \varphi u_{i,c} = h(-x\{i\}) = 0. \tag{66}$$

Since furthermore $(T_s(x\{i\}))_{\underline{c+\mu_i+a_{\epsilon,i}}} = 0$ and hence $u_{i,c} = 0$ for $q-1-\mu_i \leq c \leq q-2$ by $\sharp$-supersingularity (if $0 < \mu_i < q-1$ then $\mu_i = \epsilon_{-i}$), all the conditions on the $u_{i,c}$ in the definition of $x \in M_{\underline{a}}^{[\epsilon]}$ are satisfied.

If $\mu_i = q-1$ we have $T_s^2(T_s^2 x\{i\}) = T_s^2 x\{i\}$ and hence $k_{T_s^2 x\{i\}} = q-1$ (independently of the value of μ_i we have $(\omega^{i+1}\cdot\epsilon)_1 \equiv \mu_i$ modulo $(q-1)$), hence

$$t^{q-1}\varphi T_\omega^{-1}(T_s^2 x\{i\}) + T_s^2 x\{i\} = h(T_s^2 x\{i\}) = 0. \tag{67}$$

Similarly we see $k_{(x\{i\} - T_s^2 x\{i\})} = 0$ and hence

$$\varphi T_\omega^{-1}(x\{i\} - T_s^2 x\{i\}) + x\{i\} - T_s^2 x\{i\} = h(x\{i\} - T_s^2 x\{i\}) = 0. \tag{68}$$

We compute

$$\begin{aligned} t^{q-1}\varphi(x\{i-1\}) &= -t^{q-1}\varphi T_\omega^{-1}(x\{i\}) \\ &= -t^{q-1}\varphi T_\omega^{-1} T_s^2(x\{i\}) \\ &= T_s^2(x\{i\}) \\ &= \varphi T_\omega^{-1}(x\{i\} - T_s^2 x\{i\}) + x\{i\} \end{aligned}$$

where the second equality is the result of applying t^{q-1} to formula (68), where the third equality is formula (67) and where the fourth equality is formula (68). Thus, putting $u_{i,0} = T_\omega^{-1}(x\{i\} - T_s^2 x\{i\})$ and $u_{i,c} = 0$ for $c > 0$, we again get formula (66). Moreover, $u_{i,0}$ belongs to M_0 as $\bar{x} \in \mathcal{F}(M/M_0)^\mu$ and $\mu_i = q-1$; but it also belongs to $M^{\omega^{-1}s\omega^{i+1}\cdot\epsilon}$ since $\mu_i = q-1$ implies $\omega^{-1}s\omega^{i+1}\cdot\epsilon = \omega^i\cdot\epsilon$. By construction, $x\{d\} = (-T_\omega)^{d+1}(x)$, hence $x\{D\} = (-T_\omega)^{D+1}x = x$.

It follows that $x \in M_{\underline{a}}^{[\epsilon]}$. We have shown that any element in $\mathcal{F}(M/M_0)^\mu \cap (M/M_0)_{\underline{a}}^\epsilon$, for μ with $\epsilon_\mu = \epsilon$, lifts to an element in $M_{\underline{a}}^\epsilon \cap M_{\underline{a}}^{[\epsilon]}$. Since we have $(M/M_0)^\epsilon = \sum_{\substack{\mu \in [0,q-1]^\Phi \\ \epsilon_\mu = \epsilon}} \mathcal{F}(M/M_0)^\mu$ (see Lemma 5.7) and since this is respected by the action of $T_{e^*(\Gamma)}$, we thus have reduced our problem to showing $(M_0)^{\epsilon}_{\underline{a}} \subset M^{[\epsilon]}_{\underline{a}}$. But for this we may appeal to an induction on $\dim_k(M)$ (which we may assume to be finite).

We have shown formula (65). Now we show

$$M^{[\epsilon]}_{\underline{a}} \subset M^{\epsilon}_{\underline{a}}. \tag{69}$$

Let $x \in M^{[\epsilon]}_{\underline{a}}$, $\mu \in [0, q-1]^{\Phi}$ (with $\epsilon_\mu = \epsilon$) and $u_{i,c}$ be as in the definition of $M^{[\epsilon]}_{\underline{a}}$. Define $x\{i\}$ for $-1 \leq i \leq D$ as in that definition. By Lemma 5.9 and the proof of the inclusion (65) we find $\tilde{x} \in M^{\epsilon}_{\underline{a}}$ and $\tilde{u}_{i,c} \in M_0^{\omega^{-1}s\omega^{i+1}\cdot\epsilon}$ for $0 \leq i \leq D$ such that, after replacing x by $x - \tilde{x}$ and $u_{i,c}$ by $u_{i,c} - \tilde{u}_{i,c}$, we may assume $x \in M_0$.

Claim. *If $x \in M_0$ and if M_0 is irreducible, then there is some $x' \in (M_0)_{\underline{a}}$ with $x - x' \in (M_0)^{\epsilon}_{\underline{a}}$ and such that*

$$x'\{i\} = t^{\mu_i}\varphi(\cdots(t^{\mu_1}\varphi(t^{\mu_0}\varphi x'))\cdots)$$

belongs to M_0 for all i, and $x'\{D\} = x'$.

This follows from Lemma 5.8.

If M_0 is not irreducible, choose an $\mathcal{H}$-submodule M_{00} in M_0 such that M_0/M_{00} is irreducible. By the above claim and again invoking the proof of the inclusion (65), after modifying x by another element of $M^{\epsilon}_{\underline{a}}$ (now even of $(M_0)^{\epsilon}_{\underline{a}}$) and suitably modifying the $u_{i,c}$, we may assume $u_{i,c} \in M_{00}$. Thus, it is now enough to solve the problem for the new $x \in (M_0)_{\underline{a}}$ (and the new $u_{i,c} \in M_{00}$). We continue in this way. Since we may assume that $\dim_k(M)$ is finite, an induction on the dimension of M allows us to conclude.

We have reconstructed the decomposition (63) of M.

Now we reconstruct T_sT_ω acting on M. As we already know the decomposition (63), it is enough to reconstruct $T_sT_\omega(e)$ for $e \in M^{\epsilon'}_{\underline{a}'}$, all ϵ', a'. Given such e, let $\bar{e}$ be its class in M/M_0. By Lemma 2.4 there are then ϵ, a such that $T_\omega\bar{e} \in (M/M_0)^{\epsilon}_{\underline{a}}$.

First assume $\epsilon_1 = 0$. We then reconstruct $T_sT_\omega(e)$ as $T_sT_\omega(e) = t^{q-1}\varphi(e)$. Indeed, to see this we may assume (by Lemma 2.3) that $T_\omega(e)$ is an eigenvector for T_s^2. If $T_s^2T_\omega(e) = T_\omega(e)$ and hence $T_sT_\omega(e) = -T_\omega(e)$, the claim follows from the definition of $h(T_\omega(e))$. If $T_s^2T_\omega(e) = 0$ then in fact $T_sT_\omega(e) = 0$ (since also $\epsilon_1 = 0$), and the definition of $h(T_\omega(e))$ shows $t^{q-1}\varphi(e) = 0$.

Now assume $\epsilon_1 > 0$. This implies $T_s^2T_\omega(e) = 0$ and $k_{T_\omega(e)} = \epsilon_1$, and by $\sharp$-supersingularity we get

$$t^{k_{T_\omega(e)}+1}\varphi e = -\sum_{0 \leq c < q-1-k_{T_\omega(e)}} t^{c+1}\varphi T_\omega^{-1}((T_sT_\omega e)_{\underline{c+\epsilon_1+a}}).$$

Here $(T_sT_\omega e)_{\underline{c+\epsilon_1+a}} \in M_0^{s\cdot\epsilon}$ and $q-1-k_{T_\omega(e)} = (s\cdot\epsilon)_1$. The map

$$\bigoplus_{0 \leq c < q-1-k_{T_\omega(e)}} M_0^{s\cdot\epsilon} \to M_0, \quad (y_c)_c \mapsto \sum_{0 \leq c < q-1-k_{T_\omega(e)}} t^{c+1}\varphi T_\omega^{-1}(y_c)$$

is injective. This is first seen in the case where M_0 is irreducible; it then follows by an obvious devissage argument. We therefore see that the $(T_sT_\omega e)_{\underline{c+\epsilon_1+a}}$ for $0 \leq c < q-1-k_{T_\omega(e)}$ can be read off from $t^{k_{T_\omega(e)}+1}\varphi e$, hence also $T_sT_\omega e$ can be read off from $t^{k_{T_\omega(e)}+1}\varphi e$ (by $\sharp$-supersingularity).

The restriction of T_ω to $\{x \in M \mid T_s T_\omega(x) \in M_0\}$ is reconstructed as follows. Given $\bar{x} \in (M/M_0)^{\omega^{-1}\cdot\epsilon}_{\underline{a-\epsilon_1}}$ (for some ϵ, some a) with $T_s T_\omega \bar{x} = 0$, we use the decomposition (34) to lift $\bar{x}$ to some $x \in M^{\omega^{-1}\cdot\epsilon}_{\underline{a-\epsilon_1}}$. Since $(\omega^{-1}\cdot\epsilon)_0 = \epsilon_1$, Lemma 2.4 says $T_\omega x \in M^{\epsilon}_{\underline{a}}$. It then follows from the definitions that

$$T_\omega x = -t^{\epsilon_1}\varphi x - \sum_{c\geq 0} t^c \varphi T_\omega^{-1}((T_s T_\omega x)_{\underline{c+\epsilon_1+a}}).$$

We have now collected all the data required in Proposition 3.3 for reconstructing M as an $\mathcal{H}^\sharp$-module. □

5D. ***Full faithfulness on $\sharp$-supersingular $\mathcal{H}^\sharp$-modules.*** Let $\mathrm{Rep}(\mathrm{Gal}(\overline{F}/F))$ denote the category of representations of $\mathrm{Gal}(\overline{F}/F)$ on k-vector spaces which are projective limits of finite dimensional continuous $\mathrm{Gal}(\overline{F}/F)$-representations.

Let $\mathrm{Mod}_{ss}(\mathcal{H}^\sharp)$ denote the category of $\sharp$-supersingular $\mathcal{H}^\sharp$-modules. Let $\mathrm{Mod}_{ss}(\mathcal{H})$ and $\mathrm{Mod}_{ss}(\mathcal{H}^{\sharp\sharp})$ denote the categories of supersingular $\mathcal{H}$-modules and supersingular $\mathcal{H}^{\sharp\sharp}$-modules, respectively.

Let $M \in \mathrm{Mod}_{ss}(\mathcal{H}^{\sharp\sharp})$ with $\dim_k(M) < \infty$. By Proposition 4.3 we have $\Delta(M) \in \mathrm{Mod}^{\mathrm{ad}}(\mathfrak{O})$, thus $\Delta(M)^* \otimes_{k[[t]]} k((t)) \in \mathrm{Mod}^{\mathrm{et}}(k((t)))$ (see Proposition 1.4). Let $V(M)$ be the object in $\mathrm{Rep}(\mathrm{Gal}(\overline{F}/F))$ assigned to $\Delta(M)^* \otimes_{k[[t]]} k((t))$ by Theorem 1.2. Exhausting an object in $\mathrm{Mod}_{ss}(\mathcal{H}^{\sharp\sharp})$ by its finite dimensional subobjects we see that this construction extends to all of $\mathrm{Mod}_{ss}(\mathcal{H}^{\sharp\sharp})$.

Theorem 5.11. *The assignment*

$$\mathrm{Mod}_{ss}(\mathcal{H}^{\sharp\sharp}) \to \mathrm{Rep}(\mathrm{Gal}(\overline{F}/F)), \quad M \mapsto V(M) \tag{70}$$

is an exact contravariant functor, with $\dim_k(M) = \dim_k(V(M))$ *for any* M. *Also,*

$$\begin{aligned} \mathrm{Mod}_{ss}(\mathcal{H}^{\sharp}) \to \mathrm{Rep}(\mathrm{Gal}(\overline{F}/F)), &\quad M \mapsto V(M), \\ \mathrm{Mod}_{ss}(\mathcal{H}) \to \mathrm{Rep}(\mathrm{Gal}(\overline{F}/F)), &\quad M \mapsto V(M) \end{aligned} \tag{71}$$

are exact and fully faithful contravariant functors.

Proof. Exactness follows from exactness of $M \mapsto \Delta(M)$ (Proposition 4.3), exactness of $\Delta \mapsto \Delta^* \otimes_{k[[t]]} k((t))$ (Proposition 1.4) and exactness of the equivalence functor in Theorem 1.2. From Proposition 4.3 we get $\dim_k(M) = \dim_{k((t))}(\Delta(M)^* \otimes_{k[[t]]} k((t)))$, from Theorem 1.2 we get $\dim_{k((t))}(\Delta(M)^* \otimes_{k[[t]]} k((t))) = \dim_k(V(M))$.

To prove faithfulness on $\mathrm{Mod}_{ss}(\mathcal{H}^\sharp)$, suppose that we are given finite dimensional objects M, M' in $\mathrm{Mod}_{ss}(\mathcal{H}^\sharp)$ and a morphism $\mu : V(M') \to V(M)$ in $\mathrm{Rep}(\mathrm{Gal}(\overline{F}/F))$. By Theorem 1.2, the latter corresponds to a unique morphism of étale (φ, Γ)-modules

$$\mu : \Delta(M')^* \otimes_{k[[t]]} k((t)) \to \Delta(M)^* \otimes_{k[[t]]} k((t)).$$

By Proposition 1.17 (which applies since Proposition 4.3 tells us $\Delta(M), \Delta(M') \in \mathrm{Mod}^{\clubsuit}(\mathfrak{O})$) it is induced by a unique morphism of $\mathfrak{O}$-modules $\mu : \Delta(M) \to \Delta(M')$. Clearly μ takes $\Delta(M)[t]$ to $\Delta(M')[t]$, i.e., it takes M to M'. Applying Theorem 5.10 to both M and M' we see that $\mu : M \to M'$ is $\mathcal{H}^\sharp$-equivariant. If $M, M' \in \mathrm{Mod}_{ss}(\mathcal{H}^\sharp)$ are not necessarily finite dimensional, the same conclusion is obtained by exhausting

M, M' by its finite dimensional submodules. We deduce the stated full faithfulness on $\mathrm{Mod}_{ss}(\mathcal{H}^\sharp)$. It implies full faithfulness on $\mathrm{Mod}_{ss}(\mathcal{H})$ (see Lemma 2.5). $\square$

Example. The analogs of Proposition 3.3 and Theorem 5.11 (on the functor in formula (71)) fail for supersingular $\mathcal{H}^{\sharp\sharp}$-modules. To see this, take $d=2$, and endow the 6-dimensional k-vector space M with basis $e_0, e_1, e_2, f_0, f_1, f_2$ with the structure of an $\mathcal{H}^{\sharp\sharp}$-module as follows. T_t for each $t\in\overline{T}$ acts trivially. Put $T_s(f_0)=T_s(e_1)=T_s(e_2)=0$ and $T_s(e_0)=-e_0$, $T_s(f_1)=-f_1$, $T_s(f_2)=-f_2$. Fix $\alpha\in k$ and put $T_\omega(e_0)=e_1$, $T_\omega(e_1)=e_2$, $T_\omega(e_2)=e_0$, $T_\omega(f_0)=f_1+\alpha e_1$, $T_\omega(f_1)=f_2-\alpha e_2$, $T_\omega(f_2)=f_0$. This is even an $\mathcal{H}^\sharp$-module if and only if $\alpha=0$, if and only if it is decomposable (as an $\mathcal{H}^{\sharp\sharp}$-module). The corresponding $\mathfrak{O}$-module $\Delta(M)$ is defined by the relations $\varphi e_0=-e_1$, $\varphi e_1=-e_2$, $t^{q-1}\varphi e_2=-e_0$, $\varphi f_2=-f_0$, $t^{q-1}\varphi(f_0-\alpha e_0)-f_1$, $t^{q-1}\varphi(f_1+\alpha e_1)-f_2$. But this $\mathfrak{O}$-module is in fact independent of α, since $t^{q-1}\varphi e_1=t^{q-1}\varphi e_0=0$. Thus, an $\mathcal{H}^{\sharp\sharp}$-analog of Theorem 5.11 fails. To see that an $\mathcal{H}^{\sharp\sharp}$-analog of Proposition 3.3 fails take M_0 to be the k-subvector space of M spanned by e_0, e_1, e_2; it is stable under $\mathcal{H}^{\sharp\sharp}$. The action of $\mathcal{H}^{\sharp\sharp}$ on M_0 and on M/M_0 does not depend on α. The actions of $T_\omega^{d+1}=T_\omega^3$, of $T_{e^*(\Gamma)}$ and of T_sT_ω do not depend on α. We have $(T_sT_\omega)^{-1}(M_0)=M_0+kf_2$ and hence the restriction of T_ω to $(T_sT_\omega)^{-1}(M_0)$ does not depend on α. We have $M=\sum_\epsilon M^\epsilon$ with $M^\epsilon=0$ whenever $\epsilon_1\neq 0$. Thus, an $\mathcal{H}^{\sharp\sharp}$-analog of Proposition 3.3 would predict that also the action of T_ω (even of $\mathcal{H}^{\sharp\sharp}$) is independent of α, which however is apparently not the case.

5E. ***The essential image.***

Definition. Let $\mathrm{Hom}(\Gamma,k^\times)^\Phi$ denote the group of $(d+1)$-tuples $\alpha=(\alpha_0,\dots,\alpha_d)$ of characters $\alpha_j:\Gamma\to k^\times$. Let $\mathfrak{S}_{d+1}$ act on $\mathrm{Hom}(\Gamma,k^\times)^\Phi$ by the formulae

$$(\omega\cdot\alpha)_0=\alpha_d \quad\text{and}\quad (\omega\cdot\alpha)_i=\alpha_{i-1}\quad\text{for } 1\le i\le d,$$
$$(s\cdot\alpha)_0=\alpha_1,\quad (s\cdot\alpha)_1=\alpha_0\quad\text{and}\quad (s\cdot\alpha)_i=\alpha_i\quad\text{for } 2\le i\le d.$$

Recall the action of $\mathfrak{S}_{d+1}$ on $[0,q-2]^\Phi$. Combining both (diagonally), we obtain an action of $\mathfrak{S}_{d+1}$ on $\mathrm{Hom}(\Gamma,k^\times)^\Phi\times[0,q-2]^\Phi$.

In Lemma 1.19 we attached to each standard cyclic étale (φ,Γ)-module $\boldsymbol{D}$ of dimension $d+1$ an ordered tuple $((\alpha_0,m_0),\dots,(\alpha_d,m_d))$ (with integers $m_j\in[1-q,0]$ and characters $\alpha_j:\Gamma\to k^\times$), unique up to a cyclic permutation. Sending each m_j to the representative in $[0,q-2]$ of its class in $\mathbb{Z}/(q-1)$, the tuple $(m_0,\dots,m_d)$ gives rise to an element in $[0,q-2]^\Phi$. On the other hand, the tuple $(\alpha_0,\dots,\alpha_d)$ constitutes an element in $\mathrm{Hom}(\Gamma,k^\times)^\Phi$. Taken together we thus attach to $\boldsymbol{D}$ an element in $\mathrm{Hom}(\Gamma,k^\times)^\Phi\times[0,q-2]^\Phi$, unique up to cyclic permutation. Equivalently, we attach to $\boldsymbol{D}$ an orbit in $\mathrm{Hom}(\Gamma,k^\times)^\Phi\times[0,q-2]^\Phi$ under the action of the subgroup of $\mathfrak{S}_{d+1}$ generated by ω.

Now let $\boldsymbol{D}_1'$, $\boldsymbol{D}_2'$ be irreducible étale (φ,Γ)-modules over $k((t))$. We say that $\boldsymbol{D}_1'$, $\boldsymbol{D}_2'$ are strongly $\mathfrak{S}_{d+1}$-linked if they are subquotients of $(d+1)$-dimensional standard cyclic étale (φ,Γ)-modules $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ respectively, and if $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ give rise to the same $\mathfrak{S}_{d+1}$-orbit in $\mathrm{Hom}(\Gamma,k^\times)^\Phi\times[0,q-2]^\Phi$. We say that $\boldsymbol{D}_1'$, $\boldsymbol{D}_2'$ are $\mathfrak{S}_{d+1}$-linked if they are subquotients of $(d+1)$-dimensional standard cyclic étale

(φ, Γ)-modules $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ respectively, and if $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ give rise to the same $\mathfrak{S}_{d+1}$-orbit in $[0, q-2]^{\Phi}$ (or equivalently, if the assigned tuples (up to cyclic permutation) in $[0, q-2]^{\Phi}$ coincide as *unordered* tuples (with multiplicities)).

Remark. (a) Let $\boldsymbol{D}$ denote the étale (φ, Γ)-module over $k((t))$ corresponding to $V(M)$, for a finite dimensional supersingular $\mathcal{H}^{\sharp\sharp}$-module M. Our constructions show $M = \mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^{\natural}, k)[t]$ (where $\boldsymbol{D}^{\natural}$ is given the t-adic topology). Moreover:

(i) Consider the natural map of $k[[t]][\varphi]$-modules

$$\kappa_{\boldsymbol{D}} : k[[t]][\varphi] \otimes_{k[[t]]} M \to \mathrm{Hom}_k^{\mathrm{cont}}(\boldsymbol{D}^{\natural}, k).$$

As a $k[[t]][\varphi]$-module, $\ker(\kappa_{\boldsymbol{D}})$ is generated by $\ker(\kappa_{\boldsymbol{D}}) \cap (k \otimes M + k[[t]]\varphi \otimes M)$.

(ii) Each irreducible subquotient of $\boldsymbol{D}$ is a subquotient of a $(d+1)$-dimensional standard cyclic étale (φ, Γ)-module; more precisely:

(ii)(1) If $\boldsymbol{D}$ (or equivalently, M) is indecomposable, then any two irreducible subquotients of $\boldsymbol{D}$ are $\mathfrak{S}_{d+1}$-linked.

(ii)(2) If M is even a supersingular $\mathcal{H}$-module, and if $\boldsymbol{D}$ (or equivalently, M) is indecomposable, then any two irreducible subquotients of $\boldsymbol{D}$ are strongly $\mathfrak{S}_{d+1}$-linked.

(ii)(3) If M is even a supersingular $\mathcal{H}^{\flat}$-module, then each irreducible subquotient of $\boldsymbol{D}$ is a subquotient of a $(d+1)$-dimensional standard cyclic étale (φ, Γ)-module with parameters $m_j \in \{1-q, 0\}$ and $\alpha_j = 1$ for all j.

(iii) For any (φ, Γ)-submodule $\boldsymbol{D}_0$ of $\boldsymbol{D}$ the ψ-operator on $\boldsymbol{D}_0 \cap \boldsymbol{D}^{\natural}$ is surjective.

(b) Does property (i) mean (at least if property (iii) is assumed) that $\boldsymbol{D}$ is the reduction of a crystalline p-adic $\mathrm{Gal}(\overline{F}/F)$-representation with Hodge–Tate weights in $[-1, 0]$?

(c) Property (iii) means that the functor $\boldsymbol{D}_0 \mapsto \boldsymbol{D}_0^{\natural}$ is exact on the category of subquotients $\boldsymbol{D}_0$ of $\boldsymbol{D}$.

(d) It should not be too hard to show that properties (i), (ii)(1) and (iii) together in fact *characterize* the essential image of the functor (70).

(e) On the other hand, properties (i), (ii)(2) and (iii) together do *not* characterize the essential image of the functor (71). To see this for $d = 1$ consider the following étale (φ, Γ)-module $\boldsymbol{D}$ (which satisfies (i), (ii)(2), (iii)). It is given by a k-basis $e_0, e_1, f_0, f_1, g_0, g_1$ of $(\boldsymbol{D}^{\natural})^*[t]$ and the following relations:

$$\varphi e_1 = e_0, \quad \varphi f_1 = f_0, \quad \varphi g_1 = g_0, \quad t^{q-1}\varphi e_0 = e_1, \quad t^{q-1}\varphi f_0 = f_1 + e_1, \quad t^{q-1}\varphi g_0 = g_1 + f_0.$$

Another object not in the essential image is defined by the set of relations

$$\varphi e_1 = e_0, \quad \varphi f_1 = f_0, \quad \varphi g_1 = g_0, \quad t^{q-1}\varphi e_0 = e_1, \quad t^{q-1}\varphi f_0 = f_1 + e_0, \quad t^{q-1}\varphi g_0 = g_1 + f_1.$$

6. From G-representations to $\mathcal{H}$-modules

6A. ***Supersingular cohomology.*** Put $G = \mathrm{GL}_{d+1}(F)$, let I_0 be a pro-p-Iwahori subgroup in G, and fix an isomorphism between $\mathcal{H}$ and the pro-p-Iwahori Hecke algebra $k[I_0\backslash G/I_0]$ corresponding to $I_0 \subset G$. For a smooth G-representation Y (over k) the subspace Y^{I_0} of I_0-invariants then receives a natural action by $\mathcal{H}$. Let us denote by $H^0_{ss}(I_0, Y)$ the maximal supersingular $\mathcal{H}$-submodule of Y^{I_0}. It is clear that this defines a left exact functor

$$\mathrm{Mod}(G) \to \mathrm{Mod}_{ss}(\mathcal{H}), \quad Y \mapsto H^0_{ss}(I_0, Y)$$

where $\mathrm{Mod}(G)$ denotes the category of smooth G-representations. The category $\mathrm{Mod}(G)$ is a Grothendieck category [Schneider 2015, Lemma 1] and has enough injective objects [Vignéras 1996, I.5.9]. Let $D^+(G)$ denote the derived category of complexes of smooth G-representations vanishing in negative degrees, let $D^+_{ss}(\mathcal{H})$ denote the derived category of complexes of supersingular $\mathcal{H}$-modules vanishing in negative degrees. The above functor gives rise to a right derived functor

$$R_{ss}(I_0, .) : D^+(G) \to D^+_{ss}(\mathcal{H}). \tag{72}$$

Let $D^+(\mathrm{Gal}(\overline{F}/F)))$ denote the derived category of complexes in $\mathrm{Rep}(\mathrm{Gal}(\overline{F}/F))$ vanishing in negative degrees. Since the functor V is exact, it induces a functor

$$V : D^+_{ss}(\mathcal{H}) \to D^+(\mathrm{Gal}(\overline{F}/F))).$$

We may compose them with $R_{ss}(I_0, .)$ to obtain a functor

$$V \circ R_{ss}(I_0, .) : D^+(G) \to D^+(\mathrm{Gal}(\overline{F}/F))).$$

Remark. The functor $H^0_{ss}(I_0, .)$ is the composite of the left exact functor $\mathrm{Mod}(G) \to \mathrm{Mod}(\mathcal{H})$, $Y \mapsto Y^{I_0}$ (taking I_0-invariants) and the left exact functor $\mathrm{Mod}(\mathcal{H}) \to \mathrm{Mod}_{ss}(\mathcal{H})$, $M \mapsto M_{ss}$ which takes an $\mathcal{H}$-module to its maximal supersingular $\mathcal{H}$-submodule. Also $\mathrm{Mod}(\mathcal{H})$ is a Grothendieck category with enough injective objects. Writing $R(I_0, .)$ and $R_{ss}(.)$ for the respective right derived functors, we have a morphism $R_{ss}(I_0, .) \to R_{ss}(.) \circ R(I_0, .)$.

Remark. Of course, we expect the functor $V \circ R_{ss}(I_0, .)$ to be meaningful only when restricted to (complexes of) supersingular G-representations. The reason is the following theorem of Ollivier and Vignéras [2018]: A smooth admissible irreducible G-representation Y over an algebraic closure $\bar{k}$ of k is supersingular if and only if Y^{I_0} is a supersingular $\mathcal{H} \otimes_k \bar{k}$-module, if and only if Y^{I_0} admits a supersingular subquotient.

It is known that, beyond the case where $G = \mathrm{GL}_2(\mathbb{Q}_p)$, a smooth admissible irreducible supersingular G-representation Y over k is not uniquely determined by the $\mathcal{H}$-module Y^{I_0}. Is it perhaps uniquely determined by the derived object $R_{ss}(I_0, Y) \in D^+_{ss}(\mathcal{H})$? It would then also be uniquely determined by the derived object $V(R_{ss}(I_0, Y)) \in D^+(\mathrm{Gal}(\overline{F}/F)))$.

Remark. For the universal module $Y=\mathrm{ind}_{I_0}^G k$ we have $H^0_{ss}(I_0,Y)=0$ since $\mathcal{H}=(\mathrm{ind}_{I_0}^G k)^{I_0}$ does not contain nonzero finite dimensional $\mathcal{H}$-submodules (let alone supersingular ones).

6B. ***An exact functor from G-representations to $\mathcal{H}$-modules.*** We fix a $(d+1)$-st root of unity $\xi\in k^\times$ with $\sum_{j=0}^d \xi^j=0$.

For an $\mathcal{H}$-module M and $j\in\mathbb{Z}$ let M^{ξ^j} be the $\mathcal{H}$-module which coincides with M as a module over the k-subalgebra $k[T_s,T_t]_{t\in\overline{T}}$, but with $T_\omega|_{M^{\xi^j}}=\xi^j T_\omega|_M$.

Let $\delta: M_0\to M_1$ be a morphism of $\mathcal{H}$-modules. For $(x_0,x_1)\in M_0\oplus M_1$ put

$$\begin{aligned}T_\omega((x_0,x_1))&=(T_\omega(x_0),\,T_\omega(\delta(x_0))+\xi T_\omega(x_1)),\\ T_s((x_0,x_1))&=(T_s(x_0),\,T_s(x_1)),\\ T_t((x_0,x_1))&=(T_t(x_0),\,T_t(x_1))\quad\text{for } t\in\overline{T}.\end{aligned}$$

Lemma 6.1. *These formulae define an $\mathcal{H}$-module structure on $M_0\oplus M_1$; we denote this new $\mathcal{H}$-module by $M_0\oplus^\delta M_1$. We have an exact sequence of $\mathcal{H}$-modules*

$$0\to M_1^\xi\to M_0\oplus^\delta M_1\to M_0\to 0. \tag{73}$$

The morphism $\delta: M_0\to M_1$ can be recovered from the exact sequence (73).

If there is some $\lambda\in k^\times$ with $T_\omega^{d+1}=\lambda$ on M_0 and on M_1, then also $T_\omega^{d+1}=\lambda$ on $M_0\oplus^\delta M_1$,

Proof. By induction on i one shows

$$T_\omega^i((x_0,x_1))=\Big(T_\omega^i(x_0),\,\xi^i T_\omega^i(x_1)+\sum_{j=0}^{i-1}\xi^j T_\omega^i(\delta(x_0))\Big)$$

for $i>0$, and hence $T_\omega^{d+1}((x_0,x_1))=(T_\omega^{d+1}(x_0),T_\omega^{d+1}(x_1))$. From here, all the required relations are straightforwardly verified, showing that indeed we have defined an $\mathcal{H}$-module.

Obviously, from the exact sequence (73) both M_0 and M_1 can be recovered. That also δ can be recovered follows from the following more general consideration. Suppose that we are given $\delta: M_0\to M_1$ and $\epsilon: N_0\to N_1$ and a morphism of $\mathcal{H}$-modules $f: M_0\oplus^\delta M_1\to N_0\oplus^\epsilon N_1$ with $f(M_1^\xi)\subset N_1^\xi$. Then there are $\mathcal{H}$-module homomorphisms $f_0: M_0\to N_0$, $f_1: M_1^\xi\to N_1^\xi$ and $\tilde f: M_0\to N_1^\xi$ with $f((x_0,x_1))=(f_0(x_0),f_1(x_1)+\tilde f(x_0))$. For $x_0\in M_0$ we compute

$$\begin{aligned}f(T_\omega(x_0,0))&=f(T_\omega(x_0),T_\omega(\delta(x_0)))=(T_\omega(f_0(x_0)),\,T_\omega(f_1(\delta(x_0)))+\xi T_\omega(\tilde f(x_0))),\\ T_\omega(f(x_0,0))&=T_\omega(f_0(x_0),\tilde f(x_0))=(T_\omega(f_0(x_0)),\,T_\omega(\epsilon(f_0(x_0)))+\xi T_\omega(\tilde f(x_0))).\end{aligned}$$

As $f(T_\omega(x_0,0))=T_\omega(f(x_0,0))$ we deduce $T_\omega(\epsilon(f_0(x_0)))=T_\omega(f_1(\delta(x_0)))$, and since T_ω is an isomorphism even $\epsilon(f_0(x_0))=f_1(\delta(x_0))$. □

Let

$$(M_\bullet,\delta_\bullet)=[\cdots\xrightarrow{\delta_{-2}}M_{-1}\xrightarrow{\delta_{-1}}M_0\xrightarrow{\delta_0}M_1\xrightarrow{\delta_1}M_2\xrightarrow{\delta_2}\cdots]$$

be a complex of $\mathcal{H}$-modules.

Lemma 6.2. (a) *There is a unique $\mathcal{H}$-module $\oplus^{\delta_\bullet}_{j\in\mathbb{Z}}M_j$ with the following properties*:

- *As a k-vector space, $\oplus^{\delta_\bullet}_{j\in\mathbb{Z}}M_j=\oplus_{j\in\mathbb{Z}}M_j$.*
- *For any j we have $\tau(M_j)\subset M_j+M_{j+1}$ for each $\tau\in\mathcal{H}$; in particular, the subspace $M_{\ge j}=\oplus_{j'\ge j}M_{j'}$ is an $\mathcal{H}$-submodule.*
- *The $\mathcal{H}$-module $M_{\ge j}/M_{\ge j+2}$ is isomorphic with $M_j^{\xi^j}\oplus^{\delta_j}M_{j+1}^{\xi^j}$ as defined in Lemma 6.1.*

(b) *If there is some $\lambda\in k^\times$ with $T_\omega^{d+1}=\lambda$ on each M_j, then $T_\omega^{d+1}=\lambda$ on $\oplus^{\delta_\bullet}_{j\in\mathbb{Z}}M_j$.*

(c) *The assignment $(M_\bullet,\delta_\bullet)\mapsto(\oplus^{\delta_\bullet}_{j\in\mathbb{Z}}M_j,(M_{\ge j})_{j\in\mathbb{Z}})$ is an exact and faithful functor from the category of complexes of $\mathcal{H}$-modules to the category of filtered $\mathcal{H}$-modules. The isomorphism class of the complex $(M_\bullet,\delta_\bullet)$ can be recovered from the isomorphism class of the filtered $\mathcal{H}$-module $(\oplus^{\delta_\bullet}_{j\in\mathbb{Z}}M_j,(M_{\ge j})_{j\in\mathbb{Z}})$.*

Proof. This is clear from Lemma 6.1. □

Definition. (a) For a smooth G-representation Y over k and $i\ge 0$ let us denote by $H^i_{ss}(I_0,Y)$ the i-th cohomology group of $R_{ss}(I_0,Y)$, see formula (72).

(b) We say that a smooth G-representation Y over k is *exact* if for each $i\ge 0$ the functor $Y'\mapsto H^i_{ss}(I_0,Y')$ is exact on the category of G-subquotients Y' of Y.

(c) An exhaustive and separated decreasing filtration $(Y^j)_{j\in\mathbb{Z}}$ of a smooth G-representation Y over k is *exact* if Y^j/Y^{j+1} is exact for each j.

Example. A semisimple smooth G-representation is exact.

Let $\mathfrak{R}_G$ denote the following category: objects are smooth G-representations with an exact filtration, morphisms are G-equivariant maps respecting the filtrations (i.e., $f:Y\to W$ with $f(Y^i)\subset W^i$ for all i). We denote objects $(Y,(Y^i)_{i\in\mathbb{Z}})$ in $\mathfrak{R}_G$ simply by $Y^\bullet$.

Let $\mathfrak{E}(\mathcal{H})$ denote the category of E_1-spectral sequences in the category of $\mathcal{H}$-modules.

For $Y^\bullet\in\mathfrak{R}_G$ we have the spectral sequence

$$E(Y^\bullet)=[E_1^{m,n}(Y^\bullet)=H^{m+n}_{ss}(I_0,Y^m/Y^{m+1})\Rightarrow H^{m+n}_{ss}(I_0,Y)].$$

A morphism $f:Y^\bullet\to W^\bullet$ in $\mathfrak{R}_G$ induces morphisms $H^m_{ss}(I_0,Y^i/Y^{i+1})\to H^m_{ss}(I_0,W^i/W^{i+1})$ for any m and i, and these induce a morphism of spectral sequences $E(Y^\bullet)\to E(W^\bullet)$. We thus obtain a functor

$$\mathfrak{R}_G\to\mathfrak{E}(\mathcal{H}),\qquad Y^\bullet\mapsto E(Y^\bullet).$$

For $r\ge 1$ let $\mathcal{Y}_r$ be the set of equivalence classes of pairs of integers (m,n), where (m,n) is declared to be equivalent with (m',n') if and only if there is some $j\in\mathbb{Z}$ with $(m,n)=(m'+jr,n'-j(r-1))$. For $y\in\mathcal{Y}_r$ let $E_r^y(Y^\bullet)$ be the complex of $\mathcal{H}$-modules whose terms are the $E_r^{m,n}(Y^\bullet)$ with $(m,n)\in y$, and whose differentials $d_r:E_r^{m,n}(Y^\bullet)\to E_r^{m+r,n-r+1}(Y^\bullet)$ are given by the spectral sequence. We apply the functor of Lemma 6.2 to $E_r^y(Y^\bullet)$ to obtain a (filtered) supersingular $\mathcal{H}$-module $\boldsymbol{E}_r^y(Y^\bullet)$.

For a morphism $f:Y^\bullet\to W^\bullet$ in $\mathfrak{R}_G$ we have induced $\mathcal{H}$-linear maps $f_r:\oplus_{y\in\mathcal{Y}_r}\boldsymbol{E}_r^y(Y^\bullet)\to\oplus_{y\in\mathcal{Y}_r}\boldsymbol{E}_r^y(W^\bullet)$. Notice however that, in general, for a given $y\in\mathcal{Y}_r$ there is no $y'\in\mathcal{Y}_r$ such that $f_r(\boldsymbol{E}_r^y(Y^\bullet))\subset\boldsymbol{E}_r^{y'}(W^\bullet)$, even if $r=1$.

Lemma 6.3. *Let $Y^\bullet \to W^\bullet \to X^\bullet$ be a complex in $\mathfrak{R}_G$ such that for each i the induced sequence $0 \to Y^i/Y^{i+1} \to W^i/W^{i+1} \to X^i/X^{i+1} \to 0$ is exact. We then have an exact sequence of supersingular $\mathcal{H}$-modules*

$$0 \to \bigoplus_{y\in\mathcal{Y}_1} \boldsymbol{E}_1^y(Y^\bullet) \to \bigoplus_{y\in\mathcal{Y}_1} \boldsymbol{E}_1^y(W^\bullet) \to \bigoplus_{y\in\mathcal{Y}_1} \boldsymbol{E}_1^y(X^\bullet) \to 0.$$

Proof. This follow from the constructions. □

Remark. The analog of Lemma 6.3 is false for the maps f_r for $r > 1$.

Remark. For a smooth G-representation Y endowed with an exact filtration, we may apply the functor V of Section 5D to the supersingular $\mathcal{H}$-module $\boldsymbol{E}_r^y(Y^\bullet)$ (any r). In this way, we assign a $\mathrm{Gal}(\overline{F}/F)$-representation to Y. We propose this construction as a nonderived alternative to that of Section 6A. Of course, again it will be meaningful only on supersingular G-representations.

We expect that for $G = \mathrm{GL}_2(\mathbb{Q}_p)$, this construction, with $r = 1$, essentially recovers the restriction of Colmez's functor to all supersingular G-representations.[11]

Acknowledgements

I thank Laurent Berger, Peter Schneider and Gergely Zábrádi for helpful discussions related to this work. I thank Marie-France Vignéras for a very close reading of the text and for detailed suggestions for improvement. I thank the anonymous referees for their careful reading and helpful recommendations. I thank Rachel Ollivier for the invitation to UBC Vancouver in the spring of 2017; some progress on this work was obtained during that visit.

References

[Colmez 2010] P. Colmez, "(φ, Γ)-modules et représentations du mirabolique de $\mathrm{GL}_2(\mathbb{Q}_p)$", pp. 61–153 in *Représentations p-adiques de groupes p-adiques, II: Représentations de* $\mathrm{GL}_2(\mathbb{Q}_p)$ *et* (φ, Γ)*-modules*, edited by L. Berger et al., Astérisque **330**, Soc. Math. France, Paris, 2010. MR Zbl

[Emerton 2008] M. Emerton, "On a class of coherent rings, with applications to the smooth representation theory of $\mathrm{GL}_2(\mathbb{Q}_p)$ in characteristic p", preprint, 2008, Available at https://tinyurl.com/emerpdf.

[Fontaine 1990] J.-M. Fontaine, "Représentations p-adiques des corps locaux, I", pp. 249–309 in *The Grothendieck Festschrift, II*, edited by P. Cartier et al., Progr. Math. **87**, Birkhäuser, Boston, 1990. MR Zbl

[Grosse-Klönne 2016] E. Grosse-Klönne, "From pro-p Iwahori–Hecke modules to (φ, Γ)-modules, I", *Duke Math. J.* **165**:8 (2016), 1529–1595. MR Zbl

[Grosse-Klönne 2019] E. Grosse-Klönne, "A note on multivariable (φ, Γ)-modules", *Res. Number Theory* **5**:1 (2019), art. id. 6. MR Zbl

[Kisin and Ren 2009] M. Kisin and W. Ren, "Galois representations and Lubin–Tate groups", *Doc. Math.* **14** (2009), 441–461. MR Zbl

[Ollivier 2010] R. Ollivier, "Parabolic induction and Hecke modules in characteristic p for p-adic GL_n", *Algebra Number Theory* **4**:6 (2010), 701–742. MR Zbl

[Ollivier and Vignéras 2018] R. Ollivier and M.-F. Vignéras, "Parabolic induction in characteristic p", *Selecta Math.* (*N.S.*) **24**:5 (2018), 3973–4039. MR Zbl

[11] i.e., not only to those generated by their I_0-invariants

[Schneider 2015] P. Schneider, "Smooth representations and Hecke modules in characteristic p", *Pacific J. Math.* **279**:1-2 (2015), 447–464. MR Zbl

[Schneider 2017] P. Schneider, *Galois representations and* (φ, Γ)*-modules*, Cambridge Stud. Adv. Math. **164**, Cambridge Univ. Press, 2017. MR Zbl

[Schneider and Venjakob 2016] P. Schneider and O. Venjakob, "Coates–Wiles homomorphisms and Iwasawa cohomology for Lubin–Tate extensions", pp. 401–468 in *Elliptic curves, modular forms and Iwasawa theory* (Cambridge, 2015), edited by D. Loeffler and S. L. Zerbes, Springer Proc. Math. Stat. **188**, Springer, 2016. MR Zbl

[Vignéras 1996] M.-F. Vignéras, *Représentations* l*-modulaires d'un groupe réductif* p*-adique avec* $l \neq p$, Progr. Math. **137**, Birkhäuser, Boston, 1996. MR Zbl

[Vignéras 2005] M.-F. Vignéras, "Pro-p-Iwahori–Hecke ring and supersingular $\overline{\mathbb{F}}_p$-representations", *Math. Ann.* **331**:3 (2005), 523–556. Correction in **333**:3 (2005), 699–701. MR Zbl

[Vignéras 2017] M.-F. Vigneras, "The pro-p-Iwahori–Hecke algebra of a reductive p-adic group, III: Spherical Hecke algebras and supersingular modules", *J. Inst. Math. Jussieu* **16**:3 (2017), 571–608. MR Zbl

Communicated by Marie-France Vignéras
Received 2018-12-10 Revised 2019-05-06 Accepted 2019-09-01

gkloenne@math.hu-berlin.de *Mathematisch-Naturwissenschaftliche Fakultät, Institut für Mathematik, Humboldt-Universität zu Berlin, Germany*

ALGEBRA AND NUMBER THEORY 14:1 (2020)
dx.doi.org/10.2140/ant.2020.14.119

Stability in the homology of unipotent groups

Andrew Putman, Steven V Sam and Andrew Snowden

Let R be a (not necessarily commutative) ring whose additive group is finitely generated and let $U_n(R) \subset \mathrm{GL}_n(R)$ be the group of upper-triangular unipotent matrices over R. We study how the homology groups of $U_n(R)$ vary with n from the point of view of representation stability. Our main theorem asserts that if for each n we have representations M_n of $U_n(R)$ over a ring $\boldsymbol{k}$ that are appropriately compatible and satisfy suitable finiteness hypotheses, then the rule $[n] \mapsto \mathrm{H}_i(U_n(R), M_n)$ defines a finitely generated OI-module. As a consequence, if $\boldsymbol{k}$ is a field then $\dim \mathrm{H_i(U_n(R)}, \boldsymbol{k})$ is eventually equal to a polynomial in n. We also prove similar results for the Iwahori subgroups of $\mathrm{GL}_n(\mathcal{O})$ for number rings $\mathcal{O}$.

1. Introduction

1A. ***Homology of unipotent groups.*** Groups of the form $G(R)$, with G a linear algebraic group over a ring R, are among the most common and important groups encountered in mathematics. It is therefore a natural problem to understand their group homology, as homology is one of the most important invariants of a group. In the case where G is reductive, this problem has been studied intensively and much is known. See, for instance, [Borel 1974] for G a classical group and R a number ring, and [Quillen 1972] for $G = \mathrm{GL}_n$ and R a finite field. These computations are closely connected to algebraic K-theory.

On the other hand, when G is a unipotent group, comparatively little is known. In fact, the class of unipotent groups is fairly wild, so there might not be too much one can say in complete generality. Let $U_n \subset \mathrm{GL}_n$ be the group of upper-unitriangular matrices. These are perhaps the most important unipotent groups; for example, Engel's theorem [Borel 1969, Corollary I.4.8] shows that any unipotent group

Andrew Putman was partially supported by NSF DMS-1737434. Steven V Sam was partially supported by NSF DMS-1500069, DMS-1651327, and a Sloan research fellowship. Andrew Snowden was partially supported by NSF DMS-1303082, DMS-1453893, and a Sloan research fellowship.
MSC2010: primary 20J05; secondary 16P40.
Keywords: representation stability, unipotent groups, OI-modules, OVI-modules.

embeds into one of them. Nonetheless, the homology of even these groups is poorly understood. The purpose of this paper is to establish some new results in this direction.

To illustrate the difficulties in computing the homology of $U_n(R)$, let us consider the first few cases. We take $R = \mathbb{F}_p$ for simplicity. The group $U_1(\mathbb{F}_p)$ is trivial. The group $U_2(\mathbb{F}_p)$ is simply isomorphic to the additive group of $\mathbb{F}_p$, i.e., $\mathbb{Z}/p\mathbb{Z}$, and the homology of this group is known (it is $\mathbb{Z}$ in degree 0, $\mathbb{Z}/p\mathbb{Z}$ in odd degrees, and 0 in positive even degrees). The group $U_3(\mathbb{F}_p)$ is a nonabelian group of order p^3. It fits into an exact sequence

$$1 \to \mathbb{Z}/p\mathbb{Z} \to U_3(\mathbb{F}_p) \to (\mathbb{Z}/p\mathbb{Z})^2 \to 1,$$

where the left $\mathbb{Z}/p\mathbb{Z}$ is the center of $U_3(\mathbb{F}_p)$. We therefore have a spectral sequence (the Leray–Serre spectral sequence) that computes the homology of $U_3(\mathbb{F}_p)$ in terms of the homology of the outer groups:

$$\mathrm{E}^2_{p,q} = \mathrm{H}_p((\mathbb{Z}/p\mathbb{Z})^2, \mathrm{H}_q(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z})) \Longrightarrow \mathrm{H}_{p+q}(U_3(\mathbb{F}_p), \mathbb{Z}).$$

The action of $(\mathbb{Z}/p\mathbb{Z})^2$ on $\mathrm{H}_q(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z})$ is trivial, and so the groups on the E^2 page are easy to compute. However, it is less clear what the differentials are on the E^2 page, much less on subsequent pages, and so it is not obvious how to actually compute the homology of $U_3(\mathbb{F}_p)$ explicitly from this spectral sequence.

The analysis of $U_3(\mathbb{F}_p)$ we have just made, discouraging though it may be, does highlight a general theoretical approach to studying the homology of $U_n(\mathbb{F}_p)$: this group is nilpotent, so one can break it up into abelian groups and then use the resulting spectral sequences to study its homology. Of course, this approach becomes increasingly complicated as n grows, and there is probably little chance of understanding the spectral sequences in an explicit way in general.

The main point of this paper is that, although these spectral sequences become increasingly complicated, they exhibit a kind of regularity as n varies. The precise formulation of this statement uses the language of representation stability, and requires some preliminaries, so for the moment we simply give a sample application to the main objects of interest:

Theorem 1.1. *Let R be a (not necessarily commutative) ring whose additive group is finitely generated and let $\boldsymbol{k}$ be a field. For all $i \geq 0$, there exists some $f_i(t) \in \mathbb{Q}[t]$ such that $\dim \mathrm{H}_i(U_n(R), \boldsymbol{k}) = f_i(n)$ for $n \gg 0$.*

For the ring R in the theorem, one could take a finite field, or a number ring, or the ring of 2×2 matrices over one of these rings, for example.

Example 1.2. The case $R = \mathbb{Z}$ and $\boldsymbol{k} = \mathbb{Q}$ of Theorem 1.1 follows from work of Dwyer [1985, Theorem 1.1]. He shows that the dimension of $\mathrm{H}_i(U_n(\mathbb{Z}), \mathbb{Q})$ is the number of permutations in S_n with length i, where the length of a permutation σ is the number of pairs $i < j$ such that $\sigma(i) > \sigma(j)$. Denote this number by $I(i, n)$. We claim that $n \mapsto I(i, n)$ is a polynomial of degree i for $n \gg 0$. As an aside, this shows that the degree of the polynomials $f_i(t)$ in Theorem 1.1 cannot be bounded as we let i vary. We prove the claim by induction. For $i = 1$, we have $I(1, n) = n - 1$ for $n > 0$. In general, we have the identity

$$\sum_{i \geq 0} I(i, n) q^i = (1+q)(1+q+q^2) \cdots (1+q+q^2+\cdots+q^{n-1});$$

see [Stanley 2012, Corollary 1.3.13]. It follows that $I(i,n) - I(i,n-1) = \sum_{j=0}^{i-1} I(j,n-1)$ for $n > i$. By induction, the right hand side is a polynomial of degree $i-1$ for $n \gg 0$. Hence $I(i,n)$ is a polynomial of degree i for $n \gg 0$, as claimed.

1B. *Main results.* Our main result is a refined version of Theorem 1.1 where we allow systems of nontrivial coefficients and give a stronger conclusion. This additional generality is interesting in its own right, but is required even if one is ultimately only interested in the case of trivial coefficients. Indeed, our general approach essentially relates the i-th homology group of some system of coefficients to lower homology groups of some auxiliary systems, and the auxiliary systems can be nontrivial even if the initial system is trivial.

To formulate this general theorem, we must make sense of a "system" of representations of $U_n(R)$. For this, we introduce the category $\mathrm{OVI}(R)$. An object of $\mathrm{OVI}(R)$ is a finite rank free R-module equipped with a totally ordered basis. A morphism of $\mathrm{OVI}(R)$ is a map of R-modules that is upper-triangular with respect to the distinguished ordered bases (see Section 4A). An $\mathrm{OVI}(R)$-*module* over a commutative ring $\boldsymbol{k}$ is a functor $\mathrm{OVI}(R) \to \mathrm{Mod}_{\boldsymbol{k}}$. Every object in $\mathrm{OVI}(R)$ is isomorphic to R^n equipped with its standard basis for some n, and the automorphism group of this object is the group $U_n(R)$. Thus an $\mathrm{OVI}(R)$-module M gives rise to a sequence $\{M_n\}_{n\geq 0}$, where $M_n = M(R^n)$ is a representation of $U_n(R)$, and therefore provides a reasonable notion of a system of $U_n(R)$ representations. We are primarily interested in *finitely generated* $\mathrm{OVI}(R)$-modules (see Section 2A for the definition): indeed, it is only reasonable to expect uniform behavior of the homology in this case.

Example 1.3. (a) We have a constant $\mathrm{OVI}(R)$-module given by $R^n \mapsto \boldsymbol{k}$ for all n. Thus the sequence of trivial representations of $U_n(R)$ forms a "system" in our sense.

(b) Suppose $R = \boldsymbol{k}$. We then have an $\mathrm{OVI}(R)$-module given by $R^n \mapsto R^n$. We thus see that, in this case, the sequence of standard representations of $U_n(R)$ forms a "system." Both examples are finitely generated.

Let M be an $\mathrm{OVI}(R)$-module and fix $i \geq 0$. For each n we consider the homology group $\mathrm{H}_i(U_n(R), M_n)$. The various M_n are related by the $\mathrm{OVI}(R)$-module structure, and this should lead to relationships between these homology groups. We now examine this. Letting $[n]$ denote the ordered set $\{1, \ldots, n\}$, if $[n] \to [m]$ is an order-preserving injection of finite sets then there is an associated morphism $R^n \to R^m$ in $\mathrm{OVI}(R)$. This gives a map $M_n \to M_m$, which induces a map $\mathrm{H}_i(U_n(R), M_n) \to \mathrm{H}_i(U_m(R), M_m)$. This suggests that $[n] \mapsto \mathrm{H}_i(U_n(R), M_n)$ defines an OI-module, where OI is the category whose objects are finite totally ordered sets and whose morphisms are order-preserving injections. We show that this is indeed the case, and denote this OI-module by $\boldsymbol{H}_i(\boldsymbol{U}, M)$. (We note that OI-modules are close relatives of the well-known FI-modules introduced by Church, Ellenberg, and Farb [Church et al. 2015].)

We can now state our main theorem.

Theorem 1.4. *Let R be a ring whose additive group is finitely generated, let $\boldsymbol{k}$ be a noetherian commutative ring, and M be a finitely generated* $\mathrm{OVI}(R)$*-module over $\boldsymbol{k}$. Then $\boldsymbol{H}_i(\boldsymbol{U}, M)$ is a finitely generated* OI*-module over $\boldsymbol{k}$ for all $i \geq 0$.*

Theorem 1.1 follows immediately from this theorem by taking M_n to be the trivial representation of $U_n(R)$ for all n and appealing to the fact that a finitely generated OI-module over a field has eventually polynomial dimension (see Proposition 3.5 below).

1C. ***The noetherian result.*** As stated, to prove Theorem 1.4 we relate the homology of the $\mathrm{OVI}(R)$-module M to the homology of certain auxiliary coefficient systems constructed by various means. To ensure that these auxiliary systems are finitely generated, we require the following noetherian result, which is the primary technical result of this paper:

Theorem 1.5. *Let R be a ring whose additive group is finitely generated and let $\boldsymbol{k}$ be a noetherian commutative ring. Then the category of* $\mathrm{OVI}(R)$*-modules over $\boldsymbol{k}$ is locally noetherian, that is, any submodule of a finitely generated module is finitely generated.*

Theorem 1.5 differs from much previous work on categories of R-modules in the setting of representation stability (such as [Putman and Sam 2017; Sam and Snowden 2017]) in that it allows the ring R to be infinite. In the previous work, the automorphism groups in the categories under consideration were $\mathrm{GL}_n(R)$, and finiteness of R is necessary since the group algebra of $\mathrm{GL}_n(R)$ is not noetherian if R is infinite. In our situation, the automorphism groups are $U_n(R)$. When the additive group of R is finitely generated, these groups are virtually polycyclic, and a classical result of Philip Hall [1954] says that group rings of virtually polycyclic groups are noetherian. Our proof of Theorem 1.5 is inspired in part by Hall's proof of this fact.

Remark 1.6. It is easy to see that Theorem 1.5 is false if the additive group of R is not finitely generated (see Section 5D).

Remark 1.7. When the ring R is finite, we in fact show that the category of $\mathrm{OVI}(R)$-modules is quasi-Gröbner in the sense of [Sam and Snowden 2017, Section 4], which implies local noetherianity (but is stronger). In the general case, we do not show that the category of $\mathrm{OVI}(R)$-modules is quasi-Gröbner (and expect that it is not), and the proof of local noetherianity is far more difficult.

1D. ***Application to Iwahori groups.*** Let $\mathcal{O}$ be a number ring and let $\boldsymbol{k}$ be a commutative noetherian ring. A classical result of van der Kallen [1980] says that the homology of the group $\mathrm{GL}_n(\mathcal{O})$ stabilizes: for any fixed i the canonical map

$$\mathrm{H}_i(\mathrm{GL}_n(\mathcal{O}), \boldsymbol{k}) \to \mathrm{H}_i(\mathrm{GL}_{n+1}(\mathcal{O}), \boldsymbol{k})$$

is an isomorphism for $n \gg 0$. In particular, if $\boldsymbol{k}$ is a field then the dimension of $\mathrm{H}_i(\mathrm{GL}_n(\mathcal{O}), \boldsymbol{k})$ is eventually constant.

Now let $\mathfrak{a}$ be a nonzero proper ideal in $\mathcal{O}$ and let $\mathrm{GL}_n(\mathcal{O}, \mathfrak{a})$ be the principal congruence subgroup of level $\mathfrak{a}$, i.e., the subgroup of $\mathrm{GL}_n(\mathcal{O})$ consisting of matrices that are congruent to the identity modulo $\mathfrak{a}$. The homology of these groups does not stabilize; for instance, for $\ell \geq 2$ and $n \geq 3$ the abelianization of $\mathrm{GL}_n(\mathbb{Z}, \ell\mathbb{Z})$ is $(\mathbb{Z}/\ell)^{n^2-1}$ (see [Lee and Szczarba 1976]). Building on work of the first author [Putman 2015], Church, Ellenberg, Farb and Nagpal [Church et al. 2014] proved instead that the homology of

$\mathrm{GL}_n(\mathcal{O}, \mathfrak{a})$ satisfies a version of representation stability: the rule $[n] \mapsto \mathrm{H}_i(\mathrm{GL}_n(\mathcal{O}, \mathfrak{a}), \boldsymbol{k})$ defines a finitely generated FI-module. Consequently, when $\boldsymbol{k}$ is a field, the dimension is eventually polynomial.

The Iwahori subgroup $\mathrm{GL}_{n,0}(\mathcal{O}, \mathfrak{a})$ is the subgroup of $\mathrm{GL}_n(\mathcal{O})$ consisting of matrices that are upper-triangular modulo $\mathfrak{a}$. Using Theorem 1.4, we prove an analog of Church, Ellenberg, Farb, and Nagpal's result for $\mathrm{GL}_{n,0}(\mathcal{O}, \mathfrak{a})$.

Theorem 1.8. *Let $\mathcal{O}$ be a number ring, let $\mathfrak{a} \subset \mathcal{O}$ be a nonzero proper ideal, and let $\boldsymbol{k}$ be a commutative noetherian ring. Then the following hold for all $i \geq 0$:*

- *The rule $[n] \mapsto \mathrm{H}_i(\mathrm{GL}_{n,0}(\mathcal{O}, \mathfrak{a}), \boldsymbol{k})$ defines a finitely generated* OI*-module over $\boldsymbol{k}$.*
- *If $\boldsymbol{k}$ is a field then there is a polynomial $f \in \mathbb{Q}[t]$ such that $\dim \mathrm{H}_i(\mathrm{GL}_{n,0}(\mathcal{O}, \mathfrak{a}), \boldsymbol{k}) = f(n)$ for $n \gg 0$.*

1E. ***Outline.*** In Section 2 we review generalities on modules over categories. In Section 3 we introduce the category OI and its variants $\mathrm{OI}(d)$ and establish basic results about them. In Section 4 we introduce the category $\mathrm{OVI}(R)$ and its variants $\mathrm{OVI}(R, d)$ and establish basic results about them. In Section 5, we prove the main noetherianity result for $\mathrm{OVI}(R)$ (Theorem 1.5). In Section 6 we prove the main result of the paper (Theorem 1.4). Finally, in Section 7 we prove Theorem 1.8.

1F. ***Notation.*** Throughout, $\boldsymbol{k}$ denotes a commutative ring, typically noetherian. Unless otherwise specified, $1 \neq 0$ in all of our rings. For a fixed category $\mathcal{C}$, we write $\underline{\boldsymbol{k}}$ for the constant functor $\mathcal{C} \to \mathrm{Mod}_{\boldsymbol{k}}$ taking everything to $\boldsymbol{k}$ and all morphisms to the identity. We let $B_n \subset \mathrm{GL}_n$ be the group of upper-triangular matrices, and $U_n \subset B_n$ the subgroup where the diagonal entries are equal to 1. We use R to denote the ring appearing in the definition of $\mathrm{OVI}(R)$, and that is typically plugged in to U_n or B_n. We generally do not require it to be commutative. We set $[0] = \varnothing$, and if n is a positive integer, then $[n]$ denotes the set $\{1, \dots, n\}$.

2. Representations of categories

2A. ***Generalities.*** Let $\mathcal{C}$ be a category and let $\boldsymbol{k}$ be a noetherian commutative ring. A $\mathcal{C}$-*module* over $\boldsymbol{k}$ is a functor $M \colon \mathcal{C} \to \mathrm{Mod}_{\boldsymbol{k}}$. For an object $x \in \mathcal{C}$, we denote by M_x the image of x under M. Denote the category of $\mathcal{C}$-modules by $\mathrm{Rep}_{\boldsymbol{k}}(\mathcal{C})$. It is an abelian category. For each $x \in \mathcal{C}$, we define a $\mathcal{C}$-module P_x via the formula $(P_x)_y = \boldsymbol{k}[\mathrm{Hom}(x, y)]$. One easily sees that for any $\mathcal{C}$-module M one has a natural identification $\mathrm{Hom}(P_x, M) = M_x$. It follows that P_x is a projective $\mathcal{C}$-module; we call it the *principal projective* at x. A general $\mathcal{C}$-module M is finitely generated if and only if there exists a surjection $\bigoplus_{i=1}^k P_{x_i} \to M$ for some $x_1, \dots, x_k \in \mathcal{C}$. A $\mathcal{C}$-module is said to be *noetherian* if all of its submodules are finitely generated, and the category $\mathrm{Rep}_{\boldsymbol{k}}(\mathcal{C})$ is said to be *locally noetherian* if all finitely generated objects are noetherian.

If $\Phi \colon \mathcal{C} \to \mathcal{D}$ is a functor and M is a $\mathcal{D}$-module then the *pullback* of M along Φ, denoted $\Phi^*(M)$, is the $\mathcal{C}$-module defined via the formula $\Phi^*(M) = M \circ \Phi$, so that $\Phi^*(M)_x = M_{\Phi(x)}$. We now review how the pullback operation interacts with finite generation. The following definition is [Sam and Snowden 2017, Definition 3.2.1].

Definition 2.1. We say that a functor $\Phi\colon \mathcal{C} \to \mathcal{D}$ satisfies *property* (F) if the following condition holds for all $y \in \mathcal{D}$. There exist finitely many objects $x_1, \ldots, x_n \in \mathcal{C}$ together with morphisms $f_i\colon y \to \Phi(x_i)$ in $\mathcal{D}$ with the following property: for any $x \in \mathcal{C}$ and any morphism $f\colon y \to \Phi(x)$ in $\mathcal{D}$, there exists an i, and a morphism $g\colon x_i \to x$ in $\mathcal{C}$, such that $f = \Phi(g) \circ f_i$.

Definition 2.2. A category $\mathcal{C}$ satisfies property (F) if the diagonal $\mathcal{C} \to \mathcal{C} \times \mathcal{C}$ satisfies property (F).

The importance of these definitions is due to the following results.

Proposition 2.3. *A functor $\Phi\colon \mathcal{C} \to \mathcal{D}$ satisfies* property (F) *if and only if $\Phi^*(M)$ is a finitely generated $\mathcal{C}$-module for all finitely generated $\mathcal{D}$-modules M.*

Proof. See [Sam and Snowden 2017, Proposition 3.2.3]. □

Recall that a functor $\Phi\colon \mathcal{C} \to \mathcal{D}$ is *essentially surjective* if for all $y \in \mathcal{D}$, there exists some $x \in \mathcal{C}$ such that $\Phi(x)$ is isomorphic to y.

Proposition 2.4. *Let $\mathcal{C}$ be a category such that* $\mathrm{Rep}_{\boldsymbol{k}}(\mathcal{C})$ *is locally noetherian and let $\Phi\colon \mathcal{C} \to \mathcal{D}$ be an essentially surjective functor satisfying* property (F). *Then* $\mathrm{Rep}_{\boldsymbol{k}}(\mathcal{D})$ *is locally noetherian.*

Proof. See [Sam and Snowden 2017, Corollary 3.2.5]. □

If $\mathcal{C}$ is a category and M_1 and M_2 are $\mathcal{C}$-modules, then we define $M_1 \otimes M_2$ to be the $\mathcal{C}$-module defined by the formula $(M_1 \otimes M_2)_x = (M_1)_x \otimes (M_2)_x$ for all $x \in \mathcal{C}$.

Proposition 2.5. *Let $\mathcal{C}$ be a category that satisfies* property (F) *and let M and N be finitely generated $\mathcal{C}$-modules. Then $M \otimes N$ is finitely generated.*

Proof. See [Sam and Snowden 2017, Proposition 3.3.2]. □

We require a slight variant of the above proposition. We say that a $\mathcal{C}$-module M is *generated in finite degrees* if there exist $x_1, \ldots, x_k \in \mathcal{C}$ such that M is generated by the M_{x_i}, that is, the canonical map $\bigoplus_{i=1}^k M_{x_i} \otimes P_{x_i} \to M$ is surjective. Note that if M is generated in finite degrees and M_x is a finitely generated $\boldsymbol{k}$-module for all $x \in \mathcal{C}$ then M is finitely generated.

Proposition 2.6. *Let $\mathcal{C}$ be a category that satisfies* property (F) *and let M and N be $\mathcal{C}$-modules generated in finite degrees. Then $M \otimes N$ is generated in finite degrees.*

Proof. Observe that:

(a) A finite sum of $\mathcal{C}$-modules generated in finite degrees is generated in finite degrees.

(b) If K is a $\mathcal{C}$-module generated in finite degrees and U is any $\boldsymbol{k}$-module then $U \otimes K$ is generated in finite degrees.

(c) Any quotient of a $\mathcal{C}$-module generated in finite degrees is generated in finite degrees. Now, choose surjections $\bigoplus_{i=1}^k V_i \otimes P_{x_i} \to M$ and $\bigoplus_{j=1}^\ell W_j \otimes P_{y_j} \to N$, where the x_i and y_j are objects of $\mathcal{C}$ and the V_i and W_j are $\boldsymbol{k}$-modules (one can take $V_i = M_{x_i}$ and $W_j = N_{y_j}$). We thus have a surjection

$$\bigoplus_{i,j} V_i \otimes W_j \otimes P_{x_i} \otimes P_{y_j} \to M \otimes N.$$

Since $\mathcal{C}$ satisfies property (F), each $P_{x_i} \otimes P_{y_j}$ is finitely generated (Proposition 2.5). Thus each term in the sum is generated in finite degrees by (b); since the sum is finite, it is generated in finite degree by (a); and so we conclude $M \otimes N$ is generated in finite degrees by (c). □

Now we recall the notion of a Gröbner category. See [Sam and Snowden 2017, Section 4.3] for more details.

Definition 2.7. Let $\mathcal{C}$ be an essentially small category, i.e., there exists a set I containing a unique representative of each isomorphism class in $\mathcal{C}$. For $x \in \mathcal{C}$, define $|S_x| = \amalg_{y \in I} \operatorname{Hom}(x, y)$. Partially order $|S_x|$ by defining $f \preceq g$ if there exists a morphism h such that $g = hf$. We say that $\mathcal{C}$ is *Gröbner* if the following holds for all $x \in \mathcal{C}$:

- The poset $(|S_x|, \preceq)$ is noetherian.
- $|S_x|$ admits a total ordering $\leq$ with the following two properties:
  - The ordering $\leq$ is compatible with left composition, i.e., $f \leq g$ implies $hf \leq hg$.
  - The restriction of $\leq$ to each $\operatorname{Hom}(x, y)$ is a well-ordering.

We say that $\mathcal{C}$ is *quasi-Gröbner* if there exists a Gröbner category $\mathcal{C}'$ and an essentially surjective functor $\mathcal{C}' \to \mathcal{C}$ satisfying property (F).

The key result about quasi-Gröber categories is the following [Sam and Snowden 2017, Theorem 4.3.2]:

Theorem 2.8. *Let $\mathcal{C}$ be a quasi-Gröbner category. Then for any noetherian commutative ring $\boldsymbol{k}$, the category $\operatorname{Rep}_{\boldsymbol{k}}(\mathcal{C})$ is locally noetherian.*

2B. *Kan extension.* Let $\Phi \colon \mathcal{C} \to \mathcal{D}$ be a functor. The pullback functor Φ^* on modules admits a left adjoint $\Phi_!$ called the *left Kan extension*. It also admits a right adjoint Φ_* called the *right Kan extension*, but we will not need this.

The left Kan extension can be described explicitly as follows. Let y be an object of $\mathcal{D}$. Define a category $\mathcal{C}_{/y}$ as follows. An object of $\mathcal{C}_{/y}$ is a pair (x, f), where x is an object of $\mathcal{C}$ and $f \colon \Phi(x) \to y$ is a morphism in $\mathcal{D}$. A morphism $(x', f') \to (x, f)$ in $\mathcal{C}_{/y}$ is a morphism $g \colon x' \to x$ in $\mathcal{C}$ such that $f' = f \circ \Phi(g)$. Suppose now that M is a $\mathcal{C}$-module over $\boldsymbol{k}$. For $y \in \mathcal{D}$, define $M|_{\mathcal{C}_{/y}}$ to be the $\mathcal{C}_{/y}$-module defined via the formula $(M|_{\mathcal{C}_{/y}})_{(x,f)} = M_x$. We then have

$$\Phi_!(M)_y = \operatorname{colim}(M|_{\mathcal{C}_{/y}}).$$

That is, the value of $\Phi_!(M)$ on y is the colimit of the functor $M|_{\mathcal{C}_{/y}} \colon \mathcal{C}_{/y} \to \operatorname{Mod}_{\boldsymbol{k}}$. In certain cases, there is an even nicer description.

Proposition 2.9. *Let $\Phi \colon \mathcal{C} \to \mathcal{D}$ be a faithful functor. Assume that for all $x', x \in \mathcal{C}$, the $\operatorname{Aut}(\Phi(x))$-orbit of any element of $\operatorname{Hom}_{\mathcal{D}}(\Phi(x'), \Phi(x))$ contains an element of the form $\Phi(f)$ for some $f \in \operatorname{Hom}_{\mathcal{C}}(x', x)$ that is unique up to the action of $\operatorname{Aut}(x)$. Let M be a $\mathcal{C}$-module. Then for all $x \in \mathcal{C}$ we have a canonical isomorphism*

$$\Psi_!(M)_{\Phi(x)} = \operatorname{Ind}_{\operatorname{Aut}(x)}^{\operatorname{Aut}(\Phi(x))}(M_x).$$

Proof. Let $\{h_i\}_{i\in I}$ be a set of coset representatives for $\mathrm{Aut}(\Phi(x))/\mathrm{Aut}(x)$. For each $i \in I$, we thus have an object (x, h_i) of $\mathcal{C}_{/\Phi(x)}$. Consider an object (x', g) of $\mathcal{C}_{/\Phi(x)}$. To prove the proposition, it is enough to prove that there is a unique $i \in I$ and a unique morphism $(x', g) \to (x, h_i)$ of $\mathcal{C}_{/\Phi(x)}$.

By definition, g is a morphism $\Phi(x') \to \Phi(x)$ in $\mathcal{D}$. By assumption, we can factor g as $h\Phi(f)$ for some $h \in \mathrm{Aut}(\Phi(x'))$ and some $f \in \mathrm{Hom}_{\mathcal{C}}(x', x)$. Moreover, this factorization is unique up to the action of $\mathrm{Aut}(x)$. It follows that there is a unique factorization of the form $h_i\Phi(f)$. The morphism f now furnishes a map $(x', g) \to (x, h_i)$ in $\mathcal{C}_{/\Phi(x)}$. It is clear from the discussion that this is the unique i for which there is such a morphism, and that f is the unique such morphism. □

Left Kan extensions can be used to construct principal projectives, as follows. Let $x \in \mathcal{C}$, let pt be the point category (one object, one morphism), and let $i_x\colon \mathrm{pt} \to \mathcal{C}$ be the functor taking the object of pt to x. Regarding $\boldsymbol{k}$ as a pt-module, we have $(i_x)_!(\boldsymbol{k}) = P_x$. Indeed, if M is a $\mathcal{C}$-module, then by definition

$$\mathrm{Hom}_{\mathrm{Rep}_{\boldsymbol{k}}(\mathcal{C})}((i_x)_!(\boldsymbol{k}), M) = \mathrm{Hom}_{\mathrm{Rep}_{\boldsymbol{k}}(\mathrm{pt})}(\boldsymbol{k}, i_x^*(M)) = M_x,$$

and thus $(i_x)_!(\boldsymbol{k})$ represents the same functor as P_x.

Return now to the setting of a functor $\Phi\colon \mathcal{C} \to \mathcal{D}$. Put $y = \Phi(x)$. Then $\Phi \circ i_x = i_y$, so

$$P_y = (i_y)_!(\boldsymbol{k}) = \Phi_!((i_x)_!(\boldsymbol{k})) = \Phi_!(P_x). \tag{2.10}$$

We thus see that the left Kan extension takes principal projectives to principal projectives. Since $\Phi_!$ is right exact, it follows from this that $\Phi_!$ takes finitely generated $\mathcal{C}$-modules to finitely generated $\mathcal{D}$-modules.

2C. *$\mathcal{C}$-groups and their representations.* Let $\mathcal{C}$ be a category. A *$\mathcal{C}$-group* is a functor from $\mathcal{C}$ to the category of groups. Fix a $\mathcal{C}$-group $\boldsymbol{G}$. A *$\boldsymbol{G}$-module* over $\boldsymbol{k}$ is a $\mathcal{C}$-module M equipped with a $\boldsymbol{k}$-linear action of $\boldsymbol{G}_x$ on M_x for all $x \in \mathcal{C}$, such that for all morphisms $f\colon x \to y$ in $\mathcal{C}$ the induced morphism $f_*\colon M_x \to M_y$ is compatible with the actions via the induced homomorphism $f_*\colon \boldsymbol{G}_x \to \boldsymbol{G}_y$. In other words, for $m \in M_x$ and $g \in \boldsymbol{G}_x$ we have $f_*(gm) = f_*(g)f_*(m)$. The category $\mathrm{Rep}_{\boldsymbol{k}}(\boldsymbol{G})$ of $\boldsymbol{G}$-modules is a Grothendieck abelian category.

Let M be a $\boldsymbol{G}$-module. For $x \in \mathcal{C}$, let $\boldsymbol{H}_i(\boldsymbol{G}, M)_x$ be the group homology $\mathrm{H}_i(\boldsymbol{G}_x, M_x)$. If $f\colon x \to y$ is a morphism in $\mathcal{C}$, then the induced morphisms $f_*\colon \boldsymbol{G}_x \to \boldsymbol{G}_y$ and $f_*\colon M_x \to M_y$ together induce a morphism $f_*\colon \boldsymbol{H}_i(\boldsymbol{G}, M)_x \to \boldsymbol{H}_i(\boldsymbol{G}, M)_y$. This yields a $\mathcal{C}$-module structure on $\boldsymbol{H}_i(\boldsymbol{G}, M)$. If $\boldsymbol{k}$ is a commutative ring, then we will denote by $\underline{\boldsymbol{k}}$ the constant $\mathcal{C}$-module defined via the formula $\underline{\boldsymbol{k}}_x = \boldsymbol{k}$. We then have $\boldsymbol{H}_i(\boldsymbol{G}, \underline{\boldsymbol{k}})_x = \mathrm{H}_i(\boldsymbol{G}_x, \boldsymbol{k})$.

The following proposition concerns the homology of a semidirect product of $\mathcal{C}$-groups.

Proposition 2.11. *Let $\boldsymbol{G}$ and $\boldsymbol{E}$ be $\mathcal{C}$-groups, and let $\pi\colon \boldsymbol{G} \to \boldsymbol{E}$ and $\iota\colon \boldsymbol{E} \to \boldsymbol{G}$ be morphisms of $\mathcal{C}$-groups such that $\pi \circ \iota = \mathrm{id}$. Let $\boldsymbol{K} = \ker(\pi)$, which is also a $\mathcal{C}$-group. Then we have the following*:

(1) *$\boldsymbol{H}_i(\boldsymbol{K}, \underline{\boldsymbol{k}})$ is naturally an $\boldsymbol{E}$-module.*

(2) *As a $\mathcal{C}$-module, $\boldsymbol{H}_i(\boldsymbol{E}, \underline{\boldsymbol{k}})$ is a direct summand of $\boldsymbol{H}_i(\boldsymbol{G}, \underline{\boldsymbol{k}})$ via ι_* and π_*.*

(3) *Write* $\boldsymbol{H}_r(\boldsymbol{G}, \underline{\boldsymbol{k}}) = \mathrm{H}_r(\boldsymbol{E}, \underline{\boldsymbol{k}}) \oplus M$ *as in* (2). *Then* M *admits a* $\mathcal{C}$*-module filtration where the graded pieces are subquotients of* $\boldsymbol{H}_i(\boldsymbol{E}, \boldsymbol{H}_{r-i}(\boldsymbol{K}, \underline{\boldsymbol{k}}))$ *with* $0 \le i \le r-1$.

Proof. (1) The conjugation action of $\boldsymbol{G}$ on $\boldsymbol{K}$ is $\mathcal{C}$-linear. On homology, $\boldsymbol{K}$ acts trivially, and hence this action descends to give an $\boldsymbol{E}$-module structure on $\boldsymbol{H}_i(\boldsymbol{K}, \underline{\boldsymbol{k}})$.

(2) This is clear.

(3) For $x \in \mathcal{C}$ we have a short exact sequence of groups $1 \to \boldsymbol{K}_x \to \boldsymbol{G}_x \to \boldsymbol{E}_x \to 1$, which gives a Hochschild–Serre spectral sequence

$$\mathrm{E}^2_{p,q} = \boldsymbol{H}_p(\boldsymbol{E}_x, \boldsymbol{H}_q(\boldsymbol{K}_x, \boldsymbol{k})) \Rightarrow \boldsymbol{H}_{p+q}(\boldsymbol{G}_x, \boldsymbol{k}).$$

The spectral sequence is functorial in x, and so we get a spectral sequence of $\mathcal{C}$-modules

$$\mathrm{E}^2_{p,q} = \boldsymbol{H}_p(\boldsymbol{E}, \boldsymbol{H}_q(\boldsymbol{K}, \underline{\boldsymbol{k}})) \Rightarrow \boldsymbol{H}_{p+q}(\boldsymbol{G}, \underline{\boldsymbol{k}}).$$

In particular, $\boldsymbol{H}_r(\boldsymbol{G}, \underline{\boldsymbol{k}})$ has a filtration by subquotients of the terms $\mathrm{E}^2_{i,r-i}$. The edge map $\boldsymbol{H}_r(\boldsymbol{G}, \underline{\boldsymbol{k}}) \to \boldsymbol{H}_r(\boldsymbol{E}, \mathrm{H}_0(\boldsymbol{K}, \underline{\boldsymbol{k}}))$ coincides with the map on H_r induced by π (see [Weibel 1994, Section 6.8.2]) which we know is a split surjection, so the kernel M has a filtration by subquotients of $\mathrm{E}^2_{i,r-i}$ for $0 \le i \le r-1$. □

3. The category OI and variants

3A. ***Definitions and first results.*** Let OI be the category whose objects are finite totally ordered sets and whose morphisms are order-preserving injections. For a nonnegative integer d, we define a variant $\mathrm{OI}(d)$ as follows. An object of $\mathrm{OI}(d)$ is a pair (S, λ) where S is a totally ordered set and $\lambda = (\lambda_1 < \cdots < \lambda_d)$ is an increasing d-tuple in S. A morphism $(S, \lambda) \to (T, \mu)$ in $\mathrm{OI}(d)$ is an order-preserving injection $f\colon S \to T$ satisfying $f(\lambda) = \mu$. Note that $\mathrm{OI} = \mathrm{OI}(0)$. There is a functor $\Phi\colon \mathrm{OI}(d) \to \mathrm{OI}$ given by $\Phi(S, \lambda) = S$. We will continue to use the notation Φ for this functor throughout the paper (and use it for all values of d).

Remark 3.1. We introduce $\mathrm{OI}(d)$ to help us study an analogous category $\mathrm{OVI}(R, d)$, the motivation for which is discussed in Remark 4.1 below.

Recall that $[n]$ denotes the ordered set $\{1, \dots, n\}$. Given an OI-module M, we will write M_n for $M_{[n]}$. The category OI is equivalent to its full subcategory spanned by the $[n]$, so the data of an OI-module M is equivalent to the data of the M_n together with the maps $f_*\colon M_n \to M_m$ induced by the order preserving maps $f\colon [n] \to [m]$. Similarly, if M is an $\mathrm{OI}(d)$-module and λ is an increasing d-tuple in $[n]$, then we will write $M_{n,\lambda}$ for $M_{([n],\lambda)}$.

Proposition 3.2. *There is an equivalence of categories* $\mathrm{OI}(d) = \mathrm{OI}^{d+1}$.

Proof. Let (S, λ) be an object of $\mathrm{OI}(d)$. For $1 \le i \le d+1$, let S_i be the set of elements $x \in S$ such that $\lambda_{i-1} < x < \lambda_i$ where, by convention, $\lambda_0 < x < \lambda_{d+1}$ for all x. One easily verifies that $(S, \lambda) \mapsto (S_1, \dots, S_{d+1})$ is an equivalence. □

Corollary 3.3. *The category* $\mathrm{OI}(d)$ *is Gröbner. In particular, the category of* $\mathrm{OI}(d)$*-modules is locally noetherian.*

Proof. By [Sam and Snowden 2017, Theorem 7.1.2] the category OI is Gröbner, and by [loc. cit., Proposition 4.3.5] a finite product of Gröbner categories is Gröbner, so by Proposition 3.2 the category $\mathrm{OI}(d)$ is Gröbner. The assertion about finitely generated $\mathrm{OI}(d)$-modules now follows from Theorem 2.8. □

Corollary 3.4. *The category* $\mathrm{OI}(d)$ *satisfies* property (F). *In particular, the tensor product of finitely generated* $\mathrm{OI}(d)$*-modules is a finitely generated* $\mathrm{OI}(d)$*-module and the tensor product of* OI*-modules that are generated in finite degree is also generated in finite degree.*

Proof. The category OI satisfies property (F); this can be proved similarly to [loc. cit., Proposition 7.3.1]. One easily sees that a finite product of categories satisfying property (F) again satisfies property (F), which combined with Proposition 3.2 yields the fact that $\mathrm{OI}(d)$ satisfies property (F). The assertion about tensor products of finitely generated $\mathrm{OI}(d)$-modules now follows from Proposition 2.5, and the assertion about tensor products of OI-modules that are generated in finite degree follows from Proposition 2.6. □

Finally, we state a result about the growth of finitely generated OI-modules over fields.

Proposition 3.5. *Let M be a finitely generated* OI*-module over a field* $\boldsymbol{k}$. *Then the function* $n \mapsto \dim_{\boldsymbol{k}} M_n$ *is a polynomial function for* $n \gg 0$.

Proof. By [Sam and Snowden 2017, Theorem 7.1.2], OI is an "O-lingual category", and by [loc. cit., Theorem 6.3.2], this implies the polynomiality statement. □

3B. *Kan extension.* We now study left Kan extensions along the functor $\Phi\colon \mathrm{OI}(d) \to \mathrm{OI}$.

Proposition 3.6. *Let M be an* $\mathrm{OI}(d)$*-module. Then* $\Phi_!(M)_n = \bigoplus_\lambda M_{n,\lambda}$, *where the sum is taken over all increasing d-tuples λ in* $[n]$.

Proof. By Section 2B, we see that $\Phi_!(M)_n$ is $\mathrm{colim}(M|_{\mathrm{OI}(d)_{/[n]}})$. The category $\mathrm{OI}(d)_{/[n]}$ can be viewed as consisting of triples (S, μ, f), where $(S, \mu) \in \mathrm{OI}(d)$ and $f : S \to [n]$ is a morphism in OI. For an increasing d-tuple λ in $[n]$, let $\mathrm{OI}(d)_{/[n],\lambda}$ be the full subcategory of $\mathrm{OI}(d)_{/[n]}$ spanned by triples (S, μ, f) such that f takes μ to λ. Then $\mathrm{OI}(d)_{/[n]}$ is the disjoint union of its subcategories $\mathrm{OI}(d)_{/[n],\lambda}$. Furthermore, $([n], \lambda, \mathrm{id})$ is the final object of $\mathrm{OI}(d)_{/[n],\lambda}$. The result now follows. □

Corollary 3.7. *The functor* $\Phi_!$ *is exact.*

3C. *Shift functors.* Fix a functorial coproduct $\amalg$ on the category of finite sets. For finite sets S and T, we view $S \amalg T$ as the disjoint union of S and T; of course, this requires care when S and T share elements. Consider the functor $\Sigma_0\colon \mathrm{OI}(d) \to \mathrm{OI}(d)$ given by $\Sigma_0(S, \lambda) = (S \amalg \{\infty\}, \lambda)$, where $S \amalg \{\infty\}$ is given a total order by setting $x < \infty$ for all $x \in S$. Given an $\mathrm{OI}(d)$-module M, we define the *shift* of M, denoted $\Sigma(M)$, to be $\Sigma_0^*(M)$. There is a map $(S, \lambda) \to (S \amalg \{\infty\}, \lambda)$ in $\mathrm{OI}(d)$ induced by the inclusion $S \hookrightarrow S \amalg \{\infty\}$. This map induces a map $M \to \Sigma(M)$ of $\mathrm{OI}(d)$-modules. We let $\overline{\Sigma}(M)$ denote the cokernel of this map. We call it the *reduced shift* of M. This has the following nice property:

Proposition 3.8. *Suppose that M is an* OI*-module such that M_0 is a finitely generated $\boldsymbol{k}$-module and $\overline{\Sigma}(M)$ is a finitely generated* OI*-module. Then M is a finitely generated* OI*-module.*

Proof. By assumption, we can find $x_1, \dots, x_m$ with $x_i \in M_{n_i}$ such that the following holds. Let $\bar{x}_i \in \Sigma(M)_{n_i-1} \cong M_{n_i}$ be the associated element. Then the images of $\{\bar{x}_1, \dots, \bar{x}_m\}$ in $\overline{\Sigma}(M)$ generate $\overline{\Sigma}(M)$. We claim that $\{x_1, \dots, x_m\}$ together with a spanning set of M_0 is a generating set for M. Consider $y \in M_n$ for some $n \geq 0$. We must show that y is in the span of the indicated elements. We will do this by induction on n. The base case $n = 0$ being trivial, we can assume that $n \geq 1$. Let $\bar{y} \in \Sigma(M)_{n-1} \cong M_n$ be the associated element. The image of $\bar{y}$ in $\overline{\Sigma}(M)_n$ is in the span of the images of $\{\bar{x}_1, \dots, \bar{x}_m\}$. It follows that we can write $y = y' + y''$, where y' is in the span of $\{x_1, \dots, x_m\}$ and y'' is in the image of the composition $M_{n-1} \to \Sigma(M)_{n-1} \cong M_n$. By induction, y'' is in the span of $\{x_1, \dots, x_m\}$ together with a spanning set of M_0, so y is as well. $\square$

There is a similar functor $\Delta_0\colon \mathrm{OI}(d-1) \to \mathrm{OI}(d)$ defined by $\Delta_0(S, \lambda) = (S \amalg \{\infty\}, \lambda')$, where λ' is obtained by appending ∞ to the end of λ. For an $\mathrm{OI}(d)$-module M, we let $\Delta(M) = \Delta_0^*(M)$, which is an $\mathrm{OI}(d-1)$-module. For $d = 0$, we put $\Delta(M) = 0$ by convention.

The following result shows how the shift functor interacts with the Kan extension along the functor $\Phi\colon \mathrm{OI}(d) \to \mathrm{OI}$.

Proposition 3.9. *Let M be an* $\mathrm{OI}(d)$*-module. Then there is a natural isomorphism*

$$\Sigma(\Phi_!(M)) \cong \Phi_!(\Sigma(M)) \oplus \Phi_!(\Delta(M)).$$

Moreover, if $\alpha\colon \Phi_!(M) \to \Sigma(\Phi_!(M))$ and $\beta\colon M \to \Sigma(M)$ denote the natural maps, then the diagram

$$\begin{tikzcd} & \Phi_!(M) \arrow[dl, "\alpha"'] \arrow[dr, "\Phi_!(\beta)\oplus 0"] & \\ \Sigma(\Phi_!(M)) \arrow[rr, "\cong"] & & \Phi_!(\Sigma(M)) \oplus \Phi_!(\Delta(M)) \end{tikzcd}$$

commutes. In particular, we have a natural isomorphism

$$\overline{\Sigma}(\Phi_!(M)) = \Phi_!(\overline{\Sigma}(M)) \oplus \Phi_!(\Delta(M)).$$

Proof. Using Proposition 3.6, we have

$$\Sigma(\Phi_!(M))_n \cong \bigoplus_{\lambda} M_{n+1,\lambda},$$

where the sum is over all increasing d-tuples λ in $[n+1]$. Similarly, we have

$$\Phi_!(\Sigma(M))_n \cong \bigoplus_{\lambda} M_{n+1,\lambda},$$

where the sum is over all increasing d-tuples λ in $[n]$. Finally, using the obvious analog of Proposition 3.6 for Δ we have

$$\Phi_!(\Delta(M))_n \cong \bigoplus_{\lambda} M_{n+1,\lambda},$$

where the sum is over all increasing d-tuples λ in $[n]$ that end in $n+1$. Combining these isomorphisms, we obtain an identification

$$\Sigma(\Phi_!(M))_n \cong \Phi_!(\Sigma(M))_n \oplus \Phi_!(\Delta(M))_n.$$

It is clear that this identification comes from an isomorphism of OI-modules. The rest of the proposition follows easily. □

4. The category OVI and its variants

4A. ***Definitions.*** Fix a ring R (always assumed to be associative and unital, though not necessarily commutative). Define $\mathrm{OVI}(R)$ to be the following category. The objects are *ordered free R-modules*, that is, pairs $(V, \{v_i\}_{i\in I})$ where V is a finite rank free left R-module and $\{v_i\}$ is a basis indexed by a totally ordered set I. The morphisms $(V, \{v_i\}_{i\in I}) \to (W, \{w_j\}_{j\in J})$ are pairs (f, f_0), where $f\colon V \to W$ is a linear map and $f_0\colon I \to J$ is an order-preserving injection, such that $f(v_i) = w_{f_0(i)} + \sum_{j<f_0(i)} a_{i,j} w_j$ for scalars $a_{i,j}$. In words, f takes the i-th basis vector of V to the $f_0(i)$-th basis vector of W up to "lower order" terms. We note that f_0 can be recovered from f, so it is often omitted. Furthermore, f is necessarily a split injection. If the ring R is clear, we will just write OVI.

For a nonnegative integer n, we regard R^n as an ordered free module by endowing it with the standard basis. Every object of OVI is isomorphic to R^n for a unique n. For an OVI-module M, we write M_n for its value on R^n. The automorphism group of R^n in OVI is $U_n(R)$, which we denote simply by U_n in this section. It is the subgroup of $\mathrm{GL}_n(R)$ consisting of upper unitriangular matrices.

Let d be a nonnegative integer. We define a variant $\mathrm{OVI}(R, d) = \mathrm{OVI}(d)$ as follows. An object is a tuple $(V, \{v_i\}_{i\in I}, \lambda)$ where $(V, \{v_i\}_{i\in I})$ is an ordered free module and λ is an increasing d-tuple in I. A morphism $(V, \{v_i\}_{i\in I}, \lambda) \to (W, \{w_j\}_{j\in J}, \mu)$ is a morphism $(f, f_0)\colon (V, \{v_i\}) \to (W, \{w_j\})$ in OVI such that $f_0(\lambda) = \mu$ and such that $f(v_i) = w_{f_0(i)}$ for all i appearing in λ (i.e., no lower terms are allowed on marked basis vectors).

For a tuple $\lambda = (1 \le \lambda_1 < \cdots < \lambda_d \le n)$ we have an object (R^n, λ) of $\mathrm{OVI}(d)$. Every object of $\mathrm{OVI}(d)$ is isomorphic to a unique (R^n, λ). For an $\mathrm{OVI}(d)$-module M, we write $M_{n,\lambda}$ for its value on (R^n, λ). We let $U_{n,\lambda}$ be the automorphism group of (R^n, λ) in $\mathrm{OVI}(d)$. It is the subgroup of U_n fixing the basis vectors e_{λ_i} for $1 \le i \le d$.

Remark 4.1. We introduce $\mathrm{OVI}(d)$ as a technical device for proving Theorem 1.4, which concerns the homology groups $\boldsymbol{H}_i(\boldsymbol{U}, M)$ for $\mathrm{OVI}(R)$-modules M. We will see in Corollary 6.5 that the homology of the principal projective OVI module at d can be understood in terms of the homology of the trivial $\mathrm{OVI}(d)$-module, a helpful simplification.

There are several functors to mention:

- There is a functor $\mathrm{OI} \to \mathrm{OVI}$ taking a totally ordered set S to the ordered free module $R[S]$ with basis S. There is a similar functor $\mathrm{OI}(d) \to \mathrm{OVI}(d)$.

- There is a functor OVI $\to$ OI taking an ordered free module $(V, \{v_i\}_{i\in I})$ to the totally ordered set I and a morphism (f, f_0) to f_0. There is a similar functor $\mathrm{OVI}(d) \to \mathrm{OI}(d)$.
- There is a functor $\Psi\colon\ \mathrm{OVI}(d) \to \mathrm{OVI}$ given by forgetting λ. We continue to use the notation Ψ for this functor throughout the paper.

We have the following basic fact that follows from interpreting left multiplication by a matrix as a sequence of row operations.

Proposition 4.2. *Every morphism $\varphi\colon (R^n, \lambda) \to (R^m, \mu)$ in* $\mathrm{OI}(d)$ *has a unique factorization $\varphi = \psi f$ where $\psi \in \mathrm{Aut}(R^m, \mu)$ and f is in the image of the functor* $\mathrm{OI}(d) \to \mathrm{OVI}(d)$.

4B. ***The case where R is finite.*** The purpose of this section is to prove the following fundamental result:

Theorem 4.3. *If $|R| < \infty$, then the category* OVI *is quasi-Gröbner. In particular, by Theorem 2.8 the category* $\mathrm{Rep}_{\boldsymbol{k}}(\mathrm{OVI})$ *is locally noetherian when $\boldsymbol{k}$ is noetherian.*

Proof. An *ordered surjection* $f: S \to T$ of totally ordered finite sets is a surjection such that for all $i < j$ in T we have $\min f^{-1}(i) < \min f^{-1}(j)$. We let OS be the category whose objects are finite totally ordered sets and whose morphisms are ordered surjections. This category is known to be Gröbner [Sam and Snowden 2017, Theorem 8.1.1]. Given a totally ordered set S, we will regard the dual $R[S]^* = \mathrm{Hom}_R(R[S], R)$ as an element of OVI as follows. Let $S^* \subset R[S]^*$ be the dual basis to the basis S, and for $s \in S$, write $s^* \in S^*$ for the dual element. Then we order S^* via the rule

$$s_1^* < s_2^* \quad \text{when} \quad s_2 < s_1. \tag{4.3.a}$$

Using this convention, there is a functor $\mathrm{OS}^{\mathrm{op}} \to \mathrm{OVI}$ taking a totally ordered set S to $R[S]^*$ and an ordered surjection $T \to S$ to the dual of the induced surjective linear map $R[T] \to R[S]$. We will show that this functor satisfies property (F), which will complete the proof.

Let V be an object of OVI. Let $T_1, \ldots, T_n \in \mathrm{OS}$ be objects and $f_i: V \to R[T_i]^*$ be OVI-morphisms such that the f_i are an enumeration of all possible morphisms satisfying the following condition:

- The set T_i is a total ordering of a finite subset of V^* that spans V^* and $f_i: V \to R[T_i]^*$ is an OVI-morphism that is dual to the natural surjection $R[T_i] \to V$.

Since V is finite, there are only finitely many such f_i. Now consider some $S \in \mathrm{OS}$ and an OVI-morphism $f: V \to R[S]^*$. To prove that our functor satisfies property (F), it is enough to prove that for some $1 \le i \le n$ we can write $f = g \circ f_i$, where $g: R[T_i]^* \to R[S]^*$ is dual to an OS-morphism $S \to T_i$. Let $T \subset V^*$ be the image of S under the dual surjection $f^*: R[S] \to V^*$. Let $h: S \to T$ be the resulting surjection. Order T via the rule

$$t_1 < t_2 \quad \text{when} \quad \min h^{-1}(t_1) < \min h^{-1}(t_2), \tag{4.3.b}$$

which makes h an OS-morphism. Combining (4.3.b) with (4.3.a) (applied to order both S^* and T^*), we see that T^* has the ordering

$$t_1^* < t_2^* \quad \text{when} \quad \max\{s^* \mid s \in h^{-1}(t_1)\} < \max\{s^* \mid s \in h^{-1}(t_2)\}; \tag{4.3.c}$$

Let $g\colon R[T]^* \to R[S]^*$ be the OVI-morphism dual to h, so

$$g(t^*) = \sum_{s \in h^{-1}(t)} s^* \quad (t \in T). \tag{4.3.d}$$

Finally, let $F\colon V \to R[T]^*$ be the injection dual to the surjection $R[T] \to V^*$ induced by the inclusion $T \hookrightarrow V^*$, so $f = g \circ F$. The fact that f is an OVI-morphism together with (4.3.c) and (4.3.d) implies that F is an OVI-morphism. This implies that for some $1 \le i \le n$ we have $T = T_i$ and $F = f_i$, and we are done. □

Remark 4.4. By making use of a variant $\mathrm{OS}(d)$ of OS, one can prove a version of the above theorem for $\mathrm{OVI}(d)$. Since we do not need this, we omit the details.

4C. *Kan extension.* We now study left Kan extensions along the functor $\Psi\colon \mathrm{OVI}(d) \to \mathrm{OVI}$.

Proposition 4.5. *Let M be an* $\mathrm{OVI}(d)$*-module. Then*

$$\Psi_!(M)_n = \bigoplus_{\lambda} \mathrm{Ind}_{U_{n,\lambda}}^{U_n}(M_{n,\lambda}),$$

the sum taken over all increasing sequences $1 \le \lambda_1 < \cdots < \lambda_d \le n$.

Proof. Let $\mathrm{OVI}(d)'$ be the category whose objects are those of $\mathrm{OVI}(d)$ and where a morphism

$$(V, \{v_i\}_{i \in I}, \lambda) \to (W, \{w_j\}_{j \in J}, \mu)$$

is a morphism (f, f_0) as in OVI (ignoring the λ and μ) such that f_0 is a morphism in $\mathrm{OI}(d)$. The automorphism groups in $\mathrm{OVI}(d)'$ are the U_n. The functor Ψ factors as $\Psi_2 \circ \Psi_1$, where $\Psi_1\colon \mathrm{OVI}(d) \to \mathrm{OVI}(d)'$ and $\Psi_2\colon \mathrm{OVI}(d)' \to \mathrm{OVI}$ are the natural functors. Proposition 2.9 applies to the functor Ψ_1, and so we find

$$(\Psi_1)_!(M)_{n,\lambda} = \mathrm{Ind}_{U_{n,\lambda}}^{U_n}(M_{n,\lambda}).$$

Arguing exactly as in the proof of Proposition 3.6, we find

$$(\Psi_2)_!(N)_n = \bigoplus_{\lambda} N_{n,\lambda}$$

for any $\mathrm{OVI}(d)'$-module N. The result follows. □

4D. OVI*-modules and representations of* $\boldsymbol{U}$. Define an $\mathrm{OI}(d)$-group $\boldsymbol{U}_d$ by $(\boldsymbol{U}_d)_{n,\lambda} = U_{n,\lambda}$. If M is an $\mathrm{OVI}(d)$-module then we can regard it as an $\mathrm{OI}(d)$-module via the functor $\mathrm{OI}(d) \to \mathrm{OVI}(d)$, and as such it has the structure of a $\boldsymbol{U}_d$-module. We thus have a functor

$$\{\mathrm{OVI}(d)\text{-modules}\} \to \{\boldsymbol{U}_d\text{-modules}\}.$$

One can show that the above functor is fully faithful. We do not need this result, so we do not include a proof. We write $\boldsymbol{U}$ in place of $\boldsymbol{U}_0$.

5. Noetherianity of OVI-modules

The goal of this section is to prove Theorem 1.5, which we recall says that if R is a ring whose underlying additive group is finitely generated and $\boldsymbol{k}$ is a commutative noetherian ring, then the category of $\mathrm{OVI}(R)$-modules over $\boldsymbol{k}$ is locally noetherian, that is, any submodule of a finitely generated module is finitely generated. The ring R here is not required to be commutative. When R is finite, this follows from the much easier Theorem 4.3. We will also prove a converse to this result that says that (ignoring degenerate cases) the category $\mathrm{Rep}_{\boldsymbol{k}}(\mathrm{OVI}(R))$ is locally noetherian only if $\boldsymbol{k}$ is noetherian and the additive group of R is finitely generated. We thus have a complete characterization of when $\mathrm{Rep}_{\boldsymbol{k}}(\mathrm{OVI}(R))$ is locally noetherian.

This section has four subsections. We begin in Section 5A by describing a toy version of our proof. We then prove a technical ring-theoretic result in Section 5B. The proof of Theorem 1.5 is in the long Section 5C. Finally, in Section 5D we prove the aforementioned converse to Theorem 1.5.

5A. ***A toy version of Theorem 1.5.*** In the next sections, we prove Theorem 1.5. The proof is a bit lengthy and heavy on notation, but the idea behind it is not too complicated. In this section we sketch the proof of a simpler result that illustrates the main ideas.

Theorem 1.5 (with $R=\mathbb{Z}$) implies that the group algebra $\boldsymbol{k}[U_n(\mathbb{Z})]$ is left-noetherian, provided $\boldsymbol{k}$ is noetherian. Let us try to prove this for $n=3$. The group algebra can be identified, as a $\boldsymbol{k}$-module, with

$$Q=x_2y_3\boldsymbol{k}[x_1^{\pm 1},\, y_1^{\pm 1},\, y_2^{\pm 1}],$$

which we treat as a $\boldsymbol{k}$-submodule of the Laurent polynomial ring in the five variables. The monomials in this module correspond to the group elements in $\boldsymbol{k}[U_3(\mathbb{Z})]$; the exponents of the x's give the second column, while the exponents of the y's gives the third.

We must show that any $U_3(\mathbb{Z})$-submodule of Q is finitely generated. Let M be a given submodule. Let Q_+ be the $\boldsymbol{k}$-submodule of Q where only positive powers of the variables appear. We would like to associate to M a monomial ideal in Q_+, and then use the noetherianity of monomial ideals to conclude that M is finitely generated. By "ideal" here we really mean $\boldsymbol{k}[x_1, y_1, y_2]$-submodule. The obvious attempt at this is to first form $M_+=M\cap Q_+$ and then take its initial module $\mathrm{in}(M_+)$, the $\boldsymbol{k}$-span of the initial terms of its elements under some monomial order. The problem with this is that $\mathrm{in}(M_+)$ need not be an ideal. For example, suppose that M_+ contains the element $f=x_2y_3(y_2+1)$, with initial term $\mathrm{in}(f)=x_2y_2y_3$. Let's try to find $x_1\,\mathrm{in}(f)$ in $\mathrm{in}(M_+)$. If we apply the matrix

$$\begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix}$$

to f, we get the element $f'=x_1x_2y_3(y_1y_2+1)$, with initial term $x_1x_2y_1y_2y_3$. This is equal to $x_1y_1\,\mathrm{in}(f)$, so we now need to get rid of the y_1. We therefore apply the matrix

$$\begin{pmatrix}1&0&-1\\0&1&0\\0&0&1\end{pmatrix}$$

to f', to get the element $f'' = x_1x_2y_1^{-1}y_3(y_1y_2+1)$. This has the correct leading term. However, it no longer belongs to M_+: the power of y in the nonleading term is negative. Thus $\mathrm{in}(f'')$ does not give an element of $\mathrm{in}(M_+)$. There does not seem to be a way to produce $x_1\,\mathrm{in}(f)$ in $\mathrm{in}(M_+)$.

Remark 5.1. This approach is really attempting to show that the monoid algebra $Q_+ = \boldsymbol{k}[U_3(\mathbb{Z}_{\geq 0})]$ is noetherian. In fact, it is not noetherian. For example, the left ideal generated by the matrices

$$\begin{pmatrix} 1 & n & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

for $n \geq 0$ in Q_+ is not finitely generated.

To overcome this problem, we take a more subtle approach. Let Q_* be the submodule of Q where the exponent of y_2 is positive, but we still allow negative powers of x_1 and y_1. Given $M \subset Q$, let $M_* = M \cap Q_*$. We can then form the initial module with respect to y_2 (that is, we treat the other variables as constants); call this $\mathrm{in}_2(M_*)$. Since we allow negative powers of y_1, the issue in the previous paragraph does not arise, and $\mathrm{in}_2(M_*)$ is closed under multiplication by $x_1^{\pm 1}$, $y_1^{\pm 1}$, and y_2. We now intersect $\mathrm{in}_2(M_*)$ with M_+ and then take initial terms with respect to x_1 and y_1. The result is a monomial ideal of Q_+. Call this monomial ideal $I(M)$. One can show that if $M \subset M'$ and $I(M) = I(M')$ then $M = M'$. Since Q_+ is noetherian as a $\boldsymbol{k}[x_1, y_1, y_2]$-module, this proves that Q is noetherian as a $\boldsymbol{k}[U_3(\mathbb{Z})]$-module.

The same approach works for $\boldsymbol{k}[U_n(\mathbb{Z})]$, but the process is more involved. Let Q be the group algebra, which we identify with a $\boldsymbol{k}$-submodule of the Laurent polynomial ring in variables $x_{i,j}$ with $i \leq j$. We let $Q^{(k)}$ be the $\boldsymbol{k}$-submodule where the exponents of $x_{i,j}$ with $i \geq k$ are positive. Thus $Q^{(n)} = Q$ and $Q^{(0)}$ is what we would call Q_+. Let M be a $U_n(\mathbb{Z})$-submodule of Q. We obtain a monomial ideal in Q_+ as follows: intersect with $Q^{(n-1)}$ and take the initial submodule with respect to $x_{\bullet,n}$; then intersect with $Q^{(n-2)}$ and take the initial submodule with respect to $x_{\bullet,n-1}$; and so on. After n steps we obtain a monomial ideal in Q_+. The argument then proceeds as in the previous case.

Remark 5.2. The strategy employed here has some parallels with Hall's proof [1954, Lemma 3] that the group ring $\boldsymbol{k}[\Gamma]$ of a polycyclic group Γ is noetherian. There the key point is to take a normal subgroup Γ' such that $\Gamma/\Gamma' \cong \mathbb{Z}$ and treat each element of $\boldsymbol{k}[\Gamma]$ as a Laurent polynomial in x with coefficients in $\boldsymbol{k}[\Gamma']$ (where x is some generator for $\mathbb{Z}$) and argue by passing to initial terms.

The proof for $\mathrm{OVI}(R)$ differs from the above in only two respects. First, there is a great deal of additional bookkeeping. Second, we need a noetherianity result for the kind of OI-monomial ideals that appear in the reduction. This follows easily from Higman's lemma, and is closely related to the theorem [Cohen 1967; Aschenbrenner and Hillar 2007; Hillar and Sullivant 2012] that $\boldsymbol{k}[x_i]_{i\in\mathbb{N}}$ is $\mathrm{Inc}(\mathbb{N})$-noetherian, where $\mathrm{Inc}(\mathbb{N})$ is the monoid of increasing functions $\mathbb{N} \to \mathbb{N}$.

5B. ***Eliminating additive torsion.*** For technical reasons, Theorem 1.5 is easier to prove when R is a ring whose additive group is a finitely generated free abelian group. In this section, we show how to reduce to that case. Our main tool is the following lemma.

Lemma 5.3. *Let S be a ring and let $\boldsymbol{k}$ be a commutative ring such that the category of* $\mathrm{OVI}(S)$*-modules over $\boldsymbol{k}$ is locally noetherian. Assume that S surjects onto a ring R. Then the category of* $\mathrm{OVI}(R)$*-modules over $\boldsymbol{k}$ is locally noetherian.*

Proof. The surjection $S \to R$ induces a functor $\Phi\colon \mathrm{OVI}(S) \to \mathrm{OVI}(R)$. By Proposition 2.4, it is enough to show that Φ satisfies property (F). For some $d \geq 1$, let P_d be the principal projective $\mathrm{OVI}(R)$-module associated to R^d, so

$$(P_d)_n = \boldsymbol{k}[\mathrm{Hom}_{\mathrm{OVI}(R)}(R^d, R^n)] \quad (n \geq 1).$$

By Proposition 2.3, to prove that Φ satisfies property (F) it is enough to prove that $\Phi^*(P_d)$ is finitely generated. Since the map $S \to R$ of rings is surjective, the induced map

$$\mathrm{Hom}_{\mathrm{OVI}(S)}(S^d, S^n) \to \mathrm{Hom}_{\mathrm{OVI}(R)}(R^d, R^n)$$

is also surjective for all $n \geq 1$. This implies that there is a surjective map from the principal projective $\mathrm{OVI}(S)$-module associated to S^d to $\Phi^*(P_d)$, and thus that $\Phi^*(P_d)$ is finitely generated, as desired. □

Lemma 5.4. *Let R be a ring whose additive group is finitely generated. Then there exists a ring S and a surjection $S \to R$ such that the additive group of S is free and finitely generated.*

Proof. Let R_{tor} be the torsion subgroup of the additive group of R and let $N \geq 1$ be the exponent of R_{tor}, i.e., the minimal number such that $N R_{\mathrm{tor}} = 0$. The proof is by induction on N. In the base case where $N = 1$, the group R_{tor} is trivial and there is nothing to prove. Assume, therefore, that $N > 1$ and that the lemma is true for all smaller exponents. Let p be a prime dividing N. The ring R/pR is a finite ring. Let $\mathbb{Z}[R/pR]$ be the monoid ring of the multiplicative monoid underlying R/pR, so $\mathbb{Z}[R/pR]$ consists of finite sums of formal symbols $\{[x] \mid x \in R/pR\}$ with the ring structure defined by $[x][y] = [xy]$. The additive group of the ring $\mathbb{Z}[R/pR]$ is free abelian with basis in bijection with the elements of R/pR, and there exists a ring surjection $\mathbb{Z}[R/pR] \to R/pR$ taking $[x] \in \mathbb{Z}[R/pR]$ to $x \in R/pR$. Let R' be the fiber product of the surjections $\mathbb{Z}[R/pR] \to R/pR$ and $R \to R/pR$, so we have a cartesian square

$$\begin{tikzcd}
R' \arrow[r] \arrow[d] & \mathbb{Z}[R/pR] \arrow[d] \\
R \arrow[r] & R/pR.
\end{tikzcd}$$

Concretely,

$$R' = \{(x, r) \in \mathbb{Z}[R/pR] \times R \mid x \text{ and } r \text{ map to same element of } R/pR\}.$$

Since the maps $R \to R/pR$ and $\mathbb{Z}[R/pR] \to R/pR$ are surjective, so is the map $R' \to R$. Since the additive group underlying $\mathbb{Z}[R/pR]$ is torsion-free, the torsion subgroup $(R')_{\mathrm{tor}}$ consists of pairs $(0, r) \in \mathbb{Z}[R/pR] \times R_{\mathrm{tor}}$ such that $r \in R_{\mathrm{tor}}$ maps to 0 in R/pR. It follows that

$$(R')_{\mathrm{tor}} \cong R_{\mathrm{tor}} \cap pR = pR_{\mathrm{tor}}.$$

The exponent of $(R')_{\mathrm{tor}}$ is thus N/p, so by induction there exists a ring S whose additive group is finitely generated and free together with a surjection $S \to R'$. The desired surjection to R is then the composition $S \to R' \to R$. □

5C. *The proof of Theorem 1.5.* We now commence with the proof of Theorem 1.5, which we recall says that if R is a ring whose underlying additive group is finitely generated and $\boldsymbol{k}$ is a commutative noetherian ring, then the category of $\mathrm{OVI}(R)$-modules over $\boldsymbol{k}$ is locally noetherian. By Lemmas 5.3 and 5.4, we can assume that the additive group of R is a finitely generated free abelian group (this assumption will first be used in Substep 2a below). Fix some $d \geq 0$ and let P_d be the principal projective of $\mathrm{OVI}(R)$ defined by the formula

$$(P_d)_n = \boldsymbol{k}[\mathrm{Hom}_{\mathrm{OVI}(R)}(R^d, R^n)] \quad (n \geq 1).$$

To prove the theorem, it is enough to prove that the poset of $\mathrm{OVI}(R)$-submodules of P_d is noetherian, i.e., has no infinite strictly increasing sequences. This is trivial for $d = 0$, so we can assume that $d \geq 1$.

Say that a map $f : I \to J$ of posets is *conservative* if for all $i, i' \in I$ satisfying $i \leq i'$ and $f(i) = f(i')$, we have $i = i'$. If J is a noetherian poset and $f : I \to J$ is a conservative map, then I is also noetherian. Our strategy will be to use a sequence of conservative poset maps to reduce proving that the poset of $\mathrm{OVI}(R)$-submodules of P_d is noetherian to proving that another easier poset $\mathfrak{M}^{(0)}$ is noetherian. To help the reader understand its structure, we divide our proof into three steps (each of which is divided into a number of substeps).

Since we will introduce a lot of notation, to help the reader recall the meanings of symbols we will list the notation that is defined in each substep.

Step 1. We construct a poset $\mathfrak{M}$ and reduce the theorem to showing that $\mathfrak{M}$ is noetherian.

As in the toy version of our proof, the first step will be to relate the poset of $\mathrm{OVI}(R)$-submodules of P_d to a poset $\mathfrak{M}$ constructed using certain "generalized polynomial rings". In fact, $\mathfrak{M}$ will be a poset of certain special $\mathrm{OI}(d)$-submodules of an $\mathrm{OI}(d)$-module Q. There are three substeps: in Substep 1a we construct the $\mathrm{OI}(d)$-module Q, in Substep 1b we construct the poset $\mathfrak{M}$ of special $\mathrm{OI}(d)$-submodules of Q, and then finally in Substep 1c we construct a conservative poset map from the poset of $\mathrm{OVI}(R)$-submodules of P_d to $\mathfrak{M}$.

Substep 1a. We construct the $\mathrm{OI}(d)$-module Q.
Notation defined: Λ_n, $T^r_{i,j}$, $\boldsymbol{T}_n$, $\Lambda_n(\boldsymbol{S})$, $\Lambda_{n,\alpha}$, $\boldsymbol{T}_{n,\alpha}$, Q, $Q_{n,\alpha}$

We will want to view matrices with entries in R as certain kinds of "monomials". Since we will be focusing on P_d, the relevant matrices will have d columns and some number $n \geq 1$ of rows. To that end, we make the following definition:

- Define Λ_n to be the commutative monoid generated by the set of formal symbols $T^r_{i,j}$ with $1 \leq i \leq n$ and $1 \leq j \leq d$ and $r \in R$ subject to the relations $T^{r_1}_{i,j} T^{r_2}_{i,j} = T^{r_1+r_2}_{i,j}$, where $1 \leq i \leq n$ and $1 \leq j \leq d$ and $r_1, r_2 \in R$.

Elements of Λ_n are thus "monomials" in the $T_{i,j}^r$, and are naturally in bijection with $n \times d$ matrices with entries in R: given such a matrix $(r_{i,j})$, the associated element of Λ_n is the product of the $T_{i,j}^{r_{i,j}}$, where i ranges over $1 \leq i \leq n$ and j ranges over $1 \leq j \leq d$. The monoid product in Λ_n corresponds to matrix addition. For later use, setting $\boldsymbol{T}_n = \{T_{i,j} \mid 1 \leq i \leq n,\ 1 \leq j \leq d\}$, for $\boldsymbol{S} \subset \boldsymbol{T}_n$ we define $\Lambda_n(\boldsymbol{S})$ to be the submonoid of Λ_n generated by $\{T_{i,j}^r \mid T_{i,j} \in \boldsymbol{S},\ r \in R\}$.

Now consider an element $f \in \mathrm{Hom}_{\mathrm{OVI}(R)}(R^d, R^n)$. By definition, f is a linear map $R^d \to R^n$ such that there exists a strictly increasing sequence $\alpha = (\alpha_1, \ldots, \alpha_d)$ of d elements of $[n] = \{1, \ldots, n\}$ with the following property:

- For $1 \leq i \leq d$, the map f takes the i-th basis element of R^d to the sum of the α_i-th basis element of R^n and an R-linear combination of the basis elements of R^n that occur before α_i.

Define $\Lambda_{n,\alpha}$ to be the subset of Λ_n consisting of elements associated to $n \times d$ matrices of this form. Defining

$$\boldsymbol{T}_{n,\alpha} = \{T_{i,j} \mid 1 \leq j \leq d,\ 1 \leq i < \alpha_j\},$$

an element $\tau \in \Lambda_{n,\alpha}$ can be written as

$$\tau = T_{\alpha_1,1}^1 T_{\alpha_2,2}^1 \cdots T_{\alpha_d,d}^1 \tau' \quad \text{with } \tau' \in \Lambda_n(\boldsymbol{T}_{n,\alpha}). \tag{5.5}$$

We thus have a bijection of sets

$$\mathrm{Hom}_{\mathrm{OVI}(R)}(R^d, R^n) \cong \bigsqcup_{\alpha} \Lambda_{n,\alpha},$$

where the disjoint union ranges over the strictly increasing sequences α of d elements of $[n]$. It follows that

$$(P_d)_n = \boldsymbol{k}[\mathrm{Hom}_{\mathrm{OVI}(R)}(R^d, R^n)] = \bigoplus_{\alpha} \boldsymbol{k}[\Lambda_{n,\alpha}]. \tag{5.6}$$

The various $\boldsymbol{k}[\Lambda_{n,\alpha}]$ fit together into an $\mathrm{OI}(d)$-module Q with

$$Q_{n,\alpha} = \boldsymbol{k}[\Lambda_{n,\alpha}] \quad ((n, \alpha) \in \mathrm{OI}(d)).$$

Substep 1b. We construct a poset $\mathfrak{M}$ of $\mathrm{OI}(d)$-submodules of Q.
Notation defined: $\mathfrak{M}$, E_{i,α_j}^r

Consider an $\mathrm{OVI}(R)$-submodule M of P_d. We say that M is a *homogeneous* $\mathrm{OVI}(R)$*-submodule of* P_d if for all $n \geq 1$, the $\boldsymbol{k}$-submodule M_n of $(P_d)_n$ splits according to the decomposition (5.6), i.e., for all $(n, \alpha) \in \mathrm{OI}(d)$ there exists some $\boldsymbol{k}$-submodule $M_{n,\alpha}$ of $\boldsymbol{k}[\Lambda_{n,\alpha}]$ such that

$$M_n = \bigoplus_{\alpha} M_{n,\alpha}.$$

In this case, the various $M_{n,\alpha}$ fit together into an $\mathrm{OI}(d)$-submodule of Q. We thus get a poset injection

$$\{\text{homogeneous } \mathrm{OVI}(R)\text{-submodules of } P_d\} \hookrightarrow \{\mathrm{OI}(d)\text{-submodules of } Q\}.$$

The image of this injection consists of all $\mathrm{OI}(d)$-submodules M of Q such that each $M_{n,\alpha} \subset Q_{n,\alpha}$ is preserved by the action of $U_n(R)$, which acts on $Q_{n,\alpha}$ via the identification of $Q_{n,\alpha}$ with the set of formal $\boldsymbol{k}$-linear combinations of appropriate $n \times d$ matrices.

For the sake of our later arguments, we will actually consider a larger collection of submodules. Define $\mathfrak{M}$ to be the poset of all $\mathrm{OI}(d)$-submodules M of Q such that the following hold. Consider $(n,\alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$. Let $\{\vec{e}_1, \ldots, \vec{e}_n\}$ be the standard basis for R^n. For $1 \le j \le d$ and $1 \le i < \alpha_j$ and $r \in R$, define $E^r_{i,\alpha_j} \in U_n(R)$ to be the element that takes $\vec{e}_{\alpha_j}$ to $r\vec{e}_i + \vec{e}_{\alpha_j}$ and fixes all of the other basis vectors. We then require that $M_{n,\alpha}$ be preserved by all of the E^r_{i,α_j} for $1 \le j \le d$ and $1 \le i < \alpha_j$ and $r \in R$. The construction in the previous paragraph gives a poset injection

$$\{\text{homogeneous } \mathrm{OVI}(R)\text{-submodules of } P_d\} \hookrightarrow \mathfrak{M}. \tag{5.7}$$

Substep 1c. We construct a conservative poset map $\{\mathrm{OVI}(R)\text{-submodules of } P_d\} \to \mathfrak{M}$.
Notation defined: none

By (5.7), it is enough to construct a conservative poset map

$$\{\mathrm{OVI}(R)\text{-submodules of } P_d\} \to \{\text{homogeneous } \mathrm{OVI}(R)\text{-submodules of } P_d\}. \tag{5.8}$$

For each $n \ge 1$, put a total ordering on the set of all strictly increasing sequences α of d elements of $[n]$ using the lexicographic ordering: $\alpha < \alpha'$ if the first nonzero entry $\alpha' - \alpha$ is positive. Given a nonzero element $f \in (P_d)_n$, use the identification (5.6) to write $f = \sum_\alpha f_{n,\alpha}$ with $f_{n,\alpha} \in \boldsymbol{k}[\Lambda_{n,\alpha}]$. Define $\mathrm{in}(f) = f_{n,\alpha_0}$, where α_0 is the largest index such that $f_{n,\alpha_0} \neq 0$.

Given an $\mathrm{OVI}(R)$-submodule M of P_d and some $n \ge 1$, define $\mathrm{in}(M)_n$ to be the $\boldsymbol{k}$-span of $\{\mathrm{in}(f) \mid f \in M_n\}$. It is easy to see that $\mathrm{in}(M)$ is also an $\mathrm{OVI}(R)$-submodule of P_d. Moreover, by construction $\mathrm{in}(M)$ is homogeneous. The map $M \mapsto \mathrm{in}(M)$ is thus a poset map as in (5.8). We must prove that it is conservative. Assume otherwise, and let M and M' be $\mathrm{OVI}(R)$-submodules of P_d such that $M \subsetneq M'$ and $\mathrm{in}(M) = \mathrm{in}(M')$. Let $n \ge 1$ be such that $M_n \subsetneq M'_n$. Let $x \in M'_n \setminus M_n$ be such that $\mathrm{in}(x)$ lies in $\boldsymbol{k}[\Lambda_{n,\alpha}]$ with α as small as possible. Since $\mathrm{in}(M) = \mathrm{in}(M')$, we can find some $x' \in M_n$ with $\mathrm{in}(x) = \mathrm{in}(x')$. But then $x - x' \in M'_n \setminus M_n$, while $\mathrm{in}(x - x')$ lies in $\boldsymbol{k}[\Lambda_{n,\alpha'}]$ with $\alpha' < \alpha$, a contradiction.

Step 2. We construct a poset $\mathfrak{M}^{(0)}$ and reduce the theorem to showing that $\mathfrak{M}^{(0)}$ is noetherian.

In Step 1, we reduced the theorem to showing that the poset $\mathfrak{M}$ constructed in Substep 1b is noetherian. The goal of this step is to construct a conservative poset map from $\mathfrak{M}$ to a simpler poset $\mathfrak{M}^{(0)}$. This will be done in a sequence of steps. Recall that $\mathfrak{M}$ is a subposet of the poset of $\mathrm{OI}(d)$-submodules of an $\mathrm{OI}(d)$-module Q. In Substep 2a we will construct an $\mathrm{OI}(d)$-module filtration

$$Q^{(0)} \subset Q^{(1)} \subset \cdots \subset Q^{(d)} = Q.$$

Next, in Substeps 2b and 2c we will construct two posets $\mathfrak{M}^{(k)}$ and $\mathfrak{N}^{(k)}$ of special $\mathrm{OI}(d)$-submodules of $Q^{(k)}$ such that $\mathfrak{M}^{(d)} = \mathfrak{M}$. Finally, in Substeps 2d and 2e we will construct a sequence of conservative poset maps

$$\mathfrak{M} = \mathfrak{M}^{(d)} \to \mathfrak{N}^{(d-1)} \to \mathfrak{M}^{(d-1)} \to \mathfrak{N}^{(d-2)} \to \cdots \to \mathfrak{N}^{(0)} \to \mathfrak{M}^{(0)}.$$

This reduces the theorem to showing that the poset $\mathfrak{M}^{(0)}$ is noetherian.

Substep 2a. We construct an $\mathrm{OI}(d)$-module filtration

$$Q^{(0)} \subset Q^{(1)} \subset \cdots \subset Q^{(d)} = Q.$$

Notation defined: $(R, +) = (\mathbb{Z}^\lambda, +)$, $R_{\geq 0}$, $\Lambda_{n,\alpha,k+}$, $\Lambda_{n,\alpha,+}$, $\Lambda_{n,+}$, $\Lambda_{n,+}(\boldsymbol{S})$, $Q^{(k)}$, $Q^{(k)}_{n,\alpha}$

This step is where we use the fact that the additive group of R is a finitely generated free abelian group. Fix an identification of this additive group with $\mathbb{Z}^\lambda$ for some $\lambda \geq 1$ such that the multiplicative identity $1 \in R$ is identified with an element of $(\mathbb{Z}_{\geq 0})^\lambda$. Let $R_{\geq 0}$ be the submonoid of the additive group of R corresponding to $(\mathbb{Z}_{\geq 0})^\lambda$. The monoid $R_{\geq 0}$ contains $1 \in R$, but is not necessarily closed under multiplication.

Consider $(n, \alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$. For $0 \leq k \leq d$, define $\Lambda_{n,\alpha,k+}$ to be the set of all $\tau \in \Lambda_{n,\alpha}$ such that if $T^r_{i,j}$ appears in τ with $i \geq \alpha_k$, then $r \in R_{\geq 0}$. For $k = 0$, we use the convention $\alpha_0 = 0$, and we will also frequently omit the k, so $\Lambda_{n,\alpha,+}$ is the set of all $\tau \in \Lambda_{n,\alpha}$ such that if $T^r_{i,j}$ appears in τ, then $r \in R_{\geq 0}$. We will similarly define $\Lambda_{n,+}$ and $\Lambda_{n,+}(\boldsymbol{S})$ for $\boldsymbol{S} \subset \boldsymbol{T}_n$. We then define $Q^{(k)}$ to be the $\mathrm{OI}(d)$-submodule of Q where for all $(n, \alpha) \in \mathrm{OI}(d)$, we have

$$Q^{(k)}_{n,\alpha} = \boldsymbol{k}[\Lambda_{n,\alpha,k+}].$$

We thus have $Q^{(d)} = Q$. Moreover,

$$Q^{(0)}_{n,\alpha} = \boldsymbol{k}[\Lambda_{n,\alpha,+}].$$

Substep 2b. For $0 \leq k \leq d$, we construct a subposet $\mathfrak{M}^{(k)}$ of the poset of $\mathrm{OI}(d)$-submodules of $Q^{(k)}$ such that $\mathfrak{M}^{(d)} = \mathfrak{M}$.

Notation defined: $\mathfrak{M}^{(k)}$, (a.i$_k$), (a.ii$_k$), (b$_k$), (c$_k$)

We begin with some terminology. A $\boldsymbol{k}$-submodule X of $\boldsymbol{k}[\Lambda_n]$ is *homogeneous* with respect to $\boldsymbol{S} \subset \boldsymbol{T}_n$ if the following holds for all $x \in X$. Write

$$x = \sum_{q=1}^{m} \tau_q y_q,$$

where for all $1 \leq q \leq m$ we have the following:

- $\tau_q \in \Lambda_n(\boldsymbol{S})$, and the different τ_q are all distinct.
- $y_q \in \boldsymbol{k}[\Lambda_n(\boldsymbol{T}_n \setminus \boldsymbol{S})]$.

We then require that $\tau_q y_q \in X$ for all $1 \leq q \leq m$.

Now consider some $0 \leq k \leq d$. Define $\mathfrak{M}^{(k)}$ to be the set of all $\mathrm{OI}(d)$-submodules M of $Q^{(k)}$ such that for all $(n, \alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$, the following conditions (a.i$_k$), (a.ii$_k$), (b$_k$), and (c$_k$) hold. To simplify our notation, we will set $\alpha_0 = 0$.

(a) The $\boldsymbol{k}$-module $M_{n,\alpha} \subset \boldsymbol{k}[\Lambda_{n,\alpha,k+}]$ is closed under multiplication by the following elements:

(i$_k$) $T_{i,j}^r$ with $k \leq j \leq d$ and $1 \leq i < \alpha_k$ and $r \in R$.

(ii$_k$) $T_{i,j}^r$ with $k+1 \leq j \leq d$ and $\alpha_k \leq i < \alpha_j$ and $r \in R_{\geq 0}$.

(b$_k$) The $\boldsymbol{k}$-module $M_{n,\alpha}$ is closed under the operators E_{i,α_j}^r with $1 \leq j \leq k$ and $1 \leq i < \alpha_j$ and $r \in R$.

(c$_k$) The $\boldsymbol{k}$-module $M_{n,\alpha} \subset \boldsymbol{k}[\Lambda_{n,\alpha,k+}]$ is homogeneous with respect to

$$\{T_{\alpha_j,j} \mid 1 \leq j \leq d\} \cup \{T_{i,j} \mid k+1 \leq j \leq d \text{ and } \max(\alpha_k, 1) \leq i < \alpha_j\}.$$

We claim that $\mathfrak{M}^{(d)} = \mathfrak{M}$. Condition (b$_d$) implies that $\mathfrak{M}^{(d)} \subset \mathfrak{M}$, so we must only prove that $\mathfrak{M} \subset \mathfrak{M}^{(d)}$. Consider $M \in \mathfrak{M}$ and $(n, \alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$. We must verify that $M_{n,\alpha}$ satisfies the properties above:

- For (a.i$_d$), we must show that $M_{n,\alpha}$ is closed under multiplication by $T_{i,d}^r$ for $1 \leq i < \alpha_d$ and $r \in R$. But this can be achieved using the operator E_{i,α_d}^r, and by the definition of $\mathfrak{M}$ the $\boldsymbol{k}$-module $M_{n,\alpha}$ is closed under this operator, so (a.i$_d$) follows.
- No pairs (i, j) satisfy the conditions of (a.ii$_d$), so that condition is trivial.
- Condition (b$_d$) is a special case of the condition defining $\mathfrak{M}$, so it follows.
- The set referred to in condition (c$_d$) consists only of

$$\{T_{\alpha_j,j} \mid 1 \leq j \leq d\},$$

and by definition every element of $\boldsymbol{k}[\Lambda_{n,\alpha}]$ is homogeneous with respect to these variables (see (5.5)), so that condition follows.

Substep 2c. For $0 \leq k < d$, we construct a subposet $\mathfrak{N}^{(k)}$ of the poset of $\mathrm{OI}(d)$-submodules of $Q^{(k)}$.
Notation defined: $\mathfrak{N}^{(k)}$, (a$'$.i$'_k$), (a$'$.ii$'_k$), (b$'_k$), (c$'_k$)

Our definition of $\mathfrak{N}^{(k)}$ will be a slight modification of our definition of $\mathfrak{M}^{(k)}$. Define $\mathfrak{N}^{(k)}$ to be the set of all $\mathrm{OI}(d)$-submodules N of $Q^{(k)}$ such that for all $(n, \alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$, the following conditions (a$'$.i$'_k$), (a$'$.ii$'_k$), (b$'_k$), and (c$'_k$) hold. To simplify our notation, we will set $\alpha_0 = 0$.

(a$'$) The $\boldsymbol{k}$-module $N_{n,\alpha} \subset \boldsymbol{k}[\Lambda_{n,\alpha,k+}]$ is closed under multiplication by the following elements:

(i$'_\mathrm{k}$) $T_{i,j}^r$ with $k+1 \leq j \leq d$ and $1 \leq i < \alpha_k$ and $r \in R$.

(ii$'_\mathrm{k}$) $T_{i,j}^r$ with $k+1 \leq j \leq d$ and $\alpha_k \leq i < \alpha_j$ and $r \in R_{\geq 0}$.

(b$'_\mathrm{k}$) The $\boldsymbol{k}$-module $N_{n,\alpha}$ is closed under the operators E_{i,α_j}^r with $1 \leq j \leq k$ and $1 \leq i < \alpha_j$ and $r \in R$.

(c$'_\mathrm{k}$) The $\boldsymbol{k}$-module $N_{n,\alpha} \subset \boldsymbol{k}[\Lambda_{n,\alpha,k+}]$ is homogeneous with respect to

$$\{T_{\alpha_j,j} \mid 1 \leq j \leq d\} \cup \{T_{i,j} \mid k+2 \leq j \leq d \text{ and } \alpha_{k+1} \leq i < \alpha_j\}.$$

Substep 2d. For $1 \leq k \leq d$, we construct a conservative poset map $\mathfrak{M}^{(k)} \to \mathfrak{N}^{(k-1)}$.
Notation defined: none.

Consider $M \in \mathfrak{M}^{(k)}$, so M is an OI(d)-submodule of $Q^{(k)}$. Define $N = M \cap Q^{(k-1)}$. We claim that $N \in \mathfrak{N}^{(k-1)}$. This requires checking the conditions (a$'$.i$'_{k-1}$), (a$'$.ii$'_{k-1}$), (b$'_{k-1}$), and (c$'_{k-1}$). Consider some $(n, \alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$:

- Condition (a$'$.i$'_{k-1}$) asserts that $N_{n,\alpha}$ is closed under multiplication by $T^r_{i,j}$ with $k \leq j \leq d$ and $1 \leq i < \alpha_{k-1}$ and $r \in R$. This follows from the fact that both $M_{n,\alpha}$ and $Q^{(k-1)}_{n,\alpha}$ are closed under multiplication by these elements. This is immediate for $Q^{(k-1)}_{n,\alpha}$. For $M_{n,\alpha}$, it follows from (a.i$_k$), which says that $M_{n,\alpha}$ is closed under multiplication by $T^r_{i,j}$ with $k \leq j \leq d$ and $1 \leq i < \alpha_k$ and $r \in R$.
- Condition (a$'$.ii$'_{k-1}$) asserts that $N_{n,\alpha}$ is closed under multiplication by $T^r_{i,j}$ with $k \leq j \leq d$ and $\alpha_{k-1} \leq i < \alpha_j$ and $r \in R_{\geq 0}$. This follows from the fact that both $M_{n,\alpha}$ and $Q^{(k-1)}_{n,\alpha}$ are closed under multiplication by these elements. This is immediate for $Q^{(k-1)}_{n,\alpha}$. For $M_{n,\alpha}$, it follows from a combination of (a.i$_k$), which handles the cases where $\alpha_{k-1} \leq i < \alpha_k$ and gives the stronger conclusion that we can use $r \in R$ instead of just $r \in R_{\geq 0}$, and (a.ii$_k$), which handles the cases where $\alpha_k \leq i < \alpha_j$. Here one might worry that (a.ii$_k$) requires $k + 1 \leq j \leq d$ instead of $k \leq j \leq d$; however, the case $j = k$ is not needed since no i satisfies $\alpha_k \leq i < \alpha_k$.
- Condition (b$'_{k-1}$) asserts that $N_{n,\alpha}$ is closed under the operators E^r_{i,α_j} with $1 \leq j \leq k-1$ and $1 \leq i < \alpha_j$ and $r \in R$. This follows from the fact that both $M_{n,\alpha}$ and $Q^{(k-1)}_{n,\alpha}$ are closed under these operators. This is immediate for $Q^{(k-1)}_{n,\alpha}$. For $M_{n,\alpha}$, it follows from (b$_k$), which says that $M_{n,\alpha}$ is closed under the operators E^r_{i,α_j} with $1 \leq j \leq k$ and $1 \leq i < \alpha_j$ and $r \in R$.
- Condition (c$'_{k-1}$) asserts that $N_{n,\alpha}$ is homogeneous with respect to

  $$\{T_{\alpha_j,j} \mid 1 \leq j \leq d\} \cup \{T_{i,j} \mid k+1 \leq j \leq d \text{ and } \alpha_k \leq i < \alpha_j\}.$$

  Condition (c$_k$) says that $M_{n,\alpha}$ is homogeneous with respect to this same set, and this homogeneity is preserved when we intersect $M_{n,\alpha}$ with $Q^{(k-1)}_{n,\alpha}$.

We thus can define a poset map $\mathfrak{M}^{(k)} \to \mathfrak{N}^{(k-1)}$ taking $M \in \mathfrak{M}^{(k)}$ to $M \cap Q^{(k-1)}$. We claim that this poset map is conservative. In fact, it is even injective. Indeed, consider $M, M' \in \mathfrak{M}^{(k)}$. Let $N = M \cap Q^{(k-1)}$ and $N' = M' \cap Q^{(k-1)}$, and assume that $N = N'$. We claim that $M = M'$. By symmetry, it is enough to prove that $M \subset M'$. Consider $(n, \alpha) \in \mathrm{OI}(d)$ and $x \in M_{n,\alpha}$. We must prove that $x \in M'_{n,\alpha}$. We have $x \in Q^{(k)}_{n,\alpha}$. Setting

$$\boldsymbol{S} = \{T_{i,j} \mid 1 \leq j \leq d,\ 1 \leq i < \alpha_j,\ \alpha_{k-1} \leq i < \alpha_k\} = \{T_{i,j} \mid k \leq j \leq d,\ \alpha_{k-1} \leq i < \alpha_k\},$$

there exists some $\tau \in \Lambda_{n,\alpha}(\boldsymbol{S})$ such that $\tau x \in Q^{(k-1)}_{n,\alpha}$. By (a.i$_k$), we have $\tau x \in M_{n,\alpha}$, and thus $\tau x \in N_{n,\alpha}$. Since $N = N' \subset M'$, we deduce that $\tau x \in M'_{n,\alpha}$. Define $\tau^{-1} \in \Lambda_{n,\alpha}(\boldsymbol{S})$ to be the result of replacing all the $T^r_{i,j}$ terms in τ with $T^{-r}_{i,j}$. Another application of (a.i$_k$) shows that $\tau^{-1}\tau x = x \in M'_{n,\alpha}$, as desired.

Substep 2e. For $0 \leq k \leq d-1$, we construct a conservative poset map $\mathfrak{N}^{(k)} \to \mathfrak{M}^{(k)}$.
Notation defined: none.

Fix some $(n,\alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$. The most important difference between $\mathfrak{M}^{(k)}$ and $\mathfrak{N}^{(k)}$ is that by (c$_k$) the $\boldsymbol{k}$-modules making up $\mathfrak{M}^{(k)}$ must be homogeneous with respect to

$$\boldsymbol{S}_{n,\alpha,k} = \{T_{\alpha_j,j} \mid 1 \le j \le d\} \cup \{T_{i,j} \mid k+1 \le j \le d \text{ and } \alpha_k \le i < \alpha_j \text{ and } i \ge 1\},$$

while by (c$'_k$) the $\boldsymbol{k}$-modules making up $\mathfrak{N}^{(k)}$ must only be homogeneous with respect to

$$\boldsymbol{S}_{n,\alpha,k+1} = \{T_{\alpha_j,j} \mid 1 \le j \le d\} \cup \{T_{i,j} \mid k+2 \le j \le d \text{ and } \alpha_{k+1} \le i < \alpha_j\}.$$

The main function of our poset map $\mathfrak{N}^{(k)} \to \mathfrak{M}^{(k)}$ will be to achieve the needed increase in homogeneity.

For $x \in Q^{(k)}_{n,\alpha}$, we will define an "initial term" $\mathrm{in}(x) \in Q^{(k)}_{n,\alpha}$ as follows. Define

$$\boldsymbol{S}'_{n,\alpha,k} = \boldsymbol{S}_{n,\alpha,k} \setminus \boldsymbol{S}_{n,\alpha,k+1} = \{T_{i,j} \mid k+1 \le j \le d \text{ and } \max(\alpha_k, 1) \le i < \alpha_{k+1}\}.$$

Recall that R is identified as an additive group with $\mathbb{Z}^\lambda$ and that $R_{\ge 0} = (\mathbb{Z}_{\ge 0})^\lambda \subset R$. Using the identification $R = \mathbb{Z}^\lambda$, we will frequently speak of the coordinates of elements of R. We define a total order on $\Lambda_{n,+}(\boldsymbol{S}'_{n,\alpha,k})$ in two steps:

- We first order $\boldsymbol{S}'_{n,\alpha,k}$ by letting $T_{i,j} < T_{i',j'}$ if either $i < i'$ or if $i = i'$ and $j < j'$.
- We then order $\Lambda_{n,+}(\boldsymbol{S}'_{n,\alpha,k})$ as follows. Consider distinct $\tau, \tau' \in \Lambda_{n,+}(\boldsymbol{S}'_{n,\alpha,k})$. Enumerating the elements of $\boldsymbol{S}'_{n,\alpha,k}$ in increasing order as $T_{i_1,j_1}, \ldots, T_{i_p,j_p}$, we can uniquely write

  $$\tau = T^{r_1}_{i_1,j_1} \cdots T^{r_p}_{i_p,j_p} \quad \text{and} \quad \tau' = T^{r'_1}_{i_1,j_1} \cdots T^{r'_p}_{i_p,j_p}.$$

  for some $r_i, r'_j \in R_{\ge 0}$. Let $1 \le q \le p$ be the minimal number such that $r_q \neq r'_q$. We then say that $\tau < \tau'$ if the first nonzero coordinate of $r'_q - r_q \in R = \mathbb{Z}^\lambda$ is positive.

For nonzero $x \in Q^{(k)}_{n,\alpha}$, we can uniquely write

$$x = \sum_{q=1}^{m} \tau_q y_q,$$

where for all $1 \le q \le m$ we have the following:

- $\tau_q y_q \neq 0$ for all q.
- $\tau_q \in \Lambda_{n,+}(\boldsymbol{S}'_{n,\alpha_k})$, and the τ_q are enumerated in increasing order $\tau_1 < \tau_2 < \cdots < \tau_m$.
- $y_q \in \boldsymbol{k}[\Lambda_n(\boldsymbol{T}_n \setminus \boldsymbol{S}'_{n,\alpha,k})]$.

We then define $\mathrm{in}(x) = \tau_m y_m \in Q^{(k)}_{n,\alpha}$. We also set $\mathrm{in}(0) = 0$. We will call τ_m the *initial variable* of x, though we remark that this terminology will not be used again until the final paragraph of this substep.

We now construct the poset map $\mathfrak{N}^{(k)} \to \mathfrak{M}^{(k)}$ as follows. Consider $N \in \mathfrak{N}^{(k)}$. For $(n,\alpha) \in \mathrm{OI}(d)$, define $\mathrm{in}(N)_{n,\alpha} \subset Q^{(k)}_{n,\alpha}$ to be the $\boldsymbol{k}$-span of $\{\mathrm{in}(x) \mid x \in N_{n,\alpha}\}$. It is easy to see that $\mathrm{in}(N)$ is an $\mathrm{OI}(d)$-submodule of $Q^{(k)}$. We claim that $\mathrm{in}(N) \in \mathfrak{M}^{(k)}$. To see this, we must check the conditions (a.i$_k$), (a.ii$_k$), (b$_k$), and (c$_k$). Consider some $(n,\alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$:

- We delay (a.i$_k$) until the end, so we start by verifying condition (a.ii$_k$), which asserts that $\text{in}(N)_{n,\alpha}$ is closed under multiplication by $T^r_{i,j}$ with $k+1 \le j \le d$ and $\alpha_k \le i < \alpha_j$ and $r \in R_{\ge 0}$. This is immediate from (a$'$.ii$'_k$), which asserts that N is closed under multiplication by these same elements.

- Condition (b$_k$) asserts that $\text{in}(N)_{n,\alpha}$ is closed under the operators E^r_{i,α_j} with $1 \le j \le k$ and $1 \le i < \alpha_j$ and $r \in R$. Condition (b$'_k$) says that $N_{n,\alpha}$ is closed under these operators. To prove that this implies that $\text{in}(N)_{n,\alpha}$ is also closed under these operators, it is enough to prove that for $x \in Q^{(k)}_{n,\alpha}$, we have

$$\text{in}(E^r_{i,\alpha_j}(x)) = E^r_{i,\alpha_j}(\text{in}(x)).$$

To help the reader understand the argument below, we recommend reviewing the correspondence between elements of Λ_n and $n \times d$ matrices from Substep 1a. For nonzero x, write

$$x = \sum_{q=1}^{m} \tau_q y_q,$$

where for all $1 \le q \le m$ we have the following:

- $\tau_q y_q \neq 0$ for all q.
- $\tau_q \in \Lambda_{n,+}(\boldsymbol{S}'_{n,\alpha,k})$, and the τ_q are enumerated in increasing order $\tau_1 < \tau_2 < \cdots < \tau_m$.
- $y_q \in \boldsymbol{k}[\Lambda_n(\boldsymbol{T}_n \setminus \boldsymbol{S}'_{n,\alpha,k})]$.

Since $i < \alpha_j \le \alpha_k$, for all $1 \le q \le m$ we have

$$E^r_{i,\alpha_j}(\tau_q) = \tau_q \tau'_q \quad \text{and} \quad E^r_{i,\alpha_j}(y_q) = y_q y'_q$$

for some $\tau'_q \in \Lambda_n(\boldsymbol{T}_n \setminus \boldsymbol{S}'_{n,\alpha,k})$ and $y'_q \in \boldsymbol{k}[\Lambda_n(\boldsymbol{T}_n \setminus \boldsymbol{S}'_{n,\alpha,k})]$. We thus have

$$E^r_{i,\alpha_j}(x) = \sum_{q=1}^{m} E^r_{i,\alpha_j}(\tau_q) \cdot E^r_{i,\alpha_j}(y_q) = \sum_{q=1}^{m} \tau_q(\tau'_q y_q y'_q)$$

and

$$\text{in}(E^r_{i,\alpha_j}(x)) = \tau_m(\tau'_m y_m y'_m) = E^r_{i,\alpha_j}(\text{in}(x)),$$

as desired.

- Condition (c$_k$) asserts that $\text{in}(N)_{n,\alpha}$ is homogeneous with respect to

$$\boldsymbol{S}_{n,\alpha,k} = \{T_{\alpha_j,j} \mid 1 \le j \le d\} \cup \{T_{i,j} \mid k+1 \le j \le d \text{ and } \alpha_k \le i < \alpha_j \text{ and } i \ge 1\}.$$

By (c$'_k$), the $\boldsymbol{k}$-module $N_{n,\alpha}$ is homogeneous with respect to $\boldsymbol{S}_{n,\alpha,k+1}$, and the very definition of $\text{in}(N)_{n,\alpha}$ is designed to improve this to $\boldsymbol{S}_{n,\alpha,k}$.

- We now finally verify (a.i$_k$), which asserts that $\text{in}(N)_{n,\alpha}$ is closed under multiplication by $T^r_{i,j}$ with $k \le j \le d$ and $1 \le i < \alpha_k$ and $r \in R$. Condition (a$'$.i$'_k$) says that $N_{n,\alpha}$ is closed under multiplication by $T^r_{i,j}$ with $k+1 \le j \le d$ and $1 \le i \le \alpha_j$ and $r \in R$, and this is preserved when we pass to $\text{in}(N)_{n,\alpha}$. We thus must only verify that $\text{in}(N)_{n,\alpha}$ is closed under multiplication by $T^r_{i,k}$ with $1 \le i < \alpha_k$ and $r \in R$.

Consider some $x \in \mathrm{in}(N)_{n,\alpha}$. We must show that $T^r_{i,k} x \in \mathrm{in}(N)_{n,\alpha}$. Using the already verified condition (c_k), we can assume that $x = \tau y$ with

$$\tau \in \Lambda_{n,\alpha}(\boldsymbol{S}_{n,\alpha,k}) \quad \text{and} \quad y \in \boldsymbol{k}[\Lambda_{n,\alpha}(\boldsymbol{T}_n \setminus \boldsymbol{S}_{n,\alpha,k})].$$

Using the already verified condition (b_k), we know that $E^r_{i,\alpha_k}(x) \in \mathrm{in}(N)_{n,\alpha}$. We then calculate that

$$E^r_{i,\alpha_k}(x) = E^r_{i,\alpha_k}(\tau y) = E^r_{i,\alpha_k}(\tau) E^r_{i,\alpha_k}(y) = (\tau T^r_{i,k} \tau') \cdot y,$$

where τ' is a product of elements of $\{T^{r'}_{i,j'} \mid k+1 \le j' \le d,\ r' \in R\}$ that depends on τ and r and i and k. Letting $(\tau')^{-1}$ be the result of replacing each $T^{r'}_{i,j'}$ in τ' with $T^{-r'}_{i,j'}$, our already verified cases of ($\mathrm{a.i}_k$) imply that $\mathrm{in}(N)_{n,\alpha}$ is closed under multiplication by $(\tau')^{-1}$. In particular,

$$(\tau')^{-1} \cdot E^r_{i,\alpha_k}(x) = (\tau')^{-1} \cdot (T^r_{i,k} \tau' \tau) \cdot y = T^r_{i,k} \tau y = T^r_{i,k} x \in \mathrm{in}(N)_{n,\alpha},$$

as desired.

The map $N \mapsto \mathrm{in}(N)$ is thus a poset map from $\mathfrak{N}^{(k)}$ to $\mathfrak{M}^{(k)}$.

We claim that this is a conservative poset map. Indeed, consider $N_1, N_2 \in \mathfrak{N}^{(k)}$ such that $N_1 \subset N_2$ and $\mathrm{in}(N_1) = \mathrm{in}(N_2)$. We must prove that $N_1 = N_2$. Assume otherwise. Let $(n,\alpha) \in \mathrm{OI}(d)$ be such that $(N_1)_{n,\alpha} \subsetneq (N_2)_{n,\alpha}$. Pick $x \in (N_2)_{n,\alpha}$ such that $x \notin (N_1)_{n,\alpha}$ and such that the initial variable (see the second paragraph of this substep for the definition of this) of x is as small as possible among elements with these properties (this is possible since with the above ordering $\Lambda_{n,+}(\boldsymbol{S}'_{n,\alpha,k})$ does not have any infinite strictly decreasing chains). Since $\mathrm{in}(N_1) = \mathrm{in}(N_2)$, we can find some $x' \in (N_1)_{n,\alpha}$ such that $\mathrm{in}(x') = \mathrm{in}(x)$. But then $x - x' \in (N_2)_{n,\alpha}$ and $x - x' \notin (N_1)_{n,\alpha}$, while the initial variable of $x - x'$ is strictly smaller than the initial variable of x, a contradiction.

Step 3. We prove that $\mathfrak{M}^{(0)}$ is noetherian.

In Step 2, we reduced the theorem to showing that $\mathfrak{M}^{(0)}$ is noetherian. In this step, we will prove this. Defining

$$\Lambda_+ = \bigsqcup_{(n,\alpha) \in \mathrm{OI}(d)} \Lambda_{n,\alpha,+},$$

in Substep 3a we first construct a useful partial ordering on Λ_+ and prove that it is a well partial ordering (see below for the definition of this). In Substep 3b, we use this partial ordering to prove that $\mathfrak{M}^{(0)}$ is noetherian.

Substep 3a. We construct a partial ordering on Λ_+ and prove that it is a well partial ordering.
Notation defined: none.

We define a partial ordering on Λ_+ as follows. Consider $\tau, \tau' \in \Lambda_+$. We say that $\tau \preceq \tau'$ if the following condition is satisfied:

- Let $(n,\alpha), (n',\alpha') \in \mathrm{OI}(d)$ be such that $\tau \in \Lambda_{n,\alpha,+}$ and $\tau' \in \Lambda_{n',\alpha',+}$. We then require that there exists an $\mathrm{OI}(d)$-morphism $\iota\colon (n,\alpha) \to (n',\alpha')$ and some $\tau'' \in \Lambda_{n',\alpha',+}$ such that $\tau' = \tau'' \cdot \iota_*(\tau)$.

It is clear that this is a partial ordering.

The main goal of this substep (which we will accomplish at the end after a number of preliminaries) is to prove that this partial ordering on Λ_+ is a well partial ordering, whose definition is as follows. A poset $(\mathfrak{P}, \prec)$ is *well partially ordered* if every infinite sequence of elements of $\mathfrak{P}$ contains an infinite weakly increasing subsequence. See [Kruskal 1972] for a survey about well partial orderings. If $\mathfrak{P}$ and $\mathfrak{P}'$ are posets, then we will endow $\mathfrak{P} \times \mathfrak{P}'$ with the ordering where $(p_1, p_1') \preceq (p_2, p_2')$ if and only $p_1 \preceq p_2$ and $p_1' \preceq p_2'$. If $\mathfrak{P}$ and $\mathfrak{P}'$ are both well partially ordered, then so is $\mathfrak{P} \times \mathfrak{P}'$ (quick proof: given an infinite sequence in $\mathfrak{P} \times \mathfrak{P}'$, first pass to a subsequence to make the first coordinate weakly increasing, then pass to a further subsequence to make the second coordinate also weakly increasing).

Recall that we have identified the additive group of R with $\mathbb{Z}^\lambda$ and that $R_{\geq 0} = (\mathbb{Z}_{\geq 0})^\lambda$. Using these identifications, we will speak of the coordinates of elements of R and $R_{\geq 0}$. Endow the set $R_{\geq 0} \cup \{\spadesuit\}$ with the following partial ordering:

- $\spadesuit$ is not comparable to any element of $R_{\geq 0}$.
- For $r_1, r_2 \in R_{\geq 0}$, let $r_1 \preceq r_2$ if all the coordinates of $r_2 - r_1$ are nonnegative.

Since the usual ordering on $\mathbb{Z}_{\geq 0}$ is a well partial ordering, the restriction of our partial ordering to $R_{\geq 0} = (\mathbb{Z}_{\geq 0})^\lambda$ is also a well partial ordering. From this, it is easy to see that our partial ordering on $R_{\geq 0} \cup \{\spadesuit\}$ is also a well partial ordering. The product ordering on $(R_{\geq 0} \cup \{\spadesuit\})^d$ is thus also a well partial ordering.

Let $\mathcal{W}$ denote the set of finite words in the alphabet $(R_{\geq 0} \cup \{\spadesuit\})^d$. Endow $\mathcal{W}$ with the partial ordering where $w_1, w_2 \in \mathcal{W}$ satisfy $w_1 \preceq w_2$ if and only if the following condition is satisfied. Write $w_1 = \ell_1 \cdots \ell_n$ and $w_2 = \ell_1' \cdots \ell_{n'}'$ with each ℓ_i and $\ell_{i'}'$ an element of $(R_{\geq 0} \cup \{\spadesuit\})^d$. We then require that there exists a strictly increasing function $\iota \colon [n] \hookrightarrow [n']$ such that $\ell_i \preceq \ell_{\iota(i)}'$ for all $1 \leq i \leq n$. This partial ordering on $\mathcal{W}$ is a well partial ordering by Higman's lemma [1952, Theorem 4.3].

As promised, we now prove that the partial ordering on Λ_+ defined above is a well partial ordering. Let $\Psi \colon \Lambda_+ \to \mathcal{W}$ be the following set function. Consider $\tau \in \Lambda_{n,\alpha,+} \subset \Lambda_+$. Write $\alpha = (\alpha_1, \dots, \alpha_d)$, and expand out τ as

$$\tau = \prod_{\substack{1 \leq j \leq d \\ 1 \leq i \leq \alpha_j}} T_{i,j}^{r_{i,j}} \quad (r_{i,j} \in R_{\geq 0}).$$

For $1 \leq j \leq d$ and $1 \leq i \leq \alpha_j$, define $\bar{r}_{i,j} \in R_{\geq 0} \cup \{\spadesuit\}$ via the formula

$$\bar{r}_{i,j} = \begin{cases} r_{i,j} & \text{if } 1 \leq i < \alpha_j, \\ \spadesuit & \text{if } i = \alpha_j. \end{cases}$$

We remark that by definition we have $r_{\alpha_j,j} = 1$ for all $1 \leq j \leq d$. For $1 \leq i \leq n$, we define

$$\ell_i = (\bar{r}_{i,1}, \bar{r}_{i,2}, \dots, \bar{r}_{i,d}) \in (R_{\geq 0} \cup \{\spadesuit\})^d.$$

Finally, we define

$$\Psi(\tau) = \ell_1 \ell_2 \cdots \ell_n.$$

It is clear that Ψ is injective. What is more, it is immediate from the definitions that for all $\tau, \tau' \in \Lambda_+$ we have

$$\tau \preceq \tau' \quad \text{if and only if } \Psi(\tau) \preceq \Psi(\tau').$$

The key point here is that if we interpret elements of Λ_+ as matrices with d columns and entries in R_+, the effect of an $\mathrm{OI}(d)$-morphism on these matrices is to insert extra rows of zeros. Since Ψ is injective and $\mathcal{W}$ is well partially ordered, so is Λ_+, as claimed.

Substep 3b. We prove that the poset $\mathfrak{M}^{(0)}$ is noetherian.
Notation defined: none.

Let $(\Lambda_+, <)$ be the partially ordered set constructed in Substep 3a. By definition, $\mathfrak{M}^{(0)}$ is the poset of all $\mathrm{OI}(R)$-modules $M \subset Q^{(0)}$ such that for all $(n, \alpha) \in \mathrm{OI}(d)$ with $\alpha = (\alpha_1, \ldots, \alpha_d)$, the $\boldsymbol{k}$-module $M_{n,\alpha} \subset \boldsymbol{k}[\Lambda_{n,\alpha,+}]$ satisfies the following two properties:

(†) It is closed under multiplication by $T_{i,j}^r$ for all $1 \leq j \leq d$ and $1 \leq i < \alpha_j$ and $r \in R_{\geq 0}$.

(††) It is homogeneous with respect to all the possible $T_{i,j}$, i.e., with respect to

$$\{T_{i,j} \mid 1 \leq j \leq d \text{ and } 1 \leq i \leq \alpha_j\}.$$

Property (††) implies that $M_{n,\alpha}$ is spanned as a $\boldsymbol{k}$-module by elements of the form $c \cdot \tau$ with $c \in \boldsymbol{k}$ and $\tau \in \Lambda_{n,\alpha,+}$. Property (†) implies the following:

(† † †) Let $\tau_1 \in \Lambda_{n_1,\alpha_1,+} \subset \Lambda_+$ and $\tau_2 \in \Lambda_{n_2,\alpha_2,+} \subset \Lambda_+$ and $c \in \boldsymbol{k}$ be such that $c \cdot \tau_1 \in M_{n_1,\alpha_1}$ and $\tau_1 \leq \tau_2$. Then $c \cdot \tau_2 \in M_{n_2,\alpha_2}$.

Now assume for the sake of contradiction that $\mathfrak{M}^{(0)}$ is not noetherian. Let

$$M_1 \subsetneq M_2 \subsetneq M_3 \subsetneq \cdots$$

be an infinite strictly ascending chain in it. By (††), for all $i \geq 1$ there exists some $(n_i, \alpha_i) \in \mathrm{OI}(d)$ and some $\tau_i \in \Lambda_{n_i,\alpha_i,+}$ and some $c_i \in \boldsymbol{k}$ such that

$$c_i \cdot \tau_i \in (M_i)_{n_i,\alpha_i} \setminus (M_{i-1})_{n_i,\alpha_i}. \tag{5.9}$$

Since our partial ordering on Λ_+ is a well partial ordering, we can replace our sequence $\{M_i\}_{i=1}^{\infty}$ with a subsequence and assume that

$$\tau_1 \leq \tau_2 \leq \tau_3 \leq \cdots.$$

For $i \leq i'$, condition († † †) implies that

$$c_i \cdot \tau_{i'} \in (M_i)_{n_{i'},\alpha_{i'}}.$$

For all $q \geq 1$, applying this repeatedly with $i' = q+1$ we see that for all $1 \leq q' \leq q$ we have

$$c_{q'} \cdot \tau_{q+1} \in (M_{q'})_{n_{q+1},\alpha_{q+1}} \subset (M_q)_{n_{q+1},\alpha_{q+1}}.$$

Defining I_q to be the ideal of $\boldsymbol{k}$ generated by $\{c_1, \dots, c_q\}$, this implies that for all $d \in I_q$ we have

$$d \cdot \tau_{q+1} \in (M_q)_{n_{q+1}, \alpha_{q+1}}.$$

Since $\boldsymbol{k}$ is noetherian, we can pick $q \gg 0$ such that $I_q = I_{q+1}$; in particular, $c_{q+1} \in I_q$. But this implies that

$$c_{q+1} \cdot \tau_{q+1} \in (M_q)_{n_{q+1}, \alpha_{q+1}},$$

contradicting (5.9).

5D. ***A converse to Theorem 1.5.*** We now prove a converse to Theorem 1.5:

Proposition 5.10. *Let R be a ring and $\boldsymbol{k}$ be a commutative ring such that the category of* $\mathrm{OVI}(R)$*-modules over $\boldsymbol{k}$ is locally noetherian. Then $\boldsymbol{k}$ is noetherian and the additive group of R is finitely generated.*

Proof. Let P be the principal projective $\mathrm{OVI}(R)$-module associated to R^2 and let P^+ be the submodule of P generated by all elements lying in P_n with $n > 2$. Then P/P^+ is a finitely generated $\mathrm{OVI}(R)$-module with

$$(P/P^+)_n = \begin{cases} \boldsymbol{k}[U_2(R)] & \text{if } n = 2, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that an $\mathrm{OVI}(R)$-submodule of P/P^+ is exactly the same thing as a left ideal in $\boldsymbol{k}[U_2(R)]$, so $\boldsymbol{k}[U_2(R)]$ is a left-Noetherian ring. The group $U_2(R)$ is simply the additive group underlying R, so the proposition follows from the following lemma. □

Lemma 5.11. *Let $\boldsymbol{k}$ be a commutative ring and let A be an abelian group such that $\boldsymbol{k}[A]$ is noetherian. Then $\boldsymbol{k}$ is noetherian and A is finitely generated.*

Proof. Since $\boldsymbol{k}$ is a quotient of the noetherian ring $\boldsymbol{k}[A]$ via the augmentation homomorphism, it is noetherian. For a subgroup B of A, let I_B be the ideal of $\boldsymbol{k}[A]$ generated by $[b] - [0]$ with $b \in B$. Then $\boldsymbol{k}[A]/I_B = \boldsymbol{k}[A/B]$, and so B can be recovered from I_B as the elements $b \in A$ such that $[b] - [0] \in I_B$. Suppose that $B_\bullet$ is an ascending chain of subgroups of A. Then $I_{B_\bullet}$ is an ascending chain of ideals in $\boldsymbol{k}[A]$ and thus stabilizes. Thus the chain $B_\bullet$ stabilizes as well, and so A is noetherian (and thus finitely generated) as an abelian group. □

6. Homology of OVI-modules

In this section, R denotes a (not necessarily commutative) ring whose additive group is a finitely generated abelian group and $\boldsymbol{k}$ denotes a commutative noetherian ring. Our goal is to prove Theorem 1.4 from the introduction, which says that if M is a finitely generated OVI-module then $\boldsymbol{H}_i(\boldsymbol{U}, M)$ is a finitely generated OI-module for all $i \geq 0$. This theorem is proved in Section 6C below after some preliminaries. We then prove in Section 6D an analog of Theorem 1.4 where we allow upper triangular matrices that are not necessarily unipotent.

6A. ***Homology of some*** **OI**-***groups.*** Recall that a group Γ is of type FP over $\boldsymbol{k}$ if the trivial $\boldsymbol{k}[\Gamma]$-module $\boldsymbol{k}$ admits a projective resolution $\boldsymbol{P}_\bullet$ such that each $\boldsymbol{P}_i$ is a finitely generated $\boldsymbol{k}[\Gamma]$-module. In fact, it is equivalent to ask that each $\boldsymbol{P}_i$ be a finitely generated free module; see [Brown 1982, Theorem VIII.4.3]. Many natural classes of groups are of type FP including finite groups, finitely generated abelian groups, and lattices in semisimple Lie groups. See [Brown 1982, Chapter VIII] for more information.

Proposition 6.1. *Let A be a group of type* FP *over $\boldsymbol{k}$ and let $\boldsymbol{E}$ be the* OI*-group $[n] \mapsto A^n$. Let M be an $\boldsymbol{E}$-module which is finitely generated as an* OI*-module. The following then hold*:

(a) *The* OI*-module $\boldsymbol{H}_i(\boldsymbol{E}, M)$ is finitely generated for all $i \geq 0$.*

(b) *Suppose A is abelian. Let $C \subset A$ be a finite index subgroup, A^n_C denote the subgroup $\{(a_1, \dots, a_n) \in A^n \mid a_1 + \cdots + a_n \in C\}$, and $\boldsymbol{E}_C$ be the* OI*-group $[n] \mapsto A^n_C$. Then the* OI*-module $\boldsymbol{H}_i(\boldsymbol{E}_C, M)$ is finitely generated for all $i \geq 0$.*

Proof. Pick a free resolution $\mathbb{F}_\bullet$ of the $\boldsymbol{k}[A]$-module $\boldsymbol{k}$ such that each $\mathbb{F}_i$ is a finitely generated $\boldsymbol{k}[A]$-module and such that $\mathbb{F}_0 = \boldsymbol{k}[A]$. For each $n \geq 0$, the complex $(\mathbb{F}^{\otimes n})_\bullet$ is a free resolution of the $\boldsymbol{k}[A^n]$-module $\boldsymbol{k}$.

For each $i \geq 0$, we assemble the i-th terms of $(\mathbb{F}^{\otimes n})_\bullet$ into an OI-module $X(i)$ as follows. First, define

$$X(i)_n = (\mathbb{F}^{\otimes n})_i = \bigoplus_{i_1 + \cdots + i_n = i} \mathbb{F}_{i_1} \otimes \cdots \otimes \mathbb{F}_{i_n}.$$

Next, given an OI-morphism $f\colon [n] \to [m]$, define $f_*\colon X(i)_n \to X(i)_m$ in the following way. Consider a summand $\mathbb{F}_{i_1} \otimes \cdots \otimes \mathbb{F}_{i_n}$ of $X(i)_n$. For $1 \leq a' \leq m$, define

$$i'_{a'} = \begin{cases} i_a & \text{if } a' = f(a) \text{ for some } a \in [n], \\ 0 & \text{otherwise.} \end{cases}$$

We thus obtain a summand $\mathbb{F}_{i'_1} \otimes \cdots \otimes \mathbb{F}_{i'_m}$ of $X(i)_m$. Define $f_*\colon X(i)_n \to X(i)_m$ to be the map that takes $\mathbb{F}_{i_1} \otimes \cdots \otimes \mathbb{F}_{i_n}$ to $\mathbb{F}_{i'_1} \otimes \cdots \otimes \mathbb{F}_{i'_m}$ by inserting terms that equal $1 \in \boldsymbol{k}[A] = \mathbb{F}_0$ into the needed places.

For each $i \geq 0$, define $Y(i)$ to be the OI-module $[n] \mapsto (X(i)_n \otimes M_n)_{A^n}$, where the subscript indicates that we are taking the A^n-coinvariants. The $Y(i)$ form a complex

$$\cdots \to Y(3) \to Y(2) \to Y(1) \to Y(0) \to 0$$

of OI-modules, and the OI-module $\boldsymbol{H}_i(\boldsymbol{E}, M)$ is the i-th homology group of this complex. By the local noetherianity of OI (Corollary 3.3), to prove that $\boldsymbol{H}_i(\boldsymbol{E}, M)$ is a finitely generated OI-module for all $i \geq 0$, it is enough to prove that each $Y(i)$ is a finitely generated OI-module, which we now do.

For each $i \geq 0$, the OI-module $X(i)$ is generated in finite degree (in fact, only terms of degree at most i are needed). Since M is finitely generated as an OI-module, it is in particular generated in finite degree, so by Corollary 3.4 the OI-module $X(i) \otimes M$ is also generated in finite degree. This implies that $Y(i)$ is also generated in finite degree. Since $\mathbb{F}_i$ is a finitely generated $\boldsymbol{k}[A]$-module for each $i \geq 0$ and M_n is a $\boldsymbol{k}[A^n]$-module that is finitely generated as a $\boldsymbol{k}$-module for each $n \geq 0$, it follows that the $\boldsymbol{k}$-module $Y(i)_n = ((\mathbb{F}^{\otimes n})_i \otimes M_n)_{A^n}$ is a finitely generated $\boldsymbol{k}$-module for all $i, n \geq 0$. Combining this with the fact that each $Y(i)$ is generated in finite degree, we deduce that the OI-module $Y(i)$ is finitely generated for all $i \geq 0$, as desired.

For the second statement, the restriction of $\mathbb{F}^{\otimes n}$ to A^n_C is still finitely generated since A^n_C is a finite index subgroup in A^n, and we can proceed as before. $\square$

Proposition 6.2. *Let A be a group of type* FP *over $\boldsymbol{k}$ and let $\boldsymbol{E}'$ be the* $\mathrm{OI}(d)$*-group given by $\boldsymbol{E}'_{n,\lambda} = A^n$. Then $\boldsymbol{H}_i(\boldsymbol{E}', \underline{\boldsymbol{k}})$ is a finitely generated* $\mathrm{OI}(d)$*-module for all $i \geq 0$.*

Proof. The $\mathrm{OI}(d)$-group $\boldsymbol{E}'$ is the pullback of the OI-group $\boldsymbol{E}$ from Proposition 6.1 through the forgetful functor $\Phi\colon \mathrm{OI}(d) \to \mathrm{OI}$. Thus $\boldsymbol{H}_i(\boldsymbol{E}', \underline{\boldsymbol{k}})$ is the pullback to $\mathrm{OI}(d)$ of the OI-module $\boldsymbol{H}_i(\boldsymbol{E}, \underline{\boldsymbol{k}})$, which is finitely generated by that proposition. The result now follows from the fact that Φ satisfies property (F), which follows easily from Proposition 3.2. $\square$

6B. *A filtration.* Our goal in this section is to prove the following result. Recall that $\overline{\Sigma}$ is the reduced shift functor on OI-modules, i.e., the cokernel of the canonical map $M \to \Sigma(M)$. Also, P_d is the principal projective OVI-module associated to the object R^d of OVI.

Proposition 6.3. *The* OI*-module $\overline{\Sigma}(\boldsymbol{H}_i(\boldsymbol{U}, P_d))$ has a filtration where the graded pieces are subquotients of* OI*-modules of the form $\boldsymbol{H}_i(\boldsymbol{U}, P_e)$ with $e < d$ or $\boldsymbol{H}_j(\boldsymbol{U}, M)$ with $j < i$ and M a finitely generated* OVI*-module.*

We begin with a number of lemmas. Recall that $\Phi\colon \mathrm{OI}(d) \to \mathrm{OI}$ and $\Psi\colon \mathrm{OVI}(d) \to \mathrm{OVI}$ are the forgetful functors. Also, $\boldsymbol{U}_d$ is the $\mathrm{OI}(d)$-group $(\boldsymbol{U}_d)_{n,\lambda} = U_{n,\lambda}$, where $U_{n,\lambda}$ is the group discussed in Section 4A. Finally, the subscript ! is used to denote the left Kan extension discussed in Section 2B.

Lemma 6.4. *Let M be an* $\mathrm{OVI}(d)$*-module. We have an isomorphism of* OI*-modules $\Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_d, M)) \cong \boldsymbol{H}_i(\boldsymbol{U}, \Psi_!(M))$.*

Proof. Recall from Proposition 4.5 that

$$\Psi_!(M)_n = \bigoplus_{\lambda} \mathrm{Ind}_{U_{n,\lambda}}^{U_n}(M_{n,\lambda}).$$

Thus, by Shapiro's lemma we have

$$\boldsymbol{H}_i(\boldsymbol{U}, \Psi_!(M))_n = \mathrm{H}_i(U_n, \Psi_!(M)_n) = \bigoplus_{\lambda} \mathrm{H}_i(U_{n,\lambda}, M_{n,\lambda}),$$

and this is exactly $\Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_d, M))$ by Proposition 3.6. This shows that $\Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_d, M))$ and $\boldsymbol{H}_i(\boldsymbol{U}, \Psi_!(M))$ agree on objects, and a moment's reflection shows that they also agree on morphisms. $\square$

Corollary 6.5. *We have $\boldsymbol{H}_i(\boldsymbol{U}, P_d) = \Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))$.*

Proof. Let $x = (R^d, \{e_i\}, \lambda) \in \mathrm{OVI}(d)$ where e_i is the standard basis and $\lambda = (1 < 2 < \cdots < d)$. Set $y = (R^d, \{e_i\}) \in \mathrm{OVI}$. Then $\Psi_!(P_x) = P_y$ by (2.10). Since x is the initial object of $\mathrm{OVI}(d)$, we have $P_x(y) = \boldsymbol{k}[\mathrm{Hom}(x, y)] = \boldsymbol{k}$ for all y, so $P_x = \underline{\boldsymbol{k}}$. We thus have $\Psi_!(\underline{\boldsymbol{k}}) = P_d$. Using the fact that P_y is just another name for P_d, the result now follows from Lemma 6.4 with $M = \underline{\boldsymbol{k}}$. $\square$

Let $\boldsymbol{U}'_d = \Sigma(\boldsymbol{U}_d)$. This is the OI($d$)-group given by $(\boldsymbol{U}'_d)_{n,\lambda} = U_{n+1,\lambda}$. The group $U_{n+1,\lambda}$ is the semidirect product $U_{n,\lambda} \ltimes R^n$, and this description is functorial. More precisely, let $\boldsymbol{E}_d$ be the OI(d)-group given by $(\boldsymbol{E}_d)_{n,\lambda} = R^n$. We then have homomorphisms of OI(d)-groups $i\colon \boldsymbol{U}_d \to \boldsymbol{U}'_d$ and $p\colon \boldsymbol{U}'_d \to \boldsymbol{U}_d$ with $pi = \mathrm{id}$ and $\ker(p) = \boldsymbol{E}_d$. We observe that $\boldsymbol{E}_d$ is in fact naturally an OVI(d)-group, and thus $\boldsymbol{H}_i(\boldsymbol{E}_d, \underline{\boldsymbol{k}})$ is naturally an OVI(d)-module. Proposition 6.2 says that $\boldsymbol{H}_i(\boldsymbol{E}_d, \underline{\boldsymbol{k}})$ is finitely generated as an OI(d)-module, so it is also finitely generated as an OVI(d)-module.

Lemma 6.6. *The* OI(d)*-module* $\overline{\Sigma}(\boldsymbol{H}_r(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))$ *admits a filtration where the graded pieces are subquotients of* $\boldsymbol{H}_i(\boldsymbol{U}_d, \boldsymbol{H}_{r-i}(\boldsymbol{E}_d, \underline{\boldsymbol{k}}))$ *with* $0 \le i \le r-1$.

Proof. The module $\overline{\Sigma}(\boldsymbol{H}_r(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))$ is the cokernel of the map

$$\boldsymbol{H}_r(\boldsymbol{U}_d, \underline{\boldsymbol{k}}) \to \boldsymbol{H}_r(\boldsymbol{U}'_d, \underline{\boldsymbol{k}})$$

induced by the homomorphism $i\colon \boldsymbol{U}_d \to \boldsymbol{U}'_d$. The result therefore follows from Proposition 2.11, taking $\boldsymbol{G} = \boldsymbol{U}'_d$ and $\boldsymbol{K} = \boldsymbol{U}_d$ and $\boldsymbol{E} = \boldsymbol{E}_d$. □

Recall that if M is an OI(d)-module, then right before Proposition 3.9 we defined an OI($d-1$)-module $\Delta(M)$.

Lemma 6.7. *We have* $\Phi_!(\Delta(\boldsymbol{H}_i(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))) = \Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_{d-1}, \underline{\boldsymbol{k}}))$.

Proof. By definition,

$$\Delta(\boldsymbol{H}_i(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))_{n,\lambda} = \boldsymbol{H}_i(\boldsymbol{U}_d, \underline{\boldsymbol{k}})_{[n]\amalg\{\infty\},\lambda\amalg\{\infty\}} = \mathrm{H}_i(U_{[n]\amalg\{\infty\},\lambda\amalg\{\infty\}}, \underline{\boldsymbol{k}}).$$

Since $\{\infty\}$ is the maximal element of $[n] \amalg \{\infty\}$, we have

$$U_{[n]\amalg\{\infty\},\lambda\amalg\{\infty\}} \cong U_{n,\lambda}.$$

Thus by Proposition 3.6 we have

$$\Phi_!(\Delta(\boldsymbol{H}_i(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))) = \bigoplus_{\lambda} \mathrm{H}_i(U_{n,\lambda}, \underline{\boldsymbol{k}}),$$

the sum taken over appropriate $d-1$ tuples λ. Again using Proposition 3.6, this is exactly $\Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_{d-1}, \underline{\boldsymbol{k}}))$. □

Proof of Proposition 6.3. We have $\boldsymbol{H}_r(\boldsymbol{U}, P_d) = \Phi_!(\boldsymbol{H}_r(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))$ by Corollary 6.5. Thus by Proposition 3.9, we have

$$\overline{\Sigma}(\boldsymbol{H}_r(\boldsymbol{U}, P_d)) = \overline{\Sigma}(\Phi_!(\boldsymbol{H}_r(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))) = \Phi_!(\overline{\Sigma}(\boldsymbol{H}_r(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))) \oplus \Phi_!(\Delta(\boldsymbol{H}_r(\boldsymbol{U}_d, \underline{\boldsymbol{k}}))).$$

By Lemma 6.7, the second term on the right is $\Phi_!(\boldsymbol{H}_r(\boldsymbol{U}_{d-1}, \underline{\boldsymbol{k}}))$. By Corollary 6.5, this equals $\boldsymbol{H}_r(\boldsymbol{U}, P_{d-1})$. By Lemma 6.6, the first term admits a filtration where the graded pieces are subquotients of $\Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_d, \boldsymbol{H}_{r-i}(\boldsymbol{E}_d, \underline{\boldsymbol{k}})))$ with $0 \le i \le r-1$. Setting $N_i = \boldsymbol{H}_{r-i}(\boldsymbol{E}_d, \underline{\boldsymbol{k}})$, Proposition 6.2 implies that

N_i is a finitely generated OVI(d)-module. Set $M_i = \Psi_!(N_i)$, so M_i is a finitely generated OVI-module. By Lemma 6.4, we have

$$\Phi_!(\boldsymbol{H}_i(\boldsymbol{U}_d, N_i)) = \boldsymbol{H}_i(\boldsymbol{U}, M_i).$$

Combining all of the above, $\overline{\Sigma}(\boldsymbol{H}_r(\boldsymbol{U}, P_d))$ admits a filtration where one graded piece is $\boldsymbol{H}_r(\boldsymbol{U}, P_{d-1})$ and the other graded pieces are subquotients of $\boldsymbol{H}_i(\boldsymbol{U}, M_i)$ for $0 \leq i \leq r-1$. The result follows. $\square$

6C. ***Proof of Theorem 1.4.*** We now prove Theorem 1.4. Recall the statement: if R is a ring whose additive group is a finitely generated abelian group, $\boldsymbol{k}$ is a commutative noetherian ring, and M is a finitely generated OVI-module, then $\boldsymbol{H}_i(\boldsymbol{U}, M)$ is a finitely generated OI-module for all $i \geq 0$. Fix such $\boldsymbol{k}$ and R for the rest of this section. Consider the following statement:

(S_i) For a finitely generated OVI-module M, the OI-module $\boldsymbol{H}_i(\boldsymbol{U}, M)$ is finitely generated.

Let i be given and suppose that (S_j) is true for all $j < i$ (a vacuous condition if $i = 0$). We will prove (S_i), and this will establish the theorem.

We first show by induction on d that $\boldsymbol{H}_i(\boldsymbol{U}, P_d)$ is a finitely generated OI-module for all d. Suppose therefore that $\boldsymbol{H}_i(\boldsymbol{U}, P_e)$ is a finitely generated OI-module for $e < d$ (a vacuous condition for $d = 0$), and let us prove that $\boldsymbol{H}_i(\boldsymbol{U}, P_d)$ is a finitely generated OI-module. By Proposition 6.3, the OI-module $\overline{\Sigma}(\boldsymbol{H}_i(\boldsymbol{U}, P_d))$ has a filtration where each graded piece is a subquotient of an OI-module of the form $\boldsymbol{H}_i(\boldsymbol{U}, P_e)$ with $e < d$ or $\boldsymbol{H}_j(\boldsymbol{U}, M)$ with $j < i$ and M finitely generated. By the two inductive hypotheses in force, both of these kinds of OI-modules are finitely generated. Using the local noetherianity of OI-modules (Corollary 3.3), it follows that $\overline{\Sigma}(\boldsymbol{H}_i(\boldsymbol{U}, P_d))$ is a finitely generated OI-module. By Proposition 3.8, this implies that the OI-module $\boldsymbol{H}_i(\boldsymbol{U}, P_d)$ is finitely generated, as desired.

Let M be a finitely generated OVI-module. Consider an exact sequence

$$0 \to K \to P \to M \to 0$$

where P is a finite direct sum of principal projective OVI-modules. Since the category of OVI-modules is locally noetherian (Theorem 1.5), the OVI-module K is finitely generated. We obtain an exact sequence

$$\boldsymbol{H}_i(\boldsymbol{U}, P) \to \boldsymbol{H}_i(\boldsymbol{U}, M) \to \boldsymbol{H}_{i-1}(\boldsymbol{U}, K).$$

By the previous paragraph, the OI-module $\boldsymbol{H}_i(\boldsymbol{U}, P)$ is finitely generated. By our inductive hypothesis (S_{i-1}), the OI-module $\boldsymbol{H}_{i-1}(\boldsymbol{U}, K)$ is finitely generated. Using the local noetherianity of OI (Corollary 3.3), it follows that the OI-module $\boldsymbol{H}_i(\boldsymbol{U}, M)$ is finitely generated. We have thus established (S_i), and the proof is complete.

Remark 6.8. The dimension shifting step in the third paragraph above is the only place in the proof of the theorem where the noetherianity of OVI is used. We never need noetherianity of OVI(d).

Remark 6.9. Suppose the additive group of R is a finite rank free abelian group. We outline an alternative way to get finite generation of the OI-module $[n] \mapsto \mathrm{H}_i(U_n(R); \boldsymbol{k})$. Let $\mathfrak{u}_n(R)$ be the Lie algebra of strictly upper-triangular $n \times n$ matrices over R. By [Grünenfelder 1979, Theorem 4.3], there is a spectral sequence beginning with the Lie algebra homology of $\mathfrak{u}_n(R)$ which converges to $\mathrm{H}_i(U_n(R); \boldsymbol{k})$. The Lie algebra homology of $\mathfrak{u}_n(R)$ can be computed from the Koszul complex, whose terms are exterior powers of $\mathfrak{u}_n(R)$, and hence are finitely generated OI-modules (this is similar to the OI-structure on $\mathbb{F}^{\otimes n}$ in the proof of Proposition 6.1). By noetherianity, $\mathrm{H}_i(U_n(R); \boldsymbol{k})$ is a finitely generated OI-module.

6D. ***A variant: Relaxing unipotence.*** For each n, we let $B_n(R)$ denote the group of upper-triangular invertible $n \times n$ matrices with entries in R. We denote the OI-group $[n] \mapsto B_n(R)$ by $\boldsymbol{B}$. Also, if R is commutative and $C \subset R^\times$ is a subgroup, then let $B_n^C(R) \subset B_n(R)$ be the subgroup whose determinant lies in C. We denote the OI-subgroup $[n] \mapsto B_n^C(R)$ by $\boldsymbol{B}^C$.

The goal of this section is to prove Theorem 6.11 below, which is an analog of Theorem 1.4 for $\boldsymbol{B}^C$. This requires the following lemma:

Lemma 6.10. *If R is commutative and the additive group of R is finitely generated, then the group of units $R^\times$ is also finitely generated.*

Proof. If R is a domain, then it is either a subring of the ring of integers of a number field, in which case the statement follows from the Dirichlet unit theorem, or it is a finite field, in which case there is nothing to prove.

If R is reduced, then we have an injection $R \to \prod_P R/P$ where the product is over the finitely many associated primes of R. Thus we have an injection $R^\times \to \prod_P (R/P)^\times$, and hence $R^\times$ is finitely generated.

Finally, in general we have an exact sequence of groups

$$0 \to \mathfrak{N}(R) \to R^\times \to (R/\mathfrak{N}(R))^\times \to 0,$$

where $\mathfrak{N}(R)$ is the nilradical of R equipped with the group structure $x * y = x + y + xy$, and the first map takes x to $1 + x$. (We note that the right map is surjective since any lift of a unit in $R/\mathfrak{N}(R)$ to R is automatically a unit.) By the previous cases, the abelian group $(R/\mathfrak{N}(R))^\times$ is finitely generated. The fact that the additive group of R is finitely generated implies that R is noetherian, so $\mathfrak{N}(R)^n = 0$ for some n. For each k, the $*$ operation on $\mathfrak{N}(R)$ descends to ordinary addition on $\mathfrak{N}(R)^k/\mathfrak{N}(R)^{k+1}$. Since the additive group $\mathfrak{N}(R)^k/\mathfrak{N}(R)^{k+1}$ is a subquotient of the finitely generated additive group of R, the additive group $\mathfrak{N}(R)^k/\mathfrak{N}(R)^{k+1}$ is finitely generated. Lifting additive generators for $\mathfrak{N}(R)/\mathfrak{N}(R)^2, \mathfrak{N}(R)^2/\mathfrak{N}(R)^3, \ldots, \mathfrak{N}(R)^{n-1}/\mathfrak{N}(R)^n = \mathfrak{N}(R)^{n-1}$ to $\mathfrak{N}(R)$ gives generators for $\mathfrak{N}(R)$ with respect to the operation $*$. We conclude that $R^\times$ is a finitely generated group. □

Theorem 6.11. *Suppose that R is commutative and $C \subset R^\times$ is a subgroup. If M is a $\boldsymbol{B}$-module which is finitely generated as an* OI*-module, then $\boldsymbol{H}_i(\boldsymbol{B}^C, M)$ is a finitely generated* OI*-module for any $i \geq 0$.*

Proof. Let $(R^\times)^n_C$ denote the subgroup of $(R^\times)^n$ consisting of sequences whose product lies in C. We have a short exact sequence of groups

$$1 \to U_n(R) \to B_n(R) \to (R^\times)^n_C \to 1.$$

The group $R^\times$ is finitely generated by Lemma 6.10, and thus so is $(R^\times)^n_C$. The corollary now follows from the Hochschild–Serre spectral sequence together with Theorem 1.4 and Proposition 6.1. □

7. Application to Iwahori subgroups

The goal of this section is to prove Theorem 1.8, whose statement we now recall. Let $\mathcal{O}$ be a number ring, let $\mathfrak{a} \subset \mathcal{O}$ be a nonzero proper ideal, and let $\boldsymbol{k}$ be a commutative noetherian ring. For $i \geq 0$, let $X(i)$ be the OI-module defined by the rule $[n] \mapsto \mathrm{H}_i(\mathrm{GL}_{n,0}(\mathcal{O}, \mathfrak{a}), \boldsymbol{k})$. We must prove that $X(i)$ is a finitely generated OI-module and that if $\boldsymbol{k}$ is a field then $\dim X(i)_n$ equals a polynomial in n for $n \gg 0$. The polynomiality assertion follows from the finite generation assertion together with Proposition 3.5, so we must only prove that each $X(i)$ is a finitely generated OI-module.

Define $R = \mathcal{O}/\mathfrak{a}$ and let $C \subset R^\times$ be the image of $\mathcal{O}^\times$ under the quotient map $\mathcal{O} \to R$. Let $\mathrm{GL}_n^C(R)$ be the subgroup of $\mathrm{GL}_n(R)$ consisting of matrices whose determinant lies in C. Strong approximation (see, e.g., [Platonov and Rapinchuk 1994, Chapter 7]) implies that the map $\mathrm{SL}_n(\mathcal{O}) \to \mathrm{SL}_n(R)$ is surjective. This implies that the map $\mathrm{GL}_n(\mathcal{O}) \to \mathrm{GL}_n^C(R)$ is surjective, which implies that the map $\mathrm{GL}_{n,0}(\mathcal{O}, \alpha) \to B_n^C(R)$ is surjective.

We thus have a short exact sequence

$$1 \to \mathrm{GL}_n(\mathcal{O}, \mathfrak{a}) \to \mathrm{GL}_{n,0}(\mathcal{O}, \mathfrak{a}) \to B_n^C(R) \to 1.$$

The associated Hochschild–Serre spectral sequence is of the form

$$\mathrm{H}_i(B_n^C(R), \mathrm{H}_j(\mathrm{GL}_n(\mathcal{O}, \mathfrak{a}), \boldsymbol{k})) \Rightarrow \mathrm{H}_{i+j}(\mathrm{GL}_{n,0}(\mathcal{O}, \mathfrak{a}), \boldsymbol{k}) = X(i+j)_n.$$

Let $M(j)$ be the OVI(R)-module defined by $M(j)_n = \mathrm{H}_j(\mathrm{GL}_n(\mathcal{O}, \mathfrak{a}), \boldsymbol{k})$. Naturality of the above spectral sequence induces a spectral sequence

$$\mathrm{H}_i(B_n^C, M(j)) \Rightarrow X(i+j) \tag{7.1}$$

of OI-modules.

Letting FI be the category of finite sets and injections, the rule defining $M(j)$ also endows it with an FI-module structure, which is finitely generated by [Church et al. 2014, Theorem D]. The inclusion $\mathrm{OI} \to \mathrm{FI}$ satisfies property (F) (see [Sam and Snowden 2017, Theorem 7.1.4]), so by Proposition 2.3 the induced OI-module structure on $M(j)$ is also finitely generated. This implies in particular that $M(j)$ is a finitely generated OVI(R)-module. Theorem 6.11 now implies that $\mathrm{H}_i(B^C(R), M(j))$ is a finitely generated OI-module. Since the category of OI-modules is locally noetherian (see Corollary 3.3), we can now deduce from (7.1) that each $X(i)$ is a finitely generated OI-module, as desired.

Acknowledgments. We thank Benjamin Steinberg for pointing out a significant simplification to the proof of Lemma 5.4.

References

[Aschenbrenner and Hillar 2007] M. Aschenbrenner and C. J. Hillar, "Finite generation of symmetric ideals", *Trans. Amer. Math. Soc.* **359**:11 (2007), 5171–5192. Erratum in **361**:10 (2009), 5627. MR Zbl

[Borel 1969] A. Borel, *Linear algebraic groups*, Benjamin, 1969. MR Zbl

[Borel 1974] A. Borel, "Stable real cohomology of arithmetic groups", *Ann. Sci. École Norm. Sup.* (4) **7** (1974), 235–272. MR Zbl

[Brown 1982] K. S. Brown, *Cohomology of groups*, Grad. Texts Math. **87**, Springer, 1982. MR Zbl

[Church et al. 2014] T. Church, J. S. Ellenberg, B. Farb, and R. Nagpal, "FI-modules over Noetherian rings", *Geom. Topol.* **18**:5 (2014), 2951–2984. MR Zbl

[Church et al. 2015] T. Church, J. S. Ellenberg, and B. Farb, "FI-modules and stability for representations of symmetric groups", *Duke Math. J.* **164**:9 (2015), 1833–1910. MR Zbl

[Cohen 1967] D. E. Cohen, "On the laws of a metabelian variety", *J. Algebra* **5** (1967), 267–273. MR Zbl

[Dwyer 1985] W. G. Dwyer, "Homology of integral upper-triangular matrices", *Proc. Amer. Math. Soc.* **94**:3 (1985), 523–528. MR Zbl

[Grünenfelder 1979] L. Grünenfelder, "On the homology of filtered and graded rings", *J. Pure Appl. Algebra* **14**:1 (1979), 21–37. MR Zbl

[Hall 1954] P. Hall, "Finiteness conditions for soluble groups", *Proc. London Math. Soc.* (3) **4** (1954), 419–436. MR Zbl

[Higman 1952] G. Higman, "Ordering by divisibility in abstract algebras", *Proc. London Math. Soc.* (3) **2** (1952), 326–336. MR Zbl

[Hillar and Sullivant 2012] C. J. Hillar and S. Sullivant, "Finite Gröbner bases in infinite dimensional polynomial rings and applications", *Adv. Math.* **229**:1 (2012), 1–25. MR Zbl

[van der Kallen 1980] W. van der Kallen, "Homology stability for linear groups", *Inv. Math.* **60**:3 (1980), 269–295. MR Zbl

[Kruskal 1972] J. B. Kruskal, "The theory of well-quasi-ordering: a frequently discovered concept", *J. Combinatorial Theory Ser. A* **13** (1972), 297–305. MR Zbl

[Lee and Szczarba 1976] R. Lee and R. H. Szczarba, "On the homology and cohomology of congruence subgroups", *Inv. Math.* **33**:1 (1976), 15–53. MR Zbl

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

[Putman 2015] A. Putman, "Stability in the homology of congruence subgroups", *Inv. Math.* **202**:3 (2015), 987–1027. MR Zbl

[Putman and Sam 2017] A. Putman and S. V Sam, "Representation stability and finite linear groups", *Duke Math. J.* **166**:13 (2017), 2521–2598. MR Zbl

[Quillen 1972] D. Quillen, "On the cohomology and K-theory of the general linear groups over a finite field", *Ann. of Math.* (2) **96** (1972), 552–586. MR Zbl

[Sam and Snowden 2017] S. V Sam and A. Snowden, "Gröbner methods for representations of combinatorial categories", *J. Amer. Math. Soc.* **30**:1 (2017), 159–203. MR Zbl

[Stanley 2012] R. P. Stanley, *Enumerative combinatorics, I*, 2nd ed., Cambridge Stud. Adv. Math. **49**, Cambridge Univ. Press, 2012. MR Zbl

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Stud. Adv. Math. **38**, Cambridge Univ. Press, 1994. MR Zbl

Communicated by Victor Reiner
Received 2018-12-19 Accepted 2019-08-18

andyp@nd.edu *Department of Mathematics, University of Notre Dame, Notre Dame, IN, United States*

svs@math.wisc.edu *Mathematics Department, University of California, San Diego, La Jolla, CA, United States*

asnowden@umich.edu *Department of Mathematics, University of Michigan, Ann Arbor, MI, United States*

ALGEBRA AND NUMBER THEORY 14:1 (2020)
dx.doi.org/10.2140/ant.2020.14.155

On the orbits of multiplicative pairs

Oleksiy Klurman and Alexander P. Mangerel

Dedicated to Imre Kátai on the occasion of his 80th birthday

We characterize all pairs of completely multiplicative functions $fg : \mathbb{N} \to \mathbb{T}$, where $\mathbb{T}$ denotes the unit circle, such that

$$\overline{\{(f(n), g(n+1))\}_{n\geq 1}} \neq \mathbb{T} \times \mathbb{T}.$$

In so doing, we settle an old conjecture of Zoltán Daróczy and Imre Kátai.

1. Introduction

In this paper, we will be concerned with demonstrating yet another instance of the expected general phenomenon that the multiplicative structure of positive integers should in general be "independent" of their additive structure. Of principal focus here will be the behavior of multiplicative functions at consecutive integers. Problems of this kind are widely open in general, though spectacular progress has recently been made as a consequence of the breakthrough of Matomäki and Radziwiłł [2016], and subsequent work of Matomäki, Radziwiłł and Tao [Matomäki et al. 2015], Tao [2016] and, more recently, Tao and Teräväinen [2019]. In particular, using the work in [Matomäki and Radziwiłł 2016], Tao [2016] established a weighted version of the binary Chowla conjecture in the form

$$\sum_{n\leq x} \frac{\lambda(n)\lambda(n+h)}{n} = o(\log x)$$

for all $h \geq 1$. For a comprehensive account of the recent developments in this direction, see [Matomäki and Radziwiłł 2019]. Let $\mathbb{U}$ denote the unit disc in $\mathbb{C}$ and let $\mathbb{T}$ denote the unit circle. Let $f, g : \mathbb{N} \to \mathbb{T}$ be completely multiplicative functions. We expect that as n varies through the set of positive integers, the values $f(n)$ and $g(n+1)$ should, roughly speaking, be independently distributed unless f and g satisfy some rigid relations. To be more precise, we shall investigate the following problem: if $\{f(n)\}_n$ and $\{g(n)\}_n$ are both dense in $\mathbb{T}$, but the sequence of pairs $\{(f(n), g(n+1))\}_n$ is *not* dense in $\mathbb{T}^2$, must there be a rigid relation between f and g? This multidimensional problem continues work on rigidity problems for additive and multiplicative functions initiated by the authors in [Klurman and Mangerel 2018]. This problem has a natural dynamical flavor, which explains the title of this paper. Let $T : \mathbb{N} \to \mathbb{N}$ denote the rightward shift map $T(n) := n+1$. In this case, the above problem can be recast in terms of orbits of the

MSC2010: primary 11N37; secondary 11N64.
Keywords: multiplicative functions, Erdos discrepancy problem, Katai conjecture.

pair (f, g_T), where $f, g : \mathbb{N} \to \mathbb{T}$ are semigroup homomorphisms that fix $n = 1$, and $g_T := g \circ T$. We seek a result of the kind that, unless f and g are specially chosen maps, the orbit closure of the point 1, i.e., the closure $\overline{\{(f(n), g(n+1))\}_n}$, is expected to be the same as the product of the closures of the marginal orbits $\{f \circ T^n\}_n$ and $\{g \circ T^n\}_n$. In this connection, we quote Conjecture 3 in the survey paper by Kátai [1989] (earlier formulated in [Daróczy and Kátai 1989]).

Conjecture 1.1. *Let $f, g : \mathbb{N} \to \mathbb{T}$ be completely multiplicative. Suppose $\{(f(n), g(n+1))\}_n$ is not dense in $\mathbb{T}^2$, yet $\{f(n)\}_n$ and $\{g(n)\}_n$ are both dense in $\mathbb{T}$. Then there are integers k and l such that $f(n)^k = g(n)^l$, with $f(n) = n^{it}$ for some t.*

As stated this conjecture is easily seen to be false, as we can construct the following two types of counterexamples:

(i) Let $h_1, h_2 : \mathbb{N} \to \mathbb{T}$ be completely multiplicative functions such that there are minimal positive integers $k, l \geq 2$ for which $h_1^k = h_2^l = 1$. Fixing an arbitrary $t \in \mathbb{R}\backslash\{0\}$ and setting $f(n) := h_1(n)n^{it}$ and $g(n) := h_2(n)n^{it}$ yields a pair of completely multiplicative functions such that $\{f(n)\}_n$ and $\{g(n)\}_n$ are dense, yet $\{(f(n), g(n+1))\}_n$ cannot be dense. On the other hand, it is true in this example that $f(n)^k = n^{it'}$ and $g(n)^l = n^{it''}$ for all n, where $t' = kt$ and $t'' = lt$.

(ii) Fix a prime p and distinct irrational numbers $\alpha, \beta \in \mathbb{R}$. Let $f, g : \mathbb{N} \to \mathbb{T}$ be the completely multiplicative functions defined on primes by

$$f(q) := \begin{cases} e(\alpha) & q = p, \\ 1 & q \neq p, \end{cases} \quad \text{and} \quad g(q) := \begin{cases} e(\beta) & q = p, \\ 1 & q \neq p. \end{cases}$$

It is easy to see that the sequence $\{(f(n), g(n+1))\}_n$ belongs to the union of the sets $\mathbb{T} \times \{1\}$, $\{1\} \times \mathbb{T}$ and $\{1\} \times \{1\}$ and thus cannot be dense. On the other hand we clearly have $\overline{\{f(p^m)\}_{m\geq 1}} = \overline{\{g(p^m)\}_{m\geq 1}} = \mathbb{T}$.

Given a completely multiplicative function h, let

$$T_h := \{p \text{ prime} : h(p) \neq 1\}.$$

The collection (i) of counterexamples suggests that we should relax the conclusion of Conjecture 1.1 by allowing $f(n)^k = n^{it}$ for some $k \geq 1$. The collection (ii) of counterexamples indicates that we should add a hypothesis to exclude those functions f and g such that both $|T_{f^k}||T_{g^k}| = 1$ and $T_{f^k} = T_{g^k}$ hold for all sufficiently large k.[1] We thus prove the following amendment of Conjecture 1.1, in which the above types of counterexamples are excluded.

Theorem 1.2 (amended Daróczy–Kátai conjecture). *Let $f, g : \mathbb{N} \to \mathbb{T}$ be completely multiplicative. Suppose $\overline{\{(f(n), g(n+1))\}_n} \neq \mathbb{T}^2$, yet $\overline{\{f(n)\}_n} = \overline{\{g(n)\}_n} = \mathbb{T}$. Suppose additionally that for infinitely many m we have $|T_{f^m} \cup T_{g^m}| > 1$. Then there are integers k and l such that $f(n)^k = g(n)^l$, with $f(n)^k = n^{it}$ for some nonzero real number t.*

[1] We note that the condition that $\overline{\{f(n)\}_n} = \mathbb{T}$ implies that $|T_{f^k}| > 0$ for all $k \in \mathbb{N}$.

In the case that $f = g$, we have the following immediate corollary.

Corollary 1.3. *Let $f : \mathbb{N} \to \mathbb{T}$ be a completely multiplicative function such that $\overline{\{f(n)\}_n} = \mathbb{T}$. Then $\overline{\{(f(n), f(n+1))\}_n} \neq \mathbb{T}^2$ if, and only if, one of the following conditions holds*:

(a) *There is a positive integer l such that $|T_{f^l}| = 1$, and for $p \in T_{f^l}$, $f(p) = e(\alpha)$ with $\alpha \notin \mathbb{Q}$.*

(b) *There is a positive integer k and a nonzero real number t for which $f(n)^k = n^{it}$.*

Proof. It is easy to see that if $f : \mathbb{N} \to \mathbb{T}$ is a completely multiplicative function satisfying either of conditions (a) or (b) then $\{f(n)\}_n$ is dense in $\mathbb{T}$ but $\{(f(n), f(n+1))\}_n$ is not dense in $\mathbb{T}^2$. We now assume that $\overline{\{f(n)\}_n} = \mathbb{T}$ but $\overline{\{(f(n), f(n+1))\}_n} \neq \mathbb{T}^2$. If $|T_{f^m}| > 1$ for infinitely many m then by Theorem 1.2 there is a $k \in \mathbb{Z}$ and a $t \neq 0 \in \mathbb{R}$ for which $f(n)^k = n^{it}$ for all $n \in \mathbb{N}$, which fulfills condition (b). Conversely, if $|T_{f^m}| = 1$ for all but finitely many m then there is some $M \in \mathbb{N}$ such that for all $m \geq M$, there is exactly one prime p with $f(p)^m \neq 1$. Thus, for all $p' \neq p$, $f(p') = e(a/b)$ for some $1 \leq a, b \leq M$. Putting $l := M!$, it follows that $f(p')^l = 1$, except when $p' = p$. Moreover, $f(p)^l$, and thus also $f(p)$, must have irrational argument, otherwise we could find a larger M' for which $f(p')^{M!M'} = 1$ for all p', contradicting the denseness of $\{f(n)\}_n$ in $\mathbb{T}$; such a function satisfies condition (a). □

In order to facilitate our discussion, we distinguish completely multiplicative functions according to their values on primes as follows.

Definition 1.4. A completely multiplicative function $f : \mathbb{N} \to \mathbb{T}$ is said to be *eventually rational* if there exist positive integers k and $N_0 = N_0(k)$ such that for all $p \geq N_0$ we have $f(p)^k = 1$. We will say that f is *irrational* otherwise.

An irrational function necessarily produces a sequence $\{f(n)\}_n$ that is dense. We stress, though, that the arguments of an irrational function at primes *need not be* irrational. For example, by our definition, the completely multiplicative function defined by $f(p) = e(1/p)$ for all primes p is irrational. To prove Theorem 1.2, we will treat two cases, depending on whether or not f or g is irrational (in the sense of Definition 1.4). In Section 4 we prove the following.

Proposition 1.5. *Suppose $f, g : \mathbb{N} \to \mathbb{T}$ are eventually rational, $\overline{\{f(n)\}_n} = \overline{\{g(n)\}_n} = \mathbb{T}$ and $|T_{f^k} \cup T_{g^k}| > 1$ for infinitely many k. Then $\overline{\{(f(n), g(n+1))\}_n} = \mathbb{T}^2$.*

Proposition 1.5 asserts that the only cases satisfying the assumptions of Theorem 1.2 are those for which either f or g is irrational. In this direction we prove the following in Section 6.

Theorem 1.6. *Suppose that $f, g : \mathbb{N} \to \mathbb{T}$ are completely multiplicative functions such that $\overline{\{f(n)\}_n} = \overline{\{g(n)\}_n} = \mathbb{T}$. Suppose furthermore that at least one of f and g is irrational. If $\overline{\{(f(n), g(n+1))\}_n} \neq \mathbb{T}^2$ then there are positive integers k, l and real numbers t, t' such that $f(n)^k = n^{it}$ and $g(n)^l = n^{it'}$.*

Having shown Theorem 1.6, in order to prove Theorem 1.2 it remains to prove that the real numbers t and t' satisfy $t = \lambda t'$, for $\lambda \in \mathbb{Q}$. This is the conclusion of Proposition 1.7, which we prove in Section 3.

Proposition 1.7. *Suppose there are $k, l \in \mathbb{N}$ and $t, t' \in \mathbb{R}$ with $tt' \neq 0$ such that $f(n)^k = n^{it}$ and $g(n)^l = n^{it'}$ for all $n \in \mathbb{N}$. Then $\overline{\{(f(n), g(n+1))\}} \neq \mathbb{T}^2$ if, and only if, $t = \lambda t'$ with $\lambda \in \mathbb{Q}$.*

The proof of Theorem 1.6, which represents the major substance of our analysis, uses arguments that extend those found in the proof of Theorems 1.1 and 1.4 in [Klurman and Mangerel 2018]. Let us recall the rough outline of this argument here. In Theorem 1.4 of [loc. cit.] (the proof of which establishes Theorem 1.1 there as well), it was shown that the sequence $\{f(n)\overline{f(n+1)}\}_n$ is dense in $\mathbb{T}$, except in predictable cases.[2] The objective was to show that for every $\varepsilon > 0$ and every $z \in \mathbb{T}$, the lower bound

$$|f(n)\overline{f(n+1)} - z| \geq \varepsilon \text{ for all } n \text{ sufficiently large} \tag{1}$$

cannot hold for "generic" completely multiplicative functions $f : \mathbb{N} \to \mathbb{T}$. To establish this, we noted that (1) implies that the sequence $\{f(n)\overline{f(n+1)}\}_n$ cannot be equidistributed, which led us (via the Erdős–Turán inequality and Tao's theorem [2016] on logarithmic averages of binary correlations; see Theorem 2.1 below) to a first conclusion that f is *pseudopretentious*, i.e., such that for some (minimal) $k \in \mathbb{N}$, some Dirichlet character χ and some $t \in \mathbb{R}$ (depending at most on ε), we have

$$\mathbb{D}(f^k, \chi n^{it}; x)^2 := \sum_{p \leq x} \frac{1 - \mathrm{Re}(f(p)^k \bar{\chi}(p) p^{-it})}{p} \ll_\varepsilon 1;$$

in particular, $\mathbb{D}(f, gn^{it}; x) \ll_\varepsilon 1$ for some completely multiplicative function g such that $g(n)^k = \chi(n)$ whenever $\chi(n) \neq 0$. Here the fact that $(n+1)^{it} \approx n^{it}$ for large n is used crucially. Roughly speaking, we then conditioned on a suitable subset of n (or positive logarithmic density) such that $g(n)\overline{g(n+1)}$ is constant, and thereby reduced our work to treating the 1-pretentious function $F = f\bar{g}n^{-it}$, showing that for small enough ε,

$$|F(n)\overline{F(n+1)} - z'| = |f(n)\overline{f(n+1)} - z| + O(\varepsilon^2) \gg \varepsilon$$

for some z' (where z and z' need not be the same) is untenable for all n sufficiently large (depending at most on ε). This reduction was a key part that made that argument work. In the context of Theorem 1.2, we must face several key differences in the argument. For example:

(i) The fact that[3] $\|(f(n), g(n+1)) - (z, w)\|_{\ell^1} \geq \varepsilon$ for all large n may mean that $|f(n) - z| \geq \varepsilon/2$ on a very dense set, or a very sparse set, and we cannot exclude either of these possibilities.

(ii) If f and g are both pseudopretentious in the above sense, say f is pretentious to $h_1 n^{it}$ and g is pretentious to $h_2 n^{it'}$, then it may be that $t \neq t'$, and we must then deal with the distribution in argument of the twist $n^{i(t-t')}$, unlike in the outline of the proof of Theorem 1.4 of [Klurman and Mangerel 2018].

With some additional ideas, we are able to address these issues. See especially Section 6 for further details.

[2] That is, except when $f(n) = g(n)n^{it}$, where g is a function taking values in bounded order roots of unity.

[3] For a pair $z = (z_1, z_2) \in \mathbb{C}^2$, we write $\|z\|_{\ell^1} := |z_1| + |z_2|$.

2. Auxiliary results towards Theorem 1.2

In this section we collect the definitions and lemmata that we shall use in the proof of Theorem 1.2. A crucial result on which our method relies is the following recent breakthrough result of Tao [2016].

Theorem 2.1 (Tao). *Let $f_1, f_2 : \mathbb{N} \to \mathbb{U}$ be multiplicative functions, such that for some $j \in \{1, 2\}$, we have*

$$\inf_{|t|\le x} \mathbb{D}(f_j, \chi n^{it}; x)^2 = \inf_{|t|\le x} \sum_{p\le x} \frac{1-\mathrm{Re}(f_j(p)\overline{\chi}(p)p^{-it})}{p} \to \infty$$

for each fixed Dirichlet character χ. Then for any integers $a_1, a_2 \ge 1$ and $b_1, b_2 \ge 0$ with $a_1b_2 - a_2b_1 \ne 0$, we have

$$\frac{1}{\log x} \sum_{n\le x} \frac{f_1(a_1n+b_1) f_2(a_2n+b_2)}{n} = o(1).$$

A useful consequence of Theorem 2.1 is the following.

Proposition 2.2. *Let $f, g : \mathbb{N} \to \mathbb{U}$ be multiplicative functions. Suppose $k, l \in \mathbb{N}$ are minimal, such that*

$$\left| \frac{1}{\log x} \sum_{n\le x} \frac{f(n)^k g(n+1)^l}{n} \right| \gg 1, \tag{2}$$

as $x \to \infty$. Then there are Dirichlet characters χ_1, χ_2 with respective conductors $q_1, q_2 = O(1)$, and real numbers $t_1, t_2 = O(1)$ such that $\mathbb{D}(f, h_1 n^{it_1}, \infty) < \infty$ and $\mathbb{D}(f, h_2 n^{it_2}, \infty) < \infty$,[4] *with $h_1(n)^k = \chi_1(n/(n, q_1^\infty))$ and $h_2(n)^l = \chi_2(n/(n, q_2^\infty))$.*

Proof. From (2) and Theorem 2.1, it follows that for each x sufficiently large there is a pair (ξ_x, t_x), where ξ_x is a primitive Dirichlet character with $\mathrm{cond}(\xi_x) \ll 1$ and $t_x \ll x$, each estimate being uniform in x, such that $\mathbb{D}(f^k, \xi_x n^{it_x}; x) \ll 1$. By Lemma 2.5 of [Klurman and Mangerel 2018], it follows that there is a character χ_1 and a $t_1 \in \mathbb{R}$ such that $\mathbb{D}(f^k, \chi_1 n^{ikt_1}; x) \ll 1$. Combining this with Lemma 2.8 of [loc. cit.], it follows that there is a completely multiplicative function h_1 such that for all primes $p \nmid \mathrm{cond}(\chi_1)$ we have $h_1(p)^k = \chi_1(p)$, while if $p \mid \mathrm{cond}(\chi_1)$, $h_1(p) = 1$ and $\mathbb{D}(f, h_1 n^{it_1}, \infty) < \infty$. This implies the claim about f. The claim about g follows in the same way. □

Presieving on level sets with Archimedean twists. Recall that a 1-bounded multiplicative function f is called *pseudopretentious* if there exists a (minimal) positive integer k, a primitive Dirichlet character χ modulo q, a real number t and a completely multiplicative function $h : \mathbb{N} \to \mathbb{T}$ such that $h(n)^k = \tilde{\chi}(n)$, and $\mathbb{D}(f, hn^{it}; x) \ll 1$. Here, $\tilde{\chi}$ is the unimodular completely multiplicative function defined on primes via $\tilde{\chi}(p) = \chi(p)$ if $p \nmid q$, and $\tilde{\chi}(p) = 1$ otherwise. Henceforth, we will refer to the function h here as a *pseudocharacter*, and refer to the modulus q of χ as the *conductor* of h. We emphasize that in this definition the minimality of k implies that $h(n)^j$ is nonpretentious for all $1 \le j \le k-1$. This will be crucial in several of the arguments below. In order to control the behavior of h_1, h_2 coming from Proposition 2.2 on the corresponding progressions, we would like to eliminate the effect of the small

[4] Given $a, b \in \mathbb{N}$, we write $(a, b^\infty) := \prod_{p \mid b} p^{\nu_p(a)}$, where $a = \prod_{p \mid a} p^{\nu_p(a)}$.

primes by using presieving following the approach from [Klurman and Mangerel 2018]. To this end, given $\alpha_j \in h_j(\mathbb{N})$ for $j = 1, 2$ and $N, B \geq 1$, write

$$\mathcal{A}_{N,B}(h_1, h_2; \alpha_1, \alpha_2) := \{n \in \mathbb{N} : P^-(n(Bn+1)) > N, h_1(n) = \alpha_1, h_2(Bn+1) = \alpha_2\},$$

where $P^-(n)$ denotes the smallest prime factor dividing $n \in \mathbb{N}$. Furthermore, if $I, J \subseteq \mathbb{T}$ are arcs[5] and $u, v \in \mathbb{R}$ then

$$\mathcal{A}_{N,B,I}(h_1, h_2; \alpha_1, \alpha_2; u) := \{n \in \mathcal{A}_{N,B}(h_1, h_2; \alpha_1, \alpha_2) : n^{iu} \in I\},$$
$$\mathcal{A}_{N,B,I,J}(h_1, h_2; \alpha_1, \alpha_2; u, v) := \{n \in \mathcal{A}_{N,B}(h_1, h_2; \alpha_1, \alpha_2) : n^{iu} \in I, n^{iv} \in J\}.$$

Moreover, if $x, q \geq 1$ and a is a coprime residue class modulo q, put

$$\Phi_{N,B}(x; q, a) := |\{n \leq x : P^-(n(Bn+1)) > N \text{ and } n \equiv a(q)\}|.$$

When $q = 1$, we write $\Phi_{N,B}(x; q, a) = \Phi_{N,B}(x)$, and when $B = 1$ in addition, we write $\Phi_N(x)$. Finally, given a map $f : \mathbb{N} \to \mathbb{C}$ and a positive real $X \geq 2$, we shall write

$$\mathbb{E}^{\log}_{n \leq X} f(n) := \frac{1}{\log X} \sum_{n \leq X} \frac{f(n)}{n}.$$

Our first result, to be used throughout the paper, shows that the sets $\mathcal{A}_{N,B,I}$ and $\mathcal{A}_{N,B,I,J}$ given above, on which the values of two discrete functions h_1, h_2 are restricted as well as archimedean characters n^{iu} and n^{iv}, can be large in the sense that they have positive upper density in general.[6] Let μ_v denotes the set of v-th roots of unity.

Proposition 2.3. *Let χ_1, χ_2 be primitive Dirichlet characters of respective conductors q_1 and q_2 and orders k and l. Let $h_1, h_2 : \mathbb{N} \to \mathbb{T}$ be pseudocharacters modulo q_1 and q_2, respectively, such that $h_1(n)^r = \tilde{\chi}_1(n)$ and $h_2(n)^s = \tilde{\chi}_2(n)$, for some $r, s \in \mathbb{N}$:*

(a) *Let $u \in \mathbb{R}$, and let $(\alpha, \beta) \in \mu_{kr} \times \mu_{ls}$. Let $\delta > 0$, $z \in \mathbb{T}$ and let $I \subset \mathbb{T}$ be an arc with length δ about* 1. *Then, for any $B \geq 1$ satisfying $2q_1q_2 \mid B$,*

$$\mathbb{E}^{\log}_{n \leq x} 1_{\mathcal{A}_{N,B,I}(h_1,h_2;\alpha,\beta;u)}(n) \gg \frac{\delta}{krls} \frac{\Phi_{N,B}(x)}{x}$$

as $x \to \infty$. Moreover, if $u \neq 0$ then we may replace I above by any arc of length δ.

(b) *If $u, v \in \mathbb{R}$ are fixed and such that $u/v \notin \mathbb{Q}$, and J_1 and J_2 are arcs in $\mathbb{T}$ of respective lengths δ_1 and δ_2, then*

$$\mathbb{E}^{\log}_{n \leq x} 1_{\mathcal{A}_{N,B,J_1,J_2}(h_1,h_2;\alpha,\beta;u,v)} \gg \frac{\delta_1\delta_2}{krls} \frac{\Phi_{N,B}(x)}{x}.$$

[5]By an *arc* in $\mathbb{T}$, we mean the image of an interval $[a, b] \subseteq \mathbb{R}$ under the exponentiation map $t \mapsto e(t)$. Thus, a symmetric arc about 1, for example, refers to the image under exponentiation of any interval $[m - \eta, m + \eta]$ with $m \in \mathbb{Z}$.

[6]To be precise, Proposition 2.3 implies directly that these sets have positive upper *logarithmic density*, that is $\limsup_{x\to\infty} \mathbb{E}^{\log}_{n \leq x} 1_{\mathcal{A}}(n) > 0$; however, this also implies that the upper density $\limsup_{x\to\infty} \frac{1}{x} \sum_{n \leq x} 1_{\mathcal{A}}(n) > 0$ as well.

To prove Proposition 2.3, we need the following variant of Lemma 2.11 from [Klurman and Mangerel 2018], which is easily proven by the fundamental lemma of the sieve.

Lemma 2.4. *Let q, $B \geq 1$, $N \geq 2$ and let a be a residue class modulo q such that $(a(Ba+1), q) = 1$. Then as $x \to \infty$,*

$$\Phi_{N,B}(x; q, a) = \frac{x}{q} \prod_{\substack{p \leq N \\ p \mid B/(B,q^\infty)}} \left(1 - \frac{1}{p}\right) \prod_{\substack{3 \leq p \leq N \\ p \nmid qB}} \left(1 - \frac{2}{p}\right) + O(4^{\pi(N)}).$$

In the sequel, we write

$$\delta_{N,B,q} := \frac{1}{q} \prod_{\substack{p \leq N \\ p \mid B/(B,q^\infty)}} \left(1 - \frac{1}{p}\right) \prod_{\substack{3 \leq p \leq N \\ p \nmid qB}} \left(1 - \frac{2}{p}\right),$$

and set $\delta_{N,B} = \delta_{N,B,1}$.

Lemma 2.5. *Assume the hypotheses of Proposition 2.3. Furthermore, suppose $N > \max\{q_1, q_2\}$, and that h_1^j and h_2^m are both nonpretentious for all $1 \leq j \leq r-1$ and $1 \leq m \leq s-1$. Then as $x \to \infty$,*

$$\sum_{n \leq x} \frac{1_{A_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n^{1+iu}} = \begin{cases} \delta_{N,B}/(iukrls) + o(\log x) + O_{q_1}(|u| 4^{\pi(N)}) & \text{if } u \neq 0, \\ 1/(krls)(\delta_{N,B} + o(1)) \log x + O_{q_1}(4^{\pi(N)}) & \text{otherwise.} \end{cases}$$

Proof. Since $\alpha \in \mu_{kr}$ and $\beta \in \mu_{ls}$, we have the identities

$$1_{h_1(n)=\alpha} = \frac{1}{kr} \sum_{0 \leq j \leq kr-1} (h_1(n)\bar{\alpha})^j \quad \text{and} \quad 1_{h_2(Bn+1)=\beta} = \frac{1}{ls} \sum_{0 \leq m \leq ls-1} (h_2(Bn+1)\bar{\beta})^j.$$

It follows that

$$\begin{aligned} \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n^{1+iu}} &= \sum_{\substack{n \leq x \\ P^-(n(Bn+1)) > N}} \frac{1_{h_1(n)=\alpha} 1_{h_2(Bn+1)=\beta}}{n^{1+iu}} \\ &= \frac{1}{klrs} \sum_{0 \leq j \leq kr-1} \sum_{0 \leq m \leq ls-1} \alpha^{-j} \beta^{-m} \sum_{\substack{n \leq x \\ P^-(n(Bn+1)) > N}} \frac{h_1(n)^j h_2(Bn+1)^m n^{-iu}}{n}. \end{aligned} \tag{3}$$

As h_1^j and h_2^m are both nonpretentious for all $1 \leq j \leq r-1$ and $1 \leq m \leq s-1$, it follows that the multiplicative functions

$$\begin{aligned} \phi_{j+ar}(n) &:= h_1(n)^{j+ar} 1_{P^-(n)>N} n^{-iu} = h_1(n)^j \tilde{\chi}_1(n)^a 1_{P^-(n)>N} n^{-iu} \\ \psi_{m+bs}(n) &:= h_2(n)^{m+bs} 1_{P^-(n)>N} = h_2(n)^m \tilde{\chi}_2(n)^b 1_{P^-(n)>N} \end{aligned}$$

are both nonpretentious for such j, m, and any $0 \le a \le k-1$, $0 \le b \le s-1$. By Theorem 2.1, it follows that

$$\Big| \sum_{\substack{0\le j\le kr-1 \\ r\nmid j \text{ or } s\nmid m}} \sum_{0\le m\le ls-1} \alpha^{-j}\beta^{-m} \sum_{\substack{n\le x \\ P^-(n(Bn+1))>N}} \frac{h_1(n)^j h_2(Bn+1)^m n^{-iu}}{n} \Big|$$

$$\le \sum_{\substack{0\le j\le kr-1 \\ r\nmid j \text{ or } s\nmid m}} \sum_{0\le m\le ls-1} \Big| \sum_{n\le x} \frac{\phi_j(n)\psi_m(Bn+1)}{n} \Big| = o(\log x), \quad (4)$$

where the estimate depends on k, r, l and s. When $j = ra$ and $m = bs$, we instead have

$$\sum_{n\le x} \frac{\phi_{ar}(n)\psi_{bs}(Bn+1)}{n} = \sum_{\substack{n\le x \\ P^-(n(Bn+1))>N}} \frac{\tilde{\chi}_1(n)^a \tilde{\chi}_2(Bn+1)^b}{n^{1+iu}} = \sum_{\substack{n\le x \\ P^-(n(Bn+1))>N}} \frac{\tilde{\chi}_1(n)^a}{n^{1+iu}} \quad (5)$$

for each b, as $q_2 \mid B$. Now, as $N > q_1$, for each n with $P^-(n(Bn+1)) > B$ we must have $(n, q_1) = 1$. Thus, splitting the rightmost sum in (5) into coprime residue classes modulo q_1, the RHS of (5) becomes

$$\sum_{c(q_1)} \chi_1(c)^a \sum_{\substack{n\le x \\ P^-(n(Bn+1))>N}} \frac{1_{n\equiv c(q_1)}}{n^{1+iu}}.$$

We first consider the case $u \neq 0$. Applying the previous lemma and partial summation, we get that for each coprime residue c modulo q_1,

$$\sum_{\substack{n\le x \\ P^-(n(Bn+1))>N}} \frac{1_{n\equiv c(q_1)}}{n^{1+iu}} = \int_1^x \frac{1}{t^{1+iu}}\, d\{\Phi_{N,B}(t; q_1, c)\}$$

$$= \delta_{N,B,q_1} \int_1^x \frac{dt}{t^{1+iu}} + O_{q_1}\Big(1 + |u| 4^{\pi(N)} \int_1^x \frac{dt}{t^2}\Big) \quad (6)$$

$$= \frac{\delta_{N,B,q_1}}{iu}(1 - x^{-iu}) + O_{q_1}(1 + |u| 4^{\pi(N)}). \quad (7)$$

The main term being independent of c modulo q_1, we can invert the orders of summation and use orthogonality to get that whenever $a \neq 0$, we have

$$\sum_{c(q_1)} \chi_1(c)^a \sum_{\substack{n\le x \\ P^-(n(Bn+1))>N}} \frac{1_{n\equiv c(q_1)}}{n^{1+iu}} = \frac{\delta_{N,B,q_1}}{iu}(1 - x^{-iu}) \sum_{c(q_1)} \chi_1(c)^a + O(q_1^2(1 + |u| 4^{\pi(N)}))$$

$$\ll q_1^2((1 + |u|) 4^{\pi(N)}).$$

As such, it follows that whenever $a \neq 0$, we have

$$\sum_{n\le x} \frac{\phi_{ar}(n)\psi_{bs}(Bn+1)}{n} \ll q_1^2((1 + |u|) 4^{\pi(N)}). \quad (8)$$

In the remaining case $a = 0$, we have

$$\sum_{\substack{n \le x \\ P^-(n(Bn+1)) > N}} \frac{1}{n^{1+iu}} = \frac{\delta_{N,B}}{iu}(1 - x^{-iu}) + O((1+|u|)4^{\pi(N)}), \tag{9}$$

as this is the LHS of (6), but with q_1 replaced by 1. Combining (4), (8) and (9), and inserting these into (3), we find that

$$\sum_{n \le x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n^{1+iu}} = \frac{1}{krls}\left(\frac{\delta_{N,B}}{iu}(1 - x^{-iu}) + O(q_1(1+|u|)4^{\pi(N)}) + o(\log x)\right)$$

when $u \ne 0$, as claimed. The case $u = 0$ follows the same lines, but is simpler. The proof is now complete. $\square$

Before proceeding to the proof of Proposition 2.3, we pause to recall the following objects of relevance to it. Given $K \in \mathbb{N}$, let $B_K : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ denote the degree K *Beurling polynomial* (see Section 1.2 of [Montgomery 1994] for a definition and some of its properties). Recall that it satisfies the properties that:

(i) If $\psi(t) := \{t\} - \frac{1}{2}$ is the sawtooth function then $-B_K(-t) \le \psi(t) \le B_K(t)$. As such, given an interval $I \subseteq \mathbb{R}/\mathbb{Z}$ with endpoints $0 \le a < b < 1$, we have functions

$$f_I(t) := |I| - B_K(b-t) - B_K(t-a) \quad \text{and} \quad g_I(t) := |I| + B_K(t-b) + B_K(a-t),$$

for which

$$f_I(t) \le 1_I(t) \le g_I(t) \text{ for all } t \in \mathbb{R}/\mathbb{Z}. \tag{10}$$

(ii) the Fourier coefficients[7] $\hat{B}_K(m) := \int_0^1 B_K(t)e(-mt)\,dt$ satisfy

$$\hat{B}_K(m) = \begin{cases} 0 & \text{for } |m| > K, \\ 1/(2(K+1)) & \text{if } m = 0, \\ O(1/m) & \text{for all } m \ne 0, \end{cases}$$

the implicit constant in the last estimate being absolute.

Proof of Proposition 2.3. (a) When $u = 0$ it is clear that

$$\mathcal{A}_{N,B,I}(h_1,h_2;\alpha,\beta;u) = \{n \in \mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta) : 1 \in I\} = \mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta),$$

since $1 \in I$. By Lemma 2.4, we get

$$\delta_{N,B} = \frac{\Phi_{N,B}(x)}{x} + O(4^{\pi(N)}x^{-1}). \tag{11}$$

Proposition 2.3 in the case $u = 0$ thus follows from Lemma 2.5, since for large enough x, we have

$$\sum_{n \le x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n} = \left(\frac{\delta_{N,B,1}}{krls} + o(1)\right)\log x + O(4^{\pi(N)}) \gg \frac{\delta}{krls}\frac{\Phi_{N,B}(x)}{x}\log x.$$

[7] As usual, for $t \in \mathbb{R}$ we write $e(t) := e^{2\pi i t}$.

We now assume that $u \neq 0$. Here, we no longer assume that I is an arc around 1, but rather any arc of length δ (the proof being the same regardless of the center point of the arc). Let K be a large integer, and suppose x is sufficiently large in terms of K. Write $I' \subseteq \mathbb{R}/\mathbb{Z}$ to be the interval that maps onto I under exponentiation; thus, there is a constant $c > 0$ for which $|I'| \geq c\delta$. From (10), it follows that

$$\begin{aligned}
&\sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B,I}(h_1,h_2;\alpha,\beta;u)}(n)}{n} \\
&\quad = \sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}} \frac{1_{I'}(\{u \log n/(2\pi)\})}{n} \\
&\quad \geq |I'| \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n} \\
&\qquad\quad - \sum_{|m| \leq K} \hat{B}_K(m) \bigg(e(mb) \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n^{1+imu}} + e(-ma) \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n^{1-imu}} \bigg) \\
&\quad \geq \frac{c\delta\delta_{N,B}}{krls} \log x - 2 \sum_{|m| \leq K} |\hat{B}_K(m)| \bigg| \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n^{1+imu}} \bigg|. \qquad (12)
\end{aligned}$$

We consider the second term on the RHS of (12). Applying Lemma 2.5 for each m and summing, we get

$$\begin{aligned}
\sum_{1 \leq |m| \leq K} |\hat{B}_K(m)| \bigg| \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}(n)}{n^{1+imu}} \bigg| &\ll \sum_{1 \leq m \leq K} \frac{1}{m} \bigg(\frac{\delta_{N,B}}{m|u|} + o(\log x) + (1+m|u|)4^{\pi(N)} \bigg) \\
&\ll |u|^{-1} + o((\log K)(\log x)) + K^2 |u| 4^{\pi(N)}.
\end{aligned}$$

The term with $k = 0$ gives

$$|\hat{B}_K(0)| \sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}} \frac{1}{n} \leq \frac{\log x}{2K}.$$

As such, we get

$$\sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B,I}(h_1,h_2;u)}} \frac{1}{n} \geq \bigg(\frac{c\delta\delta_{N,B}}{krls} - \frac{1}{2K} - o(\log K) \bigg) \log x - O(|u|^{-1} + K^2 |u| 4^{\pi(N)}).$$

Choosing $K = K(x)$ tending to infinity with x, but sufficiently slow-growing so that $o((\log K)(\log x)) = o(\log x)$ and $K^2 4^{\pi(N)} = o(\log x)$, proves the claim.

(b) The case when $uv = 0$ is similar to (a) and we assume $uv \neq 0$. We are interested in

$$\sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1,h_2;\alpha,\beta)}} \frac{1_{J_1}(n^{iu}) 1_{J_2}(n^{iv})}{n}. \qquad (13)$$

We minorize each indicator function by a Beurling polynomial of degree K in a similar way as to what was done just above; we shall therefore merely sketch the argument, highlighting the differences in the argument and the relevance of the condition $u/v \notin \mathbb{Q}$. As above, let us denote the minorants of 1_{J_1} and 1_{J_2} as f_{J_1}, f_{J_2}, and the corresponding majorants as g_{J_1}, g_{J_2}. Now put[8]

$$F := f_{J_1} f_{J_2} - (g_{J_1} - f_{J_1})(g_{J_2} - f_{J_2}) = f_{J_1} g_{J_2} + f_{J_2} g_{J_1} - g_{J_1} g_{J_2}.$$

That $F \leq 1_{J_1} 1_{J_2}$ is a minorant can be seen from the inequality

$$\begin{aligned} 1_{J_1}(y) 1_{J_2}(y) - F(y) &= g_{J_2}(y)(g_{J_1}(y) - f_{J_1}(y)) + 1_{J_1}(y) 1_{J_2}(y) - f_{J_2}(y) g_{J_1}(y) \\ &\geq g_{J_2}(y)(g_{J_1}(y) - f_{J_1}(y)) - f_{J_2}(y)(g_{J_1}(y) - 1_{J_1}(y)) \\ &\geq \begin{cases} (g_{J_2}(y) - f_{J_2}(y))(g_{J_1}(y) - f_{J_1}(y)) & \text{if } f_{J_2}(y) \geq 0, \\ g_{J_2}(y)(g_{J_1}(y) - f_{J_1}(y)) & \text{if } f_{J_2}(y) < 0, \end{cases} \end{aligned}$$

which implies that $1_{J_1}(y) 1_{J_2}(y) - F(y) \geq 0$ for all y since $g_{J_2} \geq 0$ and $g_{J_r} \geq f_{J_r}$ for $r = 1, 2$. We note that for $j = 1, 2$, $\widehat{(g_{J_j} - f_{J_j})}(k) = 0$ for $|k| > K$ or $k = 0$, and otherwise $\|\widehat{g_{J_j} - f_{J_j}}\|_\infty \leq 2/(K+1)$. Noting that $k_1 u \neq -k_2 v$ for all $1 \leq |k_1|, |k_2| \leq K$ since $u/v \notin \mathbb{Q}$, Lemma 2.5 gives

$$\begin{aligned} &\Bigg| \sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1, h_2; \alpha, \beta)}} \frac{1}{n} (g_{J_1} - f_{J_1}) \left(\left\{ \frac{u \log n}{2\pi} \right\} \right) (g_{J_2} - f_{J_2}) \left(\left\{ \frac{v \log n}{2\pi} \right\} \right) \Bigg| \\ &\qquad \leq \frac{4}{(K+1)^2} \sum_{1 \leq |k_1|, |k_2| \leq K} \left| \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1, h_2; \alpha, \beta)}(n)}{n^{1 + i(k_1 u + k_2 v)}} \right| \ll \max_{1 \leq |k_1|, |k_2| \leq K} \left| \sum_{n \leq x} \frac{1_{\mathcal{A}_{N,B}(h_1, h_2; \alpha, \beta)}(n)}{n^{1 + i(k_1 u + k_2 v)}} \right| \\ &\qquad = o(\log x), \end{aligned}$$

provided that x is sufficiently large (in terms of $\max_{1 \leq |k_1|, |k_2| \leq K} |k_1 u + k_2 v|^{-1}$). Thus, by the triangle inequality, (13) is bounded from below by

$$\begin{aligned} &|J_1||J_2| \sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1, h_2; \alpha, \beta)}} \frac{1}{n} - o(\log x) \\ &\quad - \Bigg(\sum_{1 \leq |m_1| \leq K} \frac{1}{|m_1|} \Bigg| \sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1, h_2; \alpha, \beta)}} \frac{1}{n^{1 + i m_1 u}} \Bigg| + \sum_{1 \leq |m_2| \leq K} \frac{1}{|m_2|} \Bigg| \sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1, h_2; \alpha, \beta)}} \frac{1}{n^{1 + i m_2 v}} \Bigg| \Bigg) - o(\log x) \\ &\quad - \sum_{1 \leq |m_1|, |m_2| \leq K} \frac{1}{|m_1 m_2|} \Bigg| \sum_{\substack{n \leq x \\ n \in \mathcal{A}_{N,B}(h_1, h_2; \alpha, \beta)}} \frac{1}{n^{1 + i(m_1 u + m_2 v)}} \Bigg| - O\left(\frac{\log x}{K} \right) - o(\log x). \end{aligned}$$

Similarly, each of the terms here are $o(\log x)$ besides the first term. The claim of part (b) thus follows just as that of part (a) did. □

Reduction to pseudopretentious functions. Let $f, g : \mathbb{N} \to \mathbb{T}$ be completely multiplicative functions. Assume, as per the hypotheses of Theorem 1.2, that $\overline{\{(f(n), g(n+1))\}_n} \neq \mathbb{T}^2$. Using Theorem 2.1, we shall show in this subsection that f and g must both be pseudopretentious.

[8] This is inspired by the construction of vector sieve weights as found in [Brüdern and Fouvry 1996]. We thank the anonymous referee for this suggestion.

Proposition 2.6. *Let $f, g : \mathbb{N} \to \mathbb{T}$ be completely multiplicative functions. Suppose $\{(f(n), g(n+1))\}_n$ is such that there is an $\varepsilon > 0$ and a pair $(z, w) \in \mathbb{T}^2$ for which*

$$\|(f(n), g(n+1)) - (z, w)\|_{\ell_1} \geq \varepsilon \text{ for all } n \text{ sufficiently large.}$$

Then there exist primitive Dirichlet characters χ_1, χ_2 to respective moduli $q_1, q_2 = O_\varepsilon(1)$, minimal positive integers $k, l = O_\varepsilon(1)$, real numbers $t_1, t_2 = O_\varepsilon(1)$ and completely multiplicative functions $h_1, h_2 : \mathbb{N} \to \mathbb{T}$ such that:

(i) $\mathbb{D}(f, h_1 n^{it_1}; x), \mathbb{D}(g, h_2 n^{it_2}; x) \ll_\varepsilon 1$.

(ii) $h_1^k = \tilde{\chi}_1$ *and* $h_2^l = \tilde{\chi}_2$.

This is a two-dimensional analogue of the reduction argument used at the beginning of Section 2 of [Klurman and Mangerel 2018]. To appropriately formalize the arguments leading to Proposition 2.6, we recall the following (essentially standard) definitions.

Definition 2.7. Let $d, N \geq 1$ and let $\{a_n\}_n \subset \mathbb{T}^d$. The *logarithmic discrepancy* (*of height N*) of $\{a_n\}_n$ is the quantity

$$D_N(\{a_n\}_n) := \sup_{\substack{I_1,\dots,I_d \subseteq \mathbb{T} \\ I_j \text{ arcs}}} \left| \frac{1}{\log N} \sum_{\substack{n \leq N \\ a_n \in B(\boldsymbol{I})}} \frac{1}{n} - \prod_{1 \leq j \leq d} |I_j| \right|,$$

where $B(\boldsymbol{I}) := \prod_{1 \leq j \leq d} I_j$. Similarly, we let $D_N^*(\{a_n\}_n)$ denote[9] the same quantity but where the sup is over all d-tuples of arcs I_j all of whose left endpoints are 1.

Definition 2.8. A sequence $\{a_n\}_n \subset \mathbb{T}^d$ is *logarithmically equidistributed* if $D_N(\{a_n\}_n) = o(1)$, as $N \to \infty$.

It is easy to see that logarithmically equidistributed sequences in $\mathbb{T}^d$ are also dense in $\mathbb{T}^d$. Before launching into the proof Proposition 2.6, we must exclude the following degenerate case from consideration.

Proposition 2.9. *Let $f, g : \mathbb{N} \to \mathbb{T}$ be completely multiplicative. Suppose exactly one of f and g is pseudopretentious. Then $\overline{\{(f(n), g(n+1))\}_n} = \mathbb{T}^2$.*

To prove Proposition 2.9, we shall need a concentration estimate, showing that if a completely multiplicative function $F : \mathbb{N} \to \mathbb{T}$ is 1-pretentious then $F(n)$ is roughly constant for *most* $n \leq x$. This will be used in the proof of Lemma 2.11 below, which will allow us to approximate pseudopretentious functions pointwise on a positive upper density subset of integers n with $P^-(n(Bn+1)) > N$ (with B and N fixed). Given a completely multiplicative function $h : \mathbb{N} \to \mathbb{T}$ and $N \geq 1$, set

$$\mathfrak{I}_h(x; N) := \sum_{N < p \leq x} \frac{\operatorname{Im}(h(p))}{p}.$$

[9]By the triangle inequality, it is easy to check that if $\{a_n\}_n \subset \mathbb{T}^d$ then

$$D_N^*(\{a_n\}_n) \leq D_N(\{a_n\}_n) \leq 2^d D_N^*(\{a_n\}_n). \tag{14}$$

We will use this relationship between D_N and D_N^* below.

For $x \geq N \geq 1$, we define $\mathbb{D}(f, 1; N, x) := (\mathbb{D}(f, 1; x)^2 - \mathbb{D}(f, 1; N)^2)^{1/2}$.

Lemma 2.10. *Let $N, B \geq 1$ with $2 \mid B$. Then as $x \to \infty$,*

$$\sum_{\substack{n \leq x \\ P^-(n(Bn+1))>N}} |f(n) - e^{i\mathfrak{I}_f(x;N)}|^2 \ll \Phi_{N,B}(x)\left(\mathbb{D}(f, 1; N, x)^2 + \frac{(\log N)^2}{N}\right) + 4^{\pi(N)} x \frac{\log_2 x}{\log x}.$$

The same estimate holds when $f(n)$ is replaced by $f(Bn+1)$.

Proof. This is a simple extension of Proposition 2.3 in [Klurman 2017], but we give the details for the reader's convenience. The claim is trivial if $\mathbb{D}(f, 1; N, x) \geq 1$. Thus, in what follows we shall assume that $\mathbb{D}(f, 1; N, x) < 1$. Define an additive function $h : \mathbb{N} \to \mathbb{C}$ on prime powers via $h(p^k) := f(p^k) - 1$. By repeatedly applying triangle inequality we have that for all $|z_i|, |w_i| \leq 1$

$$\left| \prod_{1 \leq i \leq n} z_i - \prod_{1 \leq i \leq n} w_i \right| \leq \sum_{1 \leq i \leq n} |z_i - w_i|. \tag{15}$$

Note $e^{z-1} = z + O(|z-1|^2)$ for $|z| \leq 1$. Thus, for each $n \in \mathbb{N}$, using (15) we derive

$$\begin{aligned} f(n) &= \prod_{p^k \,\|\, n} f(p^k) \\ &= \prod_{p^k \,\|\, n} (1 + h(p^k)) \\ &= e^{h(n)} + O\left(\sum_{p^k \,\|\, n} |e^{h(p^k)} - (1 + h(p^k))|\right) \\ &= e^{h(n)} + O\left(\sum_{p^k \,\|\, n} |1 - f(p^k)|^2\right) \\ &= e^{h(n)} + O\left(\sum_{p^k \,\|\, n} (1 - \mathrm{Re}(f(p^k)))\right). \end{aligned}$$

Summing over all $n \leq x$ with $P^-(n(Bn+1)) > N$, we get

$$\begin{aligned} &\sum_{\substack{n \leq x \\ P^-(n(Bn+1))>N}} |f(n) - e^{i\mathfrak{I}_f(x;N)}|^2 \\ &\qquad \ll \sum_{\substack{n \leq x \\ P^-(n(Bn+1))>N}} |e^{h(n)} - e^{i\mathfrak{I}_f(x;N)}|^2 + \sum_{p^k \leq x} (1 - \mathrm{Re}(f(p^k))) \sum_{\substack{n \leq x \\ P^-(n(Bn+1))>N}} 1_{p^k \,\|\, n} \\ &\qquad =: T_1 + T_2. \end{aligned}$$

Using Lemma 2.4, we have

$$\begin{aligned}
T_2 &= \sum_{\substack{p\le x\\ p>N}} (1-\mathrm{Re}(f(p))) \sum_{\substack{m\le x/p\\ P^-(m(Bpm+1))>N}} 1 + O\biggl(\sum_{\substack{p^k\le x\\ p>N, k\ge 2}} \frac{1}{p^k}\biggr)\\
&= x\prod_{\substack{p'\le N\\ p'\mid B}}\biggl(1-\frac{1}{p'}\biggr)\prod_{\substack{3\le p''\le N\\ p''\nmid B}}\biggl(1-\frac{2}{p''}\biggr)\sum_{N<p\le x}\frac{1-\mathrm{Re}(f(p))}{p} + O\biggl(4^{\pi(N)}\pi(x)+\frac{x}{N}\biggr)\\
&\ll \Phi_{N,B}(x)\biggl(\mathbb{D}(f,1;N,x)^2+\frac{(\log N)^2}{N}\biggr)+4^{\pi(N)}\frac{x}{\log x}.
\end{aligned}$$

We next treat T_1. Define

$$\mu_h(x) := \sum_{\substack{p^k\le x\\ p>N}} \frac{h(p^k)}{p^k}\biggl(1-\frac{1}{p}\biggr),$$

which arises as the mean value of $h(n)$ in the estimate

$$\sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} h(n) = \mu_h(x)\Phi_{N,B}(x) + O\biggl(4^{\pi(N)}\pi(x)+\frac{x}{N}\biggr),$$

using a similar argument as that which was used to bound T_2 (but keeping track of powers p^k with $k\ge 2$ as well). Writing $h(n)=\sum_{p^k\parallel n} h(p^k)$ and noticing that $\mathrm{Re}(h(p^k))\le 0$ for all primes p and all $k\ge 1$, it follows that $\mathrm{Re}(\mu_h(x))\le 0$, and we thus have

$$\sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |e^{h(n)}-e^{\mu_h(x)}|^2 \le \sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |h(n)-\mu_h(x)|^2. \tag{16}$$

Following the usual proof of the Turán–Kubilius inequality (e.g., as in Section III.2 of [Tenenbaum 1995]), we have

$$\begin{aligned}
&\sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |h(n)-\mu_h(x)|^2\\
&\qquad= \sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |h(n)|^2 - \Phi_{N,B}(x)|\mu_h(x)|^2 + O\biggl(|\mu_h(x)|\biggl(4^{\pi(N)}\pi(x)+\frac{1}{N}\biggr)\biggr)\\
&\qquad= \sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} \sum_{p^k,q^l\parallel n} h(p^k)\overline{h(q^l)} - \Phi_{N,B}(x)|\mu_h(x)|^2 + O\biggl(|\mu_h(x)|\biggl(4^{\pi(N)}\pi(x)+\frac{1}{N}\biggr)\biggr).
\end{aligned}$$

Denote by Σ the double sum in this last expression. Splitting the terms $p = q$ from $p \neq q$, Σ can be estimated with Lemma 2.4 to give

$$\begin{aligned}\Sigma &= \sum_{\substack{p^k q^l \le x\\ p\neq q,\, p,q>N}} h(p^k)\overline{h(q^l)} \sum_{\substack{m\le x/(p^k q^l)\\ P^-(m(Bp^kq^lm+1))>N}} 1_{(m,pq)=1} + \sum_{\substack{p^k\le x\\ p>N}} |h(p^k)|^2 \sum_{\substack{m\le x/p^k\\ P^-(m(Bp^km+1))>N}} 1_{(m,p)=1}\\ &= \Phi_{N,B}(x) \sum_{\substack{p^k q^l \le x\\ p\neq q,\, p,q>N}} \frac{h(p^k)\overline{h(q^l)}}{p^k q^l}\Big(1-\frac{1}{p}\Big)\Big(1-\frac{1}{q}\Big) + \Phi_{N,B}(x)\sum_{\substack{p^k\le x\\ p>N}} \frac{|h(p^k)|^2}{p^k}\Big(1-\frac{1}{p}\Big) + O(4^{\pi(N)}\pi_2(x)),\end{aligned}$$

where $\pi_2(x)$ is the number of integers $\le x$ that are product of at most two primes. Furthermore,

$$\begin{aligned}&\sum_{\substack{p^k q^l \le x\\ p\neq q,\, p,q>N}} \frac{h(p^k)\overline{h(q^l)}}{p^k q^l}\Big(1-\frac{1}{p}\Big)\Big(1-\frac{1}{q}\Big)\\ &\qquad = |\mu_h(x)|^2 + O\bigg(\sum_{\substack{p^k,\, p^l\le x\\ p>N}} \frac{|h(p^k)||h(p^l)|}{p^{k+l}} + \sum_{\substack{p^k,q^l\le x\\ p^kq^l>x,\, p,q>N}} \frac{|h(p^k)||h(q^l)|}{p^k q^l}\bigg)\\ &\qquad = |\mu_h(x)|^2 + O\Big(\sigma_h(x)^2 + \frac{1}{N} + x^{-1/2}\sigma_h(x)^2\Big),\end{aligned}$$

where we defined

$$\sigma_h(x)^2 := \sum_{\substack{p^k\le x\\ p>N}} \frac{|h(p^k)|^2}{p^k}\Big(1-\frac{1}{p}\Big),$$

and used Cauchy–Schwarz in the last line. It follows that

$$\begin{aligned}\Sigma - \Phi_{N,B}(x)|\mu_h(x)|^2 &\ll \Phi_{N,B}(x)\Big(\sigma_h(x)^2 + \frac{1}{N}\Big) + 4^{\pi(N)}\pi_2(x)\\ &\ll \Phi_{N,B}(x)\Big(\sigma_h(x)^2 + \frac{1}{N}\Big) + 4^{\pi(N)} x\frac{\log_2 x}{\log x}.\end{aligned}$$

Hence,

$$\begin{aligned}\sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |h(n)-\mu_h(x)|^2 &\ll \Phi_{N,B}(x)\Big(\sigma_h(x)^2 + \frac{1}{N}\Big) + 4^{\pi(N)}\Big(x\frac{\log_2 x}{\log x} + |\mu_h(x)|\pi(x)\Big)\\ &\ll \Big(\mathbb{D}(f,1;N,x)^2 + \frac{1}{N}\Big)\Phi_{N,B}(x) + 4^{\pi(N)} x\frac{\log_2 x}{\log x}, \qquad (17)\end{aligned}$$

using $|\mu_h(x)| \le 2\sum_{p\le x} 1/p \le 2\log_2 x + O(1)$ and the estimate

$$\sigma_h(x)^2 = \sum_{\substack{p^k\le x\\ p>N}} \frac{|1-f(p^k)|^2}{p^k}\Big(1-\frac{1}{p}\Big) \ll \sum_{N<p\le x} \frac{1-\mathrm{Re}(f(p))}{p} + O\Big(\frac{1}{N}\Big) = \mathbb{D}(f,1;N,x)^2 + O\Big(\frac{1}{N}\Big).$$

Inserting (17) into (16), and using this in the definition of T_1, we get

$$\begin{aligned}
&\sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |e^{h(n)} - e^{i\mathfrak{I}_f(x;N)}|^2 \\
&\quad\ll \sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |e^{\mu_h(x)} - e^{i\mathfrak{I}_f(x;N)}|^2 + \sum_{\substack{n\le x\\ P^-(n(Bn+1))>N}} |e^{h(n)} - e^{\mu_h(x)}|^2 \\
&\quad\ll \Phi_{N,B}(x)|e^{\mu_h(x)} - e^{i\mathfrak{I}_f(x;N)}|^2 + \Phi_{N,B}(x)\left(\mathbb{D}(f,1;N,x)^2 + \frac{1}{N}\right) + 4^{\pi(N)}x\frac{\log_2 x}{\log x}. \qquad (18)
\end{aligned}$$

Moreover, as

$$\begin{aligned}
\mu_h(x) &= -\sum_{N<p\le x}\frac{1-\mathrm{Re}(f(p))}{p} + i\sum_{N<p\le x}\frac{\mathrm{Im}(f(p))}{p} + O(1/N) \\
&= -\mathbb{D}(f,1;N,x)^2 + i\cdot\mathfrak{I}_f(x;N) + O\left(\frac{1}{N}\right),
\end{aligned}$$

it follows that

$$|e^{i\mathfrak{I}_f(x;N)} - e^{\mu_h(x)}|^2 = |1 - e^{-\mathbb{D}(f,1;N,x)^2 + O(1/N)}|^2 \ll \mathbb{D}(f,1;N,x)^2 + \frac{1}{N}.$$

Coupled with (18), this completes the proof of the first assertion of the lemma. The assertion with $f(Bn+1)$ in place of $f(n)$ is proved similarly, and we leave the details for the reader. □

Lemma 2.11. *Let $\eta > 0$ and $C > 1$. Let $B \ge 1$ be an even integer. Suppose furthermore that $f : \mathbb{N} \to \mathbb{T}$ is a pseudopretentious multiplicative function, with f pretending to be $h(n)n^{it}$ with h a pseudocharacter and $t \in \mathbb{R}$. Then there is an infinite sequence $\{x_j\}_j$ of positive real numbers and a large parameter N (depending at most on η) such that if $j = j(N)$ is chosen sufficiently large then*

$$f(n) = h(n)n^{it} + O(\eta)$$

for all but $O_C(\eta^C\Phi_{N,B}(x_j))$ choices of $n \le x_j$ with $P^-(n(Bn+1)) > N$. Similarly, $f(Bn+1) = h(Bn+1)(Bn+1)^{it} + O(\eta)$ for all but $O_C(\eta^C\Phi_{N,B}(x_j))$ choices of $n \le x_j$ with $P^-(n(Bn+1)) > N$.

Proof. The result is trivial for any $\eta \gg_C 1$, so in what follows we shall assume that η is smaller than any fixed bound depending on C. Let $F(n) := f(n)\overline{h(n)}n^{-it}$ for each n. We consider several cases, according to the behavior of the series

$$\mathfrak{I}_F(\infty) := \lim_{x\to\infty}\mathfrak{I}_F(x;1) = \lim_{x\to\infty}\sum_{p\le x}\frac{\mathrm{Im}(F(p))}{p}.$$

First, suppose $\mathfrak{I}_F(\infty)$ converges absolutely. Then, choosing N large enough in terms of η, it follows that $\mathfrak{I}_F(x;N) < \eta/2$ for all $x > N$. In this case, we shall choose $\{x_j\}_j$ to be the set of all $x > N$. Next, suppose that $\mathfrak{I}_F(\infty)$ is unbounded. In this case, since $\mathfrak{I}_F(n+1;N) - \mathfrak{I}_F(n;N) \to 0$ as $n \to \infty$, it follows that the sequence of fractional parts of $\mathfrak{I}_f(x;N)/2\pi$ must be dense in [0, 1], for any N. In this case, we may choose any large N and $\{x_j\}_j$ to be a sequence for which $\|\mathfrak{I}_f(x_j;N)/2\pi\| \to 0$, as $j \to \infty$. Lastly, suppose $x \mapsto \mathfrak{I}_F(x;1)$ is bounded but not convergent. Let α be a limit point of $\mathfrak{I}_F(x;1)$, and choose $\{x_j\}_j$ to be a sequence for which $\mathfrak{I}_F(x_j;N) \to \alpha$. Picking j_0 sufficiently large in terms of η, then setting $N := x_{j_0}$ it follows that $\mathfrak{I}_F(x_j;N) < \eta/2$ for large enough $j > j_0$. As F is 1-pretentious, we can assume N is chosen large enough in terms of η (as in the analysis above) so that $\mathbb{D}(F,1;N,\infty)^2 + 1/N \ll \eta^{C+2}$. Furthermore, as discussed above we have selected a sequence $\{x_j\}_j$ such that $\|\mathfrak{I}_F(x_j;N)/2\pi\| < \eta/2$. Taking j sufficiently large in terms of N, Lemma 2.10 and Chebyshev's inequality give

$$\begin{aligned}
&\left|\left\{n \le x_j : P^-(n(Bn+1)) > N, |F(n) - e^{i\mathfrak{I}_F(x_j;N)}| > \frac{\eta}{2}\right\}\right| \\
&\qquad \le 4\eta^{-2} \sum_{\substack{n \le x \\ P^-(n(Bn+1))>N}} |F(n) - e^{i\mathfrak{I}_F(x_j;N)}|^2 \\
&\qquad \ll \eta^{-2}\Phi_{N,B}(x_j)\left(\mathbb{D}(F,1;N,x_j)^2 + \frac{1}{N}\right) + \eta^{-2}4^{\pi(N)} x_j \frac{\log_2 x_j}{\log x_j} \\
&\qquad \ll \Phi_{N,B}(x_j)\eta^C \\
&\qquad < \Phi_{N,B}(x_j)\frac{\eta}{2},
\end{aligned} \tag{19}$$

provided that η is sufficiently small (since $C > 1$). For all other $n \le x_j$ with $P^-(n(Bn+1)) > N$ we choose j larger if necessary so that $\|\mathfrak{I}_f(x_j;N)/2\pi\| < \eta/2$. It follows that

$$|f(n) - h(n)n^{it}| = |F(n) - 1| \le |F(n) - e^{i\mathfrak{I}_F(x_j;N)}| + |e^{i\mathfrak{I}_F(x_j;N)} - 1| < \frac{\eta}{2} + \frac{\eta}{2} = \eta,$$

which implies the first claim. The claim with $f(Bn+1)$ follows in the same way (with the same subsequence $\{x_j\}_j$) from Lemma 2.10 (which holds with $f(Bn+1)$ as well). $\square$

Remark 2.12. In the sequel, we will apply Lemma 2.11 twice in a context in which a common subsequence will be sought for a pair of functions f and g (see the proofs of Lemma 3.2 and Proposition 1.5 below), and it is not immediately clear why such a subsequence is available in general. Fortunately, in both of these contexts, the functions f and g will be finite-valued, say taking values in μ_K for some $K \in \mathbb{N}$. A particular feature of this case is that if h_1 and h_2 are pseudocharacters associated with f and g respectively, and $F := f\overline{h_1}$ and $G := g\overline{h_2}$ then the fact that $\mathbb{D}(F,1;x), \mathbb{D}(G,1;x) \ll 1$ immediately implies that $F(p) = 1 = G(p)$ except on a set S of primes for which $\sum_{p\in S} 1/p < \infty$. This means, in particular, that $\mathfrak{I}_F(x), \mathfrak{I}_G(x)$ both converge absolutely as $x \to \infty$. Thus, according to the above proof, we can take our subsequence $\{x_j\}_j$ for f to consist of all sufficiently large x, and the same subsequence is admissible for g. Thus, a choice of common infinite subsequence for both f and g necessarily exists.

Proof of Proposition 2.9. Suppose there is an $\varepsilon > 0$ and a pair $(z,w) \in \mathbb{T}^2$ such that $(f(n), g(n+1)) \notin B_\varepsilon((z,w))$ for all large n.[10] Let h be a pseudocharacter such that f is hn^{it}-pretentious, and assume it has order kr (i.e., $h : \mathbb{N} \to \mu_{kr}$). Select $\{n_j\}_j$ on which $f(2n_j) \to z$, and hence for j chosen large enough

[10] Here and elsewhere, we write $B_\varepsilon((z,w))$ to denote the ε-ball about (z,w) in $\mathbb{C}^2$ with respect to the ℓ^1-norm.

we have $f(2n_j) = z + O(\varepsilon^2)$. By Lemma 2.11, we can choose N sufficiently large in terms of ε and a sequence $\{x_\ell\}_\ell$ such that,

$$f(2n_j m) = zh(m)m^{it} + O(\varepsilon^2),$$

for all but $O(\varepsilon^3 \Phi_{N,2n_j}(x_\ell))$ integers $m \le x_\ell$ with $P^-(m(2n_j m+1)) > N$. Thus, if ε is sufficiently small then for all but a small number of exceptions, we have

$$(h(m)m^{it}, g(2n_j m+1)) \notin B_{2\varepsilon/3}((1, w)).$$

Let I be a symmetric arc of length $\varepsilon/4$ about 1, and let J be a symmetric arc of the same length about w. To yield a contradiction, we shall presently show that there exists a sufficiently dense subset of integers $m \le x_\ell$ that satisfies the following properties:

(i) $h(m) = 1$.

(ii) $P^-(m(2n_j m+1)) > N$.

(iii) $g(2mn_j + 1) \in J$.

(iv) $m^{it} \in I$.

The proof of this argument is similar to that of Proposition 2.3. For convenience, put $X := x_\ell$, for some sufficiently large index ℓ (depending at most on ε). Then we wish to bound

$$\mathfrak{G} = \sum_{\substack{m \le X \\ P^-(m(2n_j m+1)) > N}} \frac{1_I(m^{it}) 1_{h(m)=1} 1_J(g(2mn_j+1))}{m}.$$

We write the arcs I and J as the images of intervals $[a, b]$ and $[c, d]$ (say) in $\mathbb{R}$ under the map $t \mapsto e(t)$, such that $[a, b]$ contains an integer (so that I contains 1). We consider two cases, depending on t. First, if $t = 0$ then $1_I(m^{it}) = 1$ for all m. Using the properties of Beurling polynomials, we have the minorization

$$\begin{aligned} 1_J(g(2n_j m+1)) &\ge \Big(|J| - \sum_{0 \le |l| \le K} \hat{B}_K(l)(e(dl)g(2n_j m+1)^{-l} + e(-cl)g(2n_j m+1)^l)\Big) \\ &\ge \Big(|J| - \sum_{1 \le |l| \le K} \hat{B}_K(l)(e(dl)g(2n_j m+1)^{-l} + e(-cl)g(2n_j m+1)^l)\Big) - \frac{1}{K+1}, \end{aligned} \tag{20}$$

for any parameter $K = K(X)$ to be chosen. Inserting this into the definition of $\mathfrak{G}$ and summing over $m \le X$ with the given conditions, we get

$$\mathfrak{G} \ge |J| \sum_{\substack{m \le X \\ P^-(m(2n_j m+1)) > N}} \frac{1_{h(m)=1}}{m} - \sum_{1 \le |l| \le K} \hat{B}_K(l)(e(dl)M_{-l}(X) + e(-cl)M_l(X)) - O\Big(\frac{\log X}{K}\Big),$$

where for any $l \in \mathbb{Z}$ we have put

$$M_l(X) := \sum_{\substack{m \le X \\ P^-(m(2n_jm+1))>N}} \frac{1_{h(m)=1} g(2n_jm+1)^{-l}}{m}.$$

Expressing $1_{h(m)=1} = \frac{1}{kr}\sum_{\nu(kr)} h(m)^\nu$, we have

$$M_l(X) = \frac{1}{kr}\sum_{\nu(kr)} \sum_{\substack{m \le X \\ P^-(m(2n_jm+1))>N}} \frac{h(m)^\nu g(2n_jm+1)^{-l}}{m} =: \frac{1}{kr}\sum_{0\le\nu\le kr-1} M_l(X;\nu),$$

for any $l \in \mathbb{Z}$. Inserting this into our lower bound for $\mathfrak{G}$, we find

$$\begin{aligned}\mathfrak{G} \ge \frac{1}{kr}\sum_{0\le\nu\le kr-1} |J| \sum_{\substack{m \le X \\ P^-(m(2n_jm+1))>N}} \frac{h(m)^\nu}{m} - O((\log X)/K) \\ - \frac{1}{kr}\sum_{0\le\nu\le kr-1}\sum_{1\le|l|\le K} \hat{B}_K(l)(e(dl)M_{-l}(X;\nu) + e(-cl)M_l(X;\nu)) \\ =: \frac{1}{kr}\sum_{0\le\nu\le kr-1} \mathcal{S}_\nu - O((\log X)/K). \qquad (21)\end{aligned}$$

Fix $0 \le \nu \le kr-1$. Now, by hypothesis, g is not pseudopretentious, so that $g^l 1_{P^->N}$ is nonpretentious for all $l \in \mathbb{Z}\backslash\{0\}$. Hence, Theorem 2.1 implies that

$$M_l(X;\nu) = \sum_{m\le X} \frac{(h1_{P^->N})(m)^\nu (g1_{P^->N})(2n_jm+1)^l}{m} = o(\log X),$$

for all $1 \le |l| \le K$. Multiplying by $1/|l|$ and summing over $1 \le l \le K$ in the above estimate, we get

$$\mathcal{S}_\nu = |J| \sum_{\substack{m \le X \\ P^-(m(2n_jm+1))>N}} \frac{h(m)^\nu}{m} + o((\log X)(\log K)),$$

for all $0 \le \nu \le kr-1$. Recall that h^j is nonpretentious for all $1 \le j \le r-1$, and thus h^{tr+j} is nonpretentious for all $0 \le t \le q-1$. In particular, by Theorem 2.1 we have

$$\sum_{m\le X} \frac{h^{tr+v} 1_{P^->N}(m) 1_{P^->N}(2n_jm+1)}{m} = o(\log X).$$

Now suppose $\nu = vr$, for $v \ne 0$. Then $h^\nu = \chi^v$, and thus splitting m according to residue classes modulo q (noting that $N > q$), we get

$$\sum_{m\le X} \frac{\chi^v(m) 1_{P^->N}(m) 1_{P^->N}(2n_jm+1)}{m} = \sideset{}{^*}\sum_{a \pmod q} \chi(a)^v \sum_{\substack{m\le X \\ m\equiv a \pmod q}} \frac{1_{P^->N}(m) 1_{P^->N}(2n_jm+1)}{m}.$$

If $(2n_j a+1,q)=1$ then Lemma 2.4 and partial summation implies that

$$\sum_{\substack{m\le X\\ m\equiv a \pmod q}} \frac{1_{P^->N}(m)1_{P^->N}(2n_jm+1)}{m} = \delta_{N,2n_j,q}\log X + O(4^{\pi(N)}).$$

On the other hand, if $(2n_ja+1,q)>1$ then the sum is 0. It follows that

$$\begin{aligned}\sum_{1\le v\le k-1}\sum_{m\le X} &\frac{\chi^v(m)1_{P^->N}(m)1_{P^->N}(2n_jm+1)}{m}\\ &= \delta_{N,2n_j,q}(\log X)\sum_{1\le v\le k-1}\sum_{a \pmod q}\chi(a)^v\chi_0(2n_ja+1)+O_q(4^{\pi(N)})\\ &= o(\log X),\end{aligned}$$

for x suitably large relative to N, since the Jacobi sum is 0 for all $1\le j\le k-1$. It follows that when $\nu\ne 0$, we have

$$S_\nu = o((\log X)(\log K)) = o(\log X),$$

if $K=K(X)$ is chosen to be tending to infinity with X sufficiently slowly, and therefore

$$\mathfrak{G}\ge \frac{|J|}{kr}\sum_{\substack{m\le X\\ P^-(m(2n_jm+1))>N}}\frac{1}{m} - o(\log X).$$

Next, suppose $t\ne 0$. Arguing as before, this time invoking a minorization like (20) with $1_I(n^{it})$ as well and using the triangle inequality, we have (see (21) and the lines preceding it)

$$\begin{aligned}\mathfrak{G}&\ge\frac{1}{kr}\sum_{0\le\nu\le kr-1}|I||J|\sum_{\substack{m\le X\\ P^-(m(2n_jm+1))>N}}\frac{h(m)^\nu}{m}-\frac{1}{kr}\sum_{0\le\nu\le kr-1}\sum_{\substack{0\le|l_1|,|l_2|\le m\\ \max\{|l_1|,|l_2|\}\ge 1}}\frac{1}{R(l_1,l_2)}\sum_{u,v\in\{-1,+1\}}|M_{ul_1,vl_2}(X;\nu)|\\ &=:\frac{1}{kr}\sum_{0\le\nu\le kr-1}T_\nu,\end{aligned}$$

where $R(l_1,l_2):=\max\{1,|l_1|\}\max\{1,|l_2|\}$ whenever $l_1l_2\ne 0$, and we have set

$$M_{l_1,l_2}(X;\nu):=\sum_{\substack{m\le X\\ P^-(m(2n_jm+1))>N}}\frac{h(m)^\nu g(2n_jm+1)^{l_1}m^{il_2t}}{m}.$$

Since $n\mapsto g(n)^{l_2}1_{P^-(n)>N}n^{il_1t}$ is nonpretentious for all $l_2\ne 0$, and $n\mapsto h(n)^\nu 1_{P^-(n)>N}n^{-il_1t}$ is nonpretentious whenever $\max\{|l_1|,|l_2|\}\ge 1$, Theorem 2.1 implies that $M_{l_1,l_2}(X;\nu)=o(\log X)$ for all such l_1,l_2

and any ν. In particular,

$$T_\nu = |I||J| \sum_{\substack{m\le X\\ P^-(m(2n_jm+1))>N}} \frac{h(m)^\nu}{m} + o\Big((\log X)\sum_{1\le l_1,l_2\le K}\frac{1}{l_1l_2}\Big)$$
$$= |I||J| \sum_{\substack{m\le X\\ P^-(m(2n_jm+1))>N}} \frac{h(m)^\nu}{m} + o((\log X)(\log K)^2).$$

As above, if $\nu\neq 0$ then $T_\nu = o(\log X)$. Hence, we get

$$\mathfrak{G} \ge |I||J| \sum_{\substack{m\le X\\ P^-(m(2n_jm+1))>N}} \frac{1}{m} - o(\log X),$$

for $K = K(X)\to\infty$ sufficiently slowly. Invoking Lemma 2.4 (with $q = 1$) and partial summation, we get that when X is sufficiently large relative to N,

$$\sum_{\substack{m\le X\\ P^-(m(2n_jm+1))>N}} \frac{1}{m} = \delta_{N,B}\log X + o(\log X).$$

Since $|I|\ge |I||J|\gg \varepsilon^2$, this shows that

$$\mathfrak{G} \gg \frac{\varepsilon^2}{kr}\frac{\Phi_{N,B}(X)}{X}\log X.$$

Hence, assuming (as we may) that $\varepsilon \ll (kr)^{-2}$, the set of $n = 2mn_j$ with $h(m) = 1$, $|m^{it}-1| < \varepsilon/4$, $|g(mn_j+1)-w| < \varepsilon/4$ and $P^-(m(2n_jm+1)) > N$ (i.e., satisfying properties (i)–(iv) above) intersects in a set of positive upper density with the set of $n\le 2n_jX$ where $f(m) = h(m)m^{it} + O(\varepsilon^2)$. It thus follows that for a set of integers n with positive upper density, we have

$$|f(n)-z| = |f(m)-1| + O(\varepsilon^2) = |h(m)m^{it}-1| + O(\varepsilon^2) = |m^{it}-1| + O(\varepsilon^2) < \frac{\varepsilon}{3}$$

and $|g(n+1)-w| = |g(2mn_j+1)-w| < \varepsilon/4$. Consequently, $(f(n), g(n+1))\in B_{2\varepsilon/3}((z,w))$ which contradicts our initial assumption. This contradiction completes the proof. □

Having established the above proposition, we now know that if $\overline{\{(f(n), g(n+1))\}_n}\neq \mathbb{T}^2$ then either both f and g are pseudopretentious, or neither is. To complete the proof of Proposition 2.6, therefore, we shall show that the latter case is not possible.

Proof of Proposition 2.6. Let us assume for the sake of contradiction that one of f and g are not pseudopretentious. Proposition 2.9 implies that, then, both f *and* g are not pseudopretentious. Let $a, b:\mathbb{N}\to\mathbb{R}/\mathbb{Z}$ be completely additive functions for which $f(n) = e(a(n))$ and $g(n) = e(b(n))$. For each $M\in\mathbb{N}$ let

$$F_M(x_1,x_2) := \frac{1}{\log M}\sum_{\substack{n\le M\\ (a(n),b(n+1))\in[0,x_1]\times[0,x_2]}} \frac{1}{n},$$

and $G_M(x_1, x_2) = x_1 x_2$ identically for all M. Applying Theorem 2 of [Niederreiter and Philipp 1973] gives that for any $K \geq 1$ we have

$$D_M^*(\{(f(n), g(n+1))\}_n) = \sup_{x_1, x_2 \in [0,1)} |F_M(x_1, x_2) - G_M(x_1, x_2)| \tag{22}$$

$$\ll \frac{1}{K+1} + \frac{1}{\log M} \sum_{\substack{0 \leq |m_1|, |m_2| \leq K \\ (m_1, m_2) \neq (0,0)}} \frac{1}{R(m_1, m_2)} \left| \sum_{n \leq M} \frac{f(n)^{m_1} g(n+1)^{m_2}}{n} \right|, \tag{23}$$

where $R(m_1, m_2) = \max\{1, |m_1|\} \max\{1, |m_2|\}$ whenever $m_1 m_2 \neq 0$. Now, by hypothesis, there is an $\varepsilon > 0$ and a point $(z, w) \in \mathbb{T}^2$ such that $(f(n), g(n+1)) \notin B_\varepsilon((z, w))$ for all large n. In particular, this means that if I_z and J_w are, respectively, the symmetric arcs of length 2ε about z and about w, then

$$\begin{aligned} D_M(\{f(n), g(n+1)\}_n) &= \sup_{\substack{I, J \subset \mathbb{T} \\ \text{arcs}}} \left| \frac{1}{\log M} \sum_{\substack{n \leq M \\ (f(n), g(n+1)) \in I \times J}} \frac{1}{n} - |I||J| \right| \\ &\geq \left| \frac{1}{\log M} \sum_{\substack{n \leq M \\ (f(n), g(n+1)) \in I_z \times J_w}} \frac{1}{n} - B_\varepsilon((z, w)) \right| \\ &= |I_z||J_w| \\ &\gg \varepsilon^2. \end{aligned}$$

From (14), we get $D_M^*(\{(f(n), g(n+1))\}_n) \gg \varepsilon^2$. Combining this with (22) and choosing $K = \lfloor C/\varepsilon^2 \rfloor$ with $C > 0$ a sufficiently large constant, the pigeonhole principle implies that for some pair $(m_1, m_2) \neq (0, 0)$ with $|m_1|, |m_2| \leq K$

$$\left| \sum_{n \leq M} \frac{f(n)^{m_1} g(n+1)^{m_2}}{n} \right| \gg_\varepsilon \log M.$$

Note that since neither f nor g is assumed to be pseudopretentious, we must have $m_1 m_2 \neq 0$ as otherwise Halász' theorem would yield

$$\left| \sum_{n \leq M} \frac{f(n)^{m_1}}{n} \right|, \left| \sum_{n \leq M} \frac{g(n)^{m_2}}{n} \right| = o(\log M).$$

Thus, we may apply Proposition 2.2, which gives that f and g are indeed *both* pseudopretentious and the result follows. □

3. Proof of Proposition 1.7

Before addressing Propositions 1.5 and Theorem 1.6, we pause to prove Proposition 1.7, as its proof will be related to several subcases of Theorem 1.6 especially. Write $f(n) = f_0(n) n^{it}$ and $g(n) = g_0(n) n^{it'}$, where f_0 and g_0 are completely multiplicative functions taking values in μ_k and μ_l, respectively. We show in this section that if $\overline{\{f(n), g(n+1)\}_n} \neq \mathbb{T}^2$ then $t/t' \in \mathbb{Q}$. To that end, we will need the following lemma, which is a standard extension of a "repulsion" estimate for the pretentious distance due to Granville and

Soundararajan [2007] (see, for instance, Lemma C.1 in [Matomäki et al. 2015] for the case $k = 2$). We recall that for arithmetic functions $F, G : \mathbb{N} \to \mathbb{U}$, where $\mathbb{U}$ denotes the closed unit disc, and $x \geq 2$, we set

$$\mathbb{D}(F, G; x) := \left(\sum_{p \leq x} \frac{1 - \operatorname{Re}(F(p)\overline{G(p)})}{p} \right)^{1/2}.$$

Lemma 3.1. *Let $k \geq 1$ and let $f : \mathbb{N} \to \mu_k$ be a multiplicative function. Then*

$$\inf_{|\tau| \leq x} \mathbb{D}(f, n^{i\tau}; x) \geq \frac{1}{2k} \min\{\sqrt{\log_2 x}, \mathbb{D}(f, 1; x)\} - O_k(1).$$

Proof. We first assume that $\tau > 1$. The triangle inequality [Granville and Soundararajan 2007] gives

$$k\mathbb{D}(f, n^{i\tau}; x) \geq \mathbb{D}(f^k, n^{i\tau k}; x) = \mathbb{D}(1, n^{i\tau k}; x). \tag{24}$$

Now, the Vinogradov–Korobov zero-free region (see, for instance, Section 9.5 of [Montgomery 1994]) gives that

$$|\zeta(1 + 1/\log x + i\tau)| \ll (\log(2 + |\tau|))^{0.67} \text{ for all } 1 \leq |t| \leq x^2.$$

It follows that

$$\begin{aligned} \mathbb{D}(1, n^{i\tau k}; x)^2 = \sum_{p \leq x} \frac{1 - \operatorname{Re}(p^{-ik\tau})}{p} &\geq \log_2 x - \log \left| \zeta\left(1 + \frac{1}{\log x} + i\tau k\right)\right| - O(1) \\ &\geq \log_2 x - 0.67 \log_2 |k\tau| - O(1) \geq 0.33 \log_2 x - O_k(1). \end{aligned}$$

Inserting this estimate into (24), we get that

$$\mathbb{D}(f, n^{i\tau}; x) \geq \frac{1}{2k} \sqrt{\log_2 x} - O_k(1) \tag{25}$$

in this case. Suppose now that $|\tau| \leq 1$. We may assume that

$$\mathbb{D}(1, n^{i\tau k}; x) < \tfrac{1}{2}\mathbb{D}(f, 1; x) \tag{26}$$

since in the opposite case the lemma follows immediately from (24). By the prime number theorem, for each $|u| \leq k$ the asymptotic

$$\mathbb{D}(1, n^{iu}; x)^2 = \log(1 + |u| \log x) + O_k(1)$$

holds. In particular,

$$\mathbb{D}(1, n^{i\tau k}; x)^2 = \log(1 + k|\tau| \log x) + O_k(1) = \log(1 + |\tau| \log x) + O_k(1) = \mathbb{D}(1, n^{i\tau}; x)^2 + O_k(1),$$

for $|\tau| \leq 1$. Applying the triangle inequality once again, then invoking (24) and (26), we find that

$$\begin{aligned} \mathbb{D}(f, 1; x) &\leq \mathbb{D}(1, n^{i\tau}; x) + \mathbb{D}(f, n^{i\tau}; x) \\ &= \mathbb{D}(1, n^{i\tau k}; x) + \mathbb{D}(f, n^{i\tau}; x) + O_k(1) \\ &< \tfrac{1}{2}\mathbb{D}(f, 1; x) + k\mathbb{D}(f, n^{i\tau}; x) + O_k(1). \end{aligned}$$

Combined with (25), this yields the claim. $\square$

The next result shows that for most n with $P^-(n(Bn+1))$ large, we can reduce to the case that $f_0 = h_1$ and $g_0 = h_2$, where h_1, h_2 defined as in Proposition 2.6. This will allow us to use the results of previous sections in proving Proposition 1.7.

Lemma 3.2. *Let $f(n) = f_0(n)n^{it}$ and $g(n) = g_0(n)n^{it'}$, where f_0 and g_0 are completely multiplicative functions taking values in μ_k and μ_l, respectively. Let $\eta > 0$ be sufficiently small (in terms of k and l). Then there exist pseudocharacters h_1, h_2 with the following properties. Let B be an even integer. Then there is a subsequence $\{x_j\}_j$ and a parameter N depending at most on η such that if j is sufficiently large (in terms of η) then the following holds: for all but $O(\eta\Phi_{N,B}(x_j))$ integers $n \leq x_j$ with $P^-(n(Bn+1)) > N$, we have $f_0(n) = h_1(n)$ and $g_0(Bn+1) = h_2(Bn+1)$.*

Proof. Since $f(n)^k = n^{it}$, we have

$$\sum_{n\leq x} \frac{f(n)^k \overline{f(n+1)}^k}{n} = \sum_{n\leq x} \frac{1}{n} + O(k|t|) \gg \log x.$$

A similar estimate holds with g^l in place of f^k. Thus, by Proposition 2.2 we have that f and g are both pseudopretentious, and that there are real numbers $t_1, t_2 \in \mathbb{R}$ with $t_1, t_2 = O(1)$, primitive characters χ_1 and χ_2 with conductors q_1 and q_2 and $q_1, q_2 = O(1)$ and completely multiplicative functions h_1, h_2 such that $h_1(n)^k = \chi_1(n/(n, q_1^\infty))$ and $h_2(n)^l = \chi_2(n/(n, q_2^\infty))$, and such that

$$\mathbb{D}(f_0, h_1 n^{i(t_1-t)}; x) = \mathbb{D}(f, h_1(n)n^{it_1}; x) \ll 1$$
$$\mathbb{D}(g_0, h_2 n^{i(t_2-t')}; x) = \mathbb{D}(f, h_1(n)n^{it_1}; x) \ll 1.$$

Suppose that when χ_1 and χ_2 do not vanish, they take values in μ_r and μ_s, respectively. We begin by showing that $t_1 = t$ and $t_2 = t'$. Applying Lemma 3.1 with $F = f_0\overline{h_1}$,

$$\begin{aligned} 1 \gg \mathbb{D}(F, n^{i(t_1-t)}; x) &\geq \inf_{|u|\leq x} \mathbb{D}(F, n^{iu}; x) \\ &\geq \frac{1}{2kr} \min\{\sqrt{\log_2 x}, \mathbb{D}(F, 1; x)\} - O_k(1) \\ &= \frac{1}{2kr}\mathbb{D}(F, 1; x) - O_k(1), \end{aligned}$$

It follows that $\mathbb{D}(f_0, h_1; x) \ll_{k,r} 1$, and hence the triangle inequality yields

$$\mathbb{D}(1, n^{i(t_1-t)}; x) \leq \mathbb{D}(f_0\overline{h_1}, 1; x) + \mathbb{D}(f_0\overline{h_1}, n^{i(t_1-t)}; x) \ll_{k,r} 1.$$

Arguing as in the proof of Lemma 3.1 it is clear that for x sufficiently large,

$$\mathbb{D}(1, n^{i(t_1-t)}; x) = \log_2 x - \log\left|\zeta\left(1 + \frac{1}{\log x} + i(t_1 - t)\right)\right| + O(1) \geq \log_2 x - O\left(1 + \log\left(\frac{\log x}{1 + |t_1 - t|\log x}\right)\right)$$

can only be bounded if $t_1 = t$. Similarly, we must take $t_2 = t$. Now, set $F_0(n) := f_0(n)\overline{h_1(n)}$ and $G_0(n) := g_0(n)\overline{h_2(n)}$. Lemma 2.11 implies that for a suitable choice of N (depending only on η, f and

g) and a common subsequence $\{x_j\}_j$ (see Remark 2.12 for an explanation of why this exists), we get that for large enough j,

$$|\{n \le x_j : P^-(n(Bn+1)) > N, |F_0(n) - 1| > \eta\}| \ll \eta \Phi_{N,B}(x_j)$$
$$|\{n \le x_j : P^-(n(Bn+1)) > N, |G_0(Bn+1) - 1| > \eta\}| \ll \eta \Phi_{N,B}(x_j).$$

It follows that for all but $O(\eta\Phi_{N,B}(x_j))$ integers $n \le x_j$ with $P^-(n(Bn+1)) > N$, we have that $F_0(n) = 1 + O(\eta)$, and $G_0(Bn+1) = 1 + O(\eta)$, for N sufficiently large. But since F_0 and G_0 take discrete values, for these n, $F_0(n) = G_0(Bn+1) = 1$ when η is sufficiently small (in terms of k and l). In light of the definition of F_0 and G_0, this implies the claim. $\square$

Proof of Proposition 1.7. We suppose that $f(n)^k = n^{it}$ and $g(n) = n^{it'}$, for all $n \in \mathbb{N}$. Let $f(n) = f_0(n)n^{it_1}$ and $g(n) = g_0(n)n^{it_2}$, where $f_0^k = 1 = g_0^l$, $t_1 := t/k$ and $t_2 := t'/l$. As $(n+1)^{it_2} = n^{it_2} + O(|t_2|/n)$, it is equivalent to consider when the sequence of pairs $(f_0(n)n^{it_1}, g_0(n+1)n^{it_2})$ is, or is not dense for n sufficiently large. We first suppose that $t/t' = r_1/s_1 \in \mathbb{Q}\backslash\{0\}$. In this case, since

$$\overline{(f(n)^{ks_1}, g(n+1)^{lr_1})} = \overline{(n^{is_1t_1}, (n+1)^{ir_1t_2})} = \{(z, z) \in \mathbb{T}^2\} \ne \mathbb{T}^2,$$

we have $\overline{(f(n), g(n+1)} \ne \mathbb{T}^2$. Suppose now that $t/t' \notin \mathbb{Q}$, and that $\{(f_0(n)n^{it_1}, g(n+1)n^{it_2})\}_n$ is not dense in $\mathbb{T}^2$. By Proposition 2.6, we know that f and g must be pseudopretentious. Thus, let h_1 and h_2 be pseudocharacters with conductors q_1 and q_2 respectively, and $u, v \in \mathbb{R}$ such that f is h_1n^{iu}-pretentious and g is h_2n^{iv}-pretentious. From the proof of Lemma 3.2, it follows that $u = t_1$ and $v = t_2$. Now, replacing n by Bn, where $B = 2q_1q_2$, it follows that

$$(f_0(n)n^{it_1}, g_0(Bn+1)n^{it_2}) \notin B_\varepsilon(z\overline{f(B)}B^{-it_1}, wB^{-it_2}).$$

Set $(z', w') = (z\overline{f(B)}B^{-it_1}, wB^{-it_2})$. Now, according to Lemma 3.2 (applied with $\eta = \varepsilon^3$), we can choose x sufficiently large (belonging to a prescribed infinite subsequence) and N suitably large in terms of ε, such that $f_0(n) = h_1(n)$ and $g_0(Bn+1) = h_2(Bn+1)$, for all but $O(\varepsilon^3\Phi_{N,B}(x))$ elements $n \le x$ with $P^-(n(Bn+1)) > N$. Hence, we know that

$$(h_1(n)n^{it_1}, h_2(Bn+1)n^{it_2}) \notin B_\varepsilon((z', w'))$$

for all but $O(\varepsilon^3\Phi_{N,B}(x))$ of those n. On the other hand by Proposition 2.3 there are $\gg \varepsilon^2/(kl)\Phi_{N,B}(x)$ numbers $n \le x$ on which $\max\{|n^{it_1} - z'|, |n^{it_2} - w'|\} \le \varepsilon/4$ and $h_1(n) = h_2(Bn+1) = 1$, whence

$$(h_1(n)n^{it_1}, h_2(Bn+1)n^{it_2}) = (n^{it}, n^{it'}) \in B_{\varepsilon/2}((z', w')).$$

This must intersect with the set of n where $(f_0(n), g_0(Bn+1)) = (h_1(n), h_2(Bn+1))$, provided that ε is sufficiently small (in terms of k and l alone). This contradicts the earlier claim, and proves the proposition. $\square$

4. The eventually rational case

In this section, we prove Proposition 1.5. That is, we assume that f and g are both eventually rational functions, i.e., for some N sufficiently large there is an $m \in \mathbb{N}$ such that for all primes $p > N$ we have $f(p)^m = 1$, and that $\{f(n)\}_n$ and $\{g(n)\}_n$ are both dense. Appealing to Proposition 2.6 once again, we may assume that f is $h_1 n^{it_1}$-pretentious and g is $h_2 n^{it_2}$-pretentious, k and l are minimal positive integers such that $h_1^k = \tilde{\chi}_1$ and $h_2^l = \tilde{\chi}_2$. As above, we let r and s denote the respective orders of $\tilde{\chi}_1$ and $\tilde{\chi}_2$. To prove Proposition 1.5, we need the following technical result, which extends Lemma 2.12 of [Klurman and Mangerel 2018] (which is the special case $M' = 1$ and $M'' = P_N := \prod_{p \le N} p$). In what follows, for a positive integer n we write $\mathrm{rad}(n) := \prod_{p \mid n} p$ to denote the squarefree kernel of n.

Lemma 4.1. *Let $N > 2q_1q_2$ be a positive integer, and let m, m', m'' be coprime squarefree positive integers such that $P_N = mm'm''$, with $(m', 2q_1q_2) = 1$. Let M' and M'' be integers with* $\mathrm{rad}(M') = m'$ *and* $\mathrm{rad}(M'') = m''$ *and M'' odd, and suppose M is a multiple of $2q_1q_2$ with* $\mathrm{rad}(M/2q_1q_2) = m$. *Then*

$$\sum_{\substack{n \le x \\ M \mid n}} \frac{1_{n \equiv -\overline{M}(M')} 1_{(n(n+1), M'')=1} 1_{h_1(n)=1} 1_{h_2(n+1)=1}}{n} = \left(\frac{1}{MM'klrs} \prod_{p \mid M''} \left(1 - \frac{2}{p}\right) + o(1)\right) \log x.$$

Proof. The proof of Lemma 4.1 is similar to that of Lemma 2.12 of [Klurman and Mangerel 2018], and we merely sketch the proof here. If $\mathcal{S}$ denotes the LHS above then, as M, M' and M'' are coprime, using orthogonality modulo M' gives

$$\begin{aligned}\mathcal{S} &= \frac{1}{krls} \sum_{\substack{n' \le x/M \\ n'M \equiv -1 (M')}} \frac{1_{(n'(Mn'+1), M'')=1} 1_{h_1(n'M)=1} 1_{h_2(n'M+1)=1}}{n'M} \left(\sum_{0 \le a \le kr-1} \sum_{0 \le b \le ls-1} h_1(n'M)^a h_2(n'M+1)^b \right) \\ &= \frac{1}{krls} \sum_{\chi(M')} \frac{\chi(M)}{\phi(M')} \sum_{0 \le a \le kr-1} h_1(M)^a \sum_{0 \le b \le ls-1} \sum_{n' \le x/M} \frac{\chi(n') h_1(n')^a h_2(Mn'+1)^b 1_{(n', M'')=1} 1_{(Mn'+1, M'')=1}}{n'M}.\end{aligned}$$

One proceeds as in the proof of Lemma 2.12 in [Klurman and Mangerel 2018] (using Theorem 2.1) to show that whenever $ab \ne 0$, the contributions to the full sum are $o(\log x)$, irrespective of χ. In the remaining case $a = b = 0$, and χ is *nonprincipal* modulo M', the contribution to the full sum is $o(\log x)$ as well. It thus follows that

$$\begin{aligned}\mathcal{S} &= \frac{1}{M\phi(M')krls} \sum_{n' \le x/M} \frac{1_{(n', M''M')=1} 1_{(Mn'+1, M'')=1}}{n'} + o(\log x) \\ &= \frac{1}{MM'krls} \prod_{p \mid M''} \left(1 - \frac{2}{p}\right) \log x + o(\log x),\end{aligned}$$

the main term arising from the coprimality of M and M'' provided that x is large enough in terms of M (using, e.g., the fundamental lemma of the sieve). This proves the claim. □

Proof of Proposition 1.5. Let N be such that for all $p > N$, we have $f(p)^k = g(p)^l = 1$. Assume for contradiction that $(f(n), g(n+1)) \notin B_\varepsilon(z, w)$, for some $\varepsilon > 0$ and some pair $(z, w) \in \mathbb{T}^2$. By Proposition 2.6, we know that $\mathbb{D}(f, h_1 n^{it}; x), \mathbb{D}(g, h_2 n^{it'}; x) \ll_\varepsilon 1$, for some h_1, h_2 completely multiplicative taking values in roots of unity. By multiplying the characters χ_1 and χ_2 corresponding to h_1 and h_2 by the principal character modulo P_N, we may assume that h_1 and h_2 are equal to 1 at all of the primes $p \le N$. Furthermore, as $f\overline{h_1}$ and $g\overline{h_2}$ also take values in roots of unity for all but finitely many primes, the proof of Lemma 3.1 implies that $t = t' = 0$. Now since $\{f(n)\}_n$ and $\{g(n)\}_n$ are dense, and $f(p)$, $g(p)$ take values in a set of bounded order roots of unity for all but finitely many primes, we know that there is (at least) a prime p and a prime p' at which the argument of $f(p) = e(\alpha)$ and $g(p') = e(\beta)$, with $\alpha, \beta \notin \mathbb{Q}$. The additional hypothesis of the proposition implies that, in fact, $p \ne p'$. With $\varepsilon > 0$ and $z, w \in \mathbb{T}$ given above, we may apply Kronecker's theorem (with α and with β, separately) to get $a, b \in \mathbb{N}$ such that

$$\|(f(p)^a, g(p')^b) - (z\overline{f(2q_1q_2)}, w)\|_{\ell_1} < \varepsilon.$$

Now set $B = 2q_1q_2p^a$. Thus, in the remainder of the proof, in order to achieve a contradiction it will suffice to check that we can find n such that $f(Bn) = f(2q_1q_2)f(p)^a$ and $g(nB+1) = g(p')^b$. Let $\eta > 0$ be a parameter to be chosen depending at most on ε, k, r, l and s. By Lemma 2.11 (see also Remark 2.12), there is a suitable infinite sequence of x (and a possibly larger choice of N) for which we have $f(n) = h_1(n)$ and $g(Bn+1) = h_2(Bn+1)$ for all but $O(\eta\Phi_{N,B}(x/B))$ integers $n \le x$ with $P^-(n(Bn+1)) > N$. Thus, it suffices to show that there are $\gg \varepsilon^2\Phi_{N,B}(x/B)$ integers $n \le x$ with $P^-(n(Bn+1)) > N$ that satisfy:

(i) $f(nB/(nB, P_N^\infty)) = 1 = g((nB+1)/(nB+1, P_N^\infty))$.

(ii) $(p')^b \parallel (nB+1)$ but $(nB(nB+1), P_N/pp') = 2q_1q_2$.

To do this, we shall apply Lemma 4.1. Let $m = p$, $m' = p'$ and $m'' = P_N/pp'$, and set $M = B$, $M' = (p')^b$ and $M'' = m''$. As $f(nB/(nB, P_N^\infty)) = h_1(nB)$ and $g((Bn+1)/(Bn+1, P_N^\infty)) = h_2(Bn+1)$ for all but $O(\varepsilon^3\Phi_{N,B}(x/B))$ choices of $n \le x/B$, Lemma 4.1 implies that the number of $n \le x$ with $B \mid n$ and satisfying the other required properties is

$$\gg \frac{1}{B(p')^{b-1}(p'-1)krls} \prod_{\substack{p'' \le N \\ p'' \nmid pp'B}} \left(1 - \frac{2}{p}\right) \gg \frac{1}{(p')^b krls}\Phi_{N,B}(x/B).$$

Choosing $\eta = C((p')^b krls)^{-1}$ with a sufficiently small constant $C = C(\varepsilon) > 0$ (noting that this quantity depends only on ε) implies the claim. $\square$

5. Some preliminary cases of Theorem 1.6

In this section, we shall dispose of several special cases of Theorem 1.6. In this way, we shall be able to reduce our work in proving Theorem 1.6 in Section 6 to focusing on functions with certain prescribed behavior. In this section, we shall assume that either $f(p)^k = p^{it}$ or $g(p)^l = p^{it'}$ for all *sufficiently large* primes p. Since the argument in the first case (i.e., with f) is symmetric to that with g, we shall focus only on the case with f. We consider three subcases of this case, according to the behavior of g:

subcase (i): We assume that $f(p)^k = p^{it}$ for all large primes p, and g is an eventually rational function (see Lemma 5.1).

subcase (ii): We assume that $f(p)^k = p^{it}$ holds *only* for large primes p and $g(p)^l = p^{it'}$ holds for large primes p (see Lemma 5.2 for a more precise formulation).

subcase (iii): We assume that $f(p)^k = p^{it}$ for *all* p, and $g(p)^l = p^{it'}$ *only* for large p (see the discussion following the proof of Lemma 5.2). In what follows, we recall that for each $n \in \mathbb{N}$, $f_0(n) := f(n)n^{-it/k}$ an $g_0(n) := g(n)n^{-it'/l}$.

Lemma 5.1. *Suppose there are positive integers N and k and a real t such that for all $p > N$, $f(p)^k = p^{it}$. Suppose moreover that g is eventually rational. Then $\{(f(n), g(n+1))\}_n$ is dense in $\mathbb{T}^2$.*

The same result holds with the roles of f and g reversed (and replacing $n \mapsto n+1$ in the analysis below by $n \mapsto n-1$).

Proof. By choosing N larger if necessary, and replacing k by some multiple of k, we may assume that $g(p)^k = 1$ for all $p > N$. Now suppose $(f(n-1), g(n)) \notin B_\varepsilon(z, w)$ for n sufficiently large. By Proposition 2.6, f and g are both pseudopretentious, and following the argument in the previous section, one can show that f is $h_1 n^{it/k}$-pretentious, with h_1 a pseudocharacter with conductor q_1, and similarly g is h_2-pretentious where h_2 is a pseudocharacter with conductor q_2.[11] We may assume that $t \neq 0$ here, otherwise f and g are both eventually rational, which is a case dealt with by Proposition 1.5, proven in the previous section. Since $\{g(n)\}_n$ is dense, but with rational argument for all but finitely many primes, there exists a prime p be such that $g(p) = e(\alpha)$ with $\alpha \notin \mathbb{Q}$. We now fix an even integer $B = 2q_1q_2p^r$, where the exponent r is chosen (via Kronecker's theorem) such that $g(p)^r = \overline{g(2q_1q_2)}w + O(\varepsilon^2)$. Now, if m is chosen so that $P^-(m(Bm-1)) > N$, it follows that for m sufficiently large,

$$f(Bm-1) = f_0(Bm-1)(Bm-1)^{it/k} = f_0(Bm-1)B^{it/k}m^{it/k} + O(\varepsilon^2).$$

Thus, it follows that

$$(f_0(Bm-1)m^{it/k}, g_0(m)) \notin B_{\varepsilon/2}(zB^{-it/k}, 1) \tag{27}$$

for all m sufficiently large with $P^-(m(Bm-1)) > N$. Now, by Lemma 3.2,[12] for all but $O(\varepsilon^3 \Phi_{N,B}(x))$ integers $m \leq x$ (with x chosen from appropriate infinite increasing sequence and N suitably large) with $P^-(m(Bm-1)) > N$ we have

$$(f_0(Bm-1), g_0(m)) = (h_1(Bm-1), h_2(m)) + O(\varepsilon^2).$$

[11] An application of Lemma 3.1, as in the previous section, implies that $g\bar{h}_2$, which takes finite roots of unity as values, is 1-pretentious.

[12] Strictly speaking, Lemmata 3.2 and 2.3 deal only with the pair of linear forms $(n \mapsto n, n \mapsto Bn+1)$; however, the same results can be derived with $(m \mapsto Bm-1, m \mapsto m)$ with minimal change to the proofs.

Moreover, by Proposition 2.3 there are $\gg \varepsilon^2 \Phi_{N,B}(x)$ such integers for which $(h_1(Bm-1), h_2(m)) = (1, 1)$, and such that $m^{it/k} = zB^{-it/k} + O(\varepsilon^2)$. Hence, it follows that

$$(f_0(Bm-1)m^{it/k}, g_0(m)) = (h_1(Bm-1)m^{it/k}, h_2(m)) = (zB^{-it/k}, 1) + O(\varepsilon^2),$$

for a positive upper density set of m. But this contradicts (27). The contradiction implies

$$\overline{\{f(n-1), g(n)\}_n} = \overline{\{f(n), g(n+1)\}_n} = \mathbb{T}^2,$$

as claimed. □

Lemma 5.2. *Suppose there are positive integers N, k, l and real numbers t, t' such that for all $p > N$, $f(p)^k = p^{it}$ and $g(p)^l = p^{it'}$, but for any $m \in \mathbb{N}$ and any $u \in \mathbb{R}$ there is an integer n such that $f(n)^m \neq n^{iu}$. Then $\{(f(n), g(n+1))\}_n$ is dense.*

Proof. If $t/t' \notin \mathbb{Q}$ then the proof is the same as the corresponding case in the proof of Proposition 1.7 (which only used data about integers n such that $n(Bn+1) > N$ for suitable B). Conversely, suppose $t = at'/b$ for some $a, b \in \mathbb{Z}$, with $b \neq 0$, and suppose there is an $\varepsilon > 0$ and a pair $(z, w) \in \mathbb{T}^2$ such that $(f(n), g(n+1)) \notin B_\varepsilon(z, w)$ for all n large. We may assume that $tt' \neq 0$, since otherwise f or g is eventually rational, and this was covered in the previous lemma. Furthermore, since we may replace k by kb and l by la, we may assume that $t = t'$. By hypothesis, we can choose a prime $p_0 \leq N$ for which $f(p_0)^k/p_0^{it}$ is not a root of unity of any order. Indeed, if every $p \leq N$ satisfied $f(p)^k/p^{it} \in \mu_m$ for some $m \geq 1$, we could replace k with km and t with tm so that $f(p)^{km} = p^{itm}$. Making these replacements for k and for t iteratively at every prime $p \leq N$ would imply that $f(n)^{k'} = n^{iu}$ for *all* n, where $k' \in \mathbb{N}$ and $u \in \mathbb{R}$. This contradicts our initial hypothesis regarding f. Thus, as per the above paragraph we can choose $p_0 \leq N$ such that $f(2p_0)(2p_0)^{-it/k} = e(\alpha)$, where $\alpha \notin \mathbb{Q}$. We may thus choose ℓ such that

$$(f(2p_0)(2p_0)^{-it})^\ell = e(\alpha\ell) = z\overline{w} + O(\varepsilon^2).$$

Now,[13] we pick m with $P^-(m((2p_0)^\ell m + 1)) > N$ and such that $m^{it} = w + O(\varepsilon^2)$. By assumption, it follows that

$$g((2p_0)^\ell m + 1) = ((2p_0)^\ell m + 1)^{it} = (2p_0)^{i\ell t} m^{it} + O(\varepsilon^2),$$

for sufficiently large m. Then, setting $n = (2p_0)^\ell m$ and assuming that m is sufficiently large, we have

$$\begin{aligned}(f(n), g(n+1)) &= (f(2p_0)^\ell m^{it}, (n+1)^{it}) \\ &= n^{it}((f(2p_0)(2p_0)^{-it})^\ell, 1) + O(\varepsilon^2) \\ &= w \cdot (z\overline{w}, 1) + O(\varepsilon^2) \\ &= (z, w) + O(\varepsilon^2),\end{aligned}$$

in contradiction to the claim. □

[13] This can be done, for example, by applying Proposition 2.3 with every pair of values of $h_1(n)$ and $h_2(Bn+1)$, then combining all of these contributions.

Turning to subcase (iii), it remains to consider the case that $f(n)^k = n^{it}$ for some $k \in \mathbb{N}$ and $t \in \mathbb{R}$, and $g(p)^l = p^{it'}$ for all $p > N$ but such that for each $m \in \mathbb{N}$ and $u \in \mathbb{R}$ there is $n \in \mathbb{N}$ with $g(n)^m \neq n^{iu}$. The argument is symmetric to that given in subcase (ii) (perhaps up to considering the sequence $\{f(Bn-1), g(n)\}_n$, for a suitable choice of B). We leave the details to the interested reader.

6. Proof of Theorem 1.6: The remaining cases

In this section, we prove Theorem 1.6, except in the exceptional cases that were dealt with in the previous section. By symmetry, we may assume that for one of the functions, say f, we have that for all $k \in \mathbb{N}$ and $t \in \mathbb{R}$ there are infinitely many primes p such that $f(p)^k \neq p^{it}$. First, suppose for contradiction that there is an $\varepsilon > 0$ and a point $(z, w) \in \mathbb{T}^2$ such that

$$(f(n), g(n+1)) \notin B_\varepsilon((z, w)) \text{ for all sufficiently large} n.$$

Applying Proposition 2.6, we know that f and g are, respectively, $h_1 n^{it}$- and $h_2 n^{it'}$-pseudopretentious, where the conductors of the pseudocharacters h_1 and h_2 are q_1 and q_2, respectively. We need two technical results in order to proceed in the proof of Theorem 1.6. The first will allow us to simultaneously control the angular distribution of the values of the irrational function f at prime powers p^m, as well as the angular distribution of p^{imt}, for $t \in \mathbb{R}$.

Lemma 6.1. *Let $f : \mathbb{N} \to \mathbb{T}$ be completely multiplicative function such that for all $t \in \mathbb{R}$ and $l \in \mathbb{N}$ there are infinitely many primes p such that $f(p)^l \neq p^{it}$. Let $z \in \mathbb{T}$, $u \in \mathbb{R}$ and $\delta, \eta \in (0, 1)$. Let $I \subset \mathbb{T}$ be the symmetric arc about 1 with length 2δ. Then for any $k, N \in \mathbb{N}$ there exists $n \in \mathbb{N}$ with $P^-(n) > N$ and $m \in \mathbb{N}$ such that the following hold*:

(i) $|f(n)^m - z| < \eta$.

(ii) $n^{ium} \in I$.

(iii) $k \mid m$.

Moreover if $u \neq 0$ then for any $w \in \mathbb{T}$ we can choose I to be a symmetric arc about w of length 2δ.

Proof. We consider two cases. First, suppose there is a prime $p > N$ for which $f(p)$ and p^{iu} are such that if $f(p) = e(\alpha)$ and $p^{iu} = e(\beta)$ then $\{1, \alpha, \beta\}$ are $\mathbb{Q}$-linearly independent (thus, $u \neq 0$ necessarily). Equivalently, $\{1, k\alpha, k\beta\}$ is $\mathbb{Q}$-linearly independent. In this case, we can apply Kronecker's theorem to find m_0 such that

$$\|(f(p)^{km_0}, p^{iukm_0}) - (z, w)\|_{\ell_1} \leq \min\{\eta, \delta\}^2,$$

and the claim follows with $m = km_0$ and $n = p$. Now, suppose that for all primes $p > N$ we have that $f(p)$ and p^{iu} have $\mathbb{Q}$-linearly dependent arguments. Equivalently, we can find a root of unity ξ_p such that $f(p) = \xi_p p^{iu}$ for each prime $p > N$. We consider two subcases of this case. First, if $\{\xi_p\}_{p>N}$ is a set of roots of unity of *bounded* order, say M, then it follows that $f(p)^{M!} = p^{iuM!}$ for all $p > N$. This contradicts our initial assumption that $f(p)^k \neq p^{it}$ for infinitely many p (with $k = M!$ and $t = uM!$).

Next, suppose that $\{\xi_p\}_{p>N}$ is such that for any $M \geq 1$ we can find a prime $p > N$ such that $\xi_p = e(a/b)$ with $b > Mk$ and $(a, b) = 1$. Pick $M > 2\eta^{-2}$, and choose $p = p(M)$ in this way. For $g, r \in \mathbb{N}$ parameters to be chosen, note that

$$f(p)^{k(gb+r)} = e(ark/b)p^{ik(gb+r)u}.$$

Writing $z = e(\gamma)$, we may select $0 \leq \ell \leq b-1$ such that $\gamma \in [k\ell/b, k(\ell+1)/b]$, whence

$$|\gamma - k\ell/b| \leq k/b \leq 1/M < \tfrac{1}{2}\eta^2.$$

We will thus pick r such that $ar \equiv \ell(b)$, so that $|f(p)^{k(gb+r)} - zp^{ik(gb+r)u}| < \eta^2/2$. Now, if $u = 0$ and $w = 1$ then we are done, as we can take $n = p$ and $m = kr$ (i.e., with $g = 0$). Otherwise, if $u \neq 0$ (and w is not necessarily 1), the sequence $\{p^{ikbug}\}_g$ is dense in $\mathbb{T}$. Thus, having already chosen r, it follows that we can pick g such that $|p^{ikgbu} - wp^{-ikru}| \leq \frac{1}{2}\min\{\eta^2, \delta^2\}$. Hence, we get that

$$\begin{aligned}|p^{ik(gb+r)u} - w| &= |p^{ikgbu} - wp^{-ikru}| \leq \delta^2,\\ |f(p)^{k(gb+r)} - z| &\leq |f(p)^{k(gb+r)} - zp^{-ik(gb+r)u}| + |z||p^{-ik(gb+r)u} - 1| \leq \eta^2.\end{aligned}$$

Thus, selecting $m = k(gb+r)$ and $n = p$ suffices to prove the claim in this case. □

Lemma 6.2. *Let $f, g : \mathbb{N} \to \mathbb{T}$ be completely multiplicative functions such that $\overline{\{f(n)\}_n} = \overline{\{g(n)\}_n} = \mathbb{T}$, but that $\overline{\{(f(n), g(n+1))\}_n} \neq \mathbb{T}^2$, and furthermore that f satisfies the hypotheses of Lemma 6.1. Suppose that f and g are, respectively, $h_1 n^{it}$- and $h_2 n^{it'}$-pretentious, where h_1 and h_2 are pseudocharacters of conductor q_1 and q_2, respectively, and $h_1^{kr} = h_2^{ls} = 1$. Finally, let $N, B \geq 1$, with $2q_1q_2 \mid B$, and let $u \in \mathbb{R}$ and $z \in \mathbb{T}$. Then for any $\eta > 0$ sufficiently small (in terms of u) there is a positive upper density subset $T_z = T_z(\eta)$ of $\mathcal{A}_{N,B,I}(h_1, h_2; 1, 1; u)$ on which $|f(n) - z| < \eta$, where I is the symmetric arc of length 2η about* 1.

In words, the lemma states that we can find a "large" set of integers n satisfying $P^-(n(Bn+1)) > N$ such that $f(n)$ is close to z, $h_1(n) = h_2(n) = 1$, and $n^{iu} \in I$.

Proof. Let $w \in \mathbb{T}$ be a parameter to be chosen (based on t and u). By Lemma 6.1, choose n_0 with $P^-(n_0) > N$ and m a multiple of kr such that $f(n_0)^m = z + O(\eta^2)$ and $(n_0)^{imu} = w + O(\eta^2)$, and set $B' := Bn_0^m$. Since f is pretentious to $h_1(n)n^{it}$, by Lemma 2.11 we can choose x from a suitable infinite sequence (and N slightly larger if necessary) such that for all but $O(\eta^3/(klrsn_0^m)\Phi_{N,B'}(x))$ integers $n \leq x$ with $P^-(n(B'n+1)) > N$ we have $f(n) = h_1(n)n^{it} + O(\eta^2)$. Now, furthermore, Proposition 2.3 gives

$$\left|\mathcal{A}_{N,B',I,I}(h_1, h_2; 1, 1; t, u) \cap \left[1, \frac{x}{n_0^m}\right]\right| \gg \frac{\eta^2}{klrs}\Phi_{N,B'}(x/n_0^m) \asymp \frac{\eta^2}{klrsn_0^m}\Phi_{N,B'}(x),$$

provided that $t/u \notin \mathbb{Q}$. Choosing $w = 1$ in this case, it follows that for all n in this set we have $f(nn_0^m) = zh(n)n^{it} + O(\eta^2) = z + O(\eta^2)$, and moreover $(nn_0^m)^{iu} = 1 + O(\eta^2)$. We define, in the case $t/u \notin \mathbb{Q}$, the set

$$T_z := \{nn_0^m : n \in \mathcal{A}_{N,B',I,I}(h_1, h_2; 1, 1; t, u)\}.$$

If, instead, $t/u = a/b$ and $a < b$ and $\eta^{-1/2} > b/a$ (without loss of generality), we can conclude that if $n^{it} = 1 + O(\eta^2 b/a)$ then $n^{iu} = \zeta_a^c + O(\eta^{3/2})$, where ζ_a is a fixed primitive a-th root of unity, and $0 \le c \le a-1$. By Proposition 2.3 once again, we have

$$\left|\mathcal{A}_{N,B,I'}(h_1,h_2;1,1;t) \cap \left[1, \frac{x}{n_0^m}\right]\right| \gg \frac{\eta^2}{klrsn_0^m}\Phi_{N,B'}(x),$$

where $I' \subseteq I$ is a symmetric subinterval of I about 1 of length η. By the pigeonhole principle, there is a choice $0 \le c_0 \le a-1$ such that $n^{iu} = \zeta_a^{c_0} + O(\eta^{3/2})$ for at least a proportion $\ge 1/a$ of these n. Choosing $w = \zeta_a^{-c_0}$ in this case, it follows that for all $n \in \mathcal{A}_{N,B,I'}(h_1,h_2;1,1;t)$ we have $f(nn_0^m) = f(n_0)^m h(n) n^{it} = 1 + O(\eta^2)$ and $(nn_0^m)^{iu} = 1 + O(\eta^{3/2})$. In this case, we define

$$T_z := \{nn_0^m : n \in \mathcal{A}_{N,B,I'}(h_1,h_2;1,1;t)\}.$$

We now observe that for any $n' \in T_z$ (regardless of the nature of t/u), $h_1(n') = h_1(n_0)^m h_1(n) = 1$, since $kr \mid m$ and $h^{kr} = 1$. Furthermore, we have $h_2(Bn'+1) = h_2(B'n+1) = 1$ by choice of B', and as we saw above, $f(n') = z + O(\eta^2)$ and $(n')^{iu} = 1 + O(\eta^{3/2})$ simultaneously. It follows that T_z has positive upper density, that $T_z \subseteq \mathcal{A}_{N,B,I}(h_1,h_2;1,1;u)$ (provided η is small enough) and that on T_z we have $f(n') = z + O(\eta^2)$. This implies the claim if η is small enough. □

Proof of Theorem 1.6, Part 2. We assume here that f is both irrational, and such that for each $k \in \mathbb{N}$ and $t \in \mathbb{R}$, there are infinitely many primes p for which $f(p)^k \neq p^{it}$; we can do this since the other cases where f is irrational were treated in the previous section. Now, suppose that $\{(f(n), g(n+1)\}_n$ stays outside of an arc of radius ε (in ℓ^1) about $(a,b) \in \mathbb{T}^2$. This means that $|f(n)-a| + |g(n+1)-b| \ge \varepsilon$ for all large n. By Proposition 2.6, there are minimal positive integers k, l, r and s, pseudocharacters h_1, h_2 with conductors $q_1, q_2 = O_\varepsilon(1)$ taking values in μ_{kr} and μ_{ls} respectively, and $t_1, t_2 \in \mathbb{R}$ such that $\mathbb{D}(f, h_1 n^{it_1}; x), \mathbb{D}(g, h_2 n^{it_2}; x) \ll_\varepsilon 1$. Let I be a symmetric interval of length ε^2 about 1, and let B be an integer of our choosing, chosen subject only to the constraint that $2q_1q_2 \mid B$. Consider those $n \in \mathcal{A}_{N,B,I}(h_1,h_2;1,1;t_1-t_2) =: T$. By Proposition 2.3, T has positive upper density. We have three possible scenarios:

(i) There is a positive upper density subset of T on which $|f(Bn)-a| \ge 3\varepsilon/4$ and $|g(Bn+1)-b| < \varepsilon/4$.

(ii) There is a positive upper density subset of T on which $|g(Bn+1)-b| \ge 3\varepsilon/4$ and $|f(Bn)-a| < \varepsilon/4$.

(iii) Except on an upper density zero subset of T we have $|f(Bn)-a| \ge \varepsilon/4$ and $|g(Bn+1)-b| \ge \varepsilon/4$.

Consider alternative (iii). It is clearly true that $|f(n) - a\overline{f(B)}| \ge \varepsilon/4$ for all but a zero upper density subset of $\mathbb{N}$. By Lemma 6.2 (taking $u = t_1 - t_2$ and $z = a\overline{f(B)}$ in the notation there), we can find a positive upper density subset $T_{a\overline{f(B)}} \subset T$ on which $|f(n) - a\overline{f(B)}| < \varepsilon/4$, contradicting the conditions of case (iii). Thus, (iii) can clearly not occur. Suppose next that we are in case (i). Then, writing

$$f(Bn)\overline{g(Bn+1)} - a\bar{b} = \overline{g(Bn+1)}(f(Bn)-a) + a\overline{(g(Bn+1)-b)}$$

and using the unimodularity of $g(Bn+1)$ and a together with the triangle inequality, we get that

$$|f(Bn)\overline{g(Bn+1)} - a\bar{b}| \geq |f(Bn) - a| - |g(Bn+1) - b| \geq \frac{\varepsilon}{2}$$

on a subset of positive upper density in T. The same condition occurs in case (ii) as well (by essentially the same argument). Now, put $F(n) := f(n)\overline{h_1(n)}n^{-it_1}$ and $G(n) := g(n)\overline{h_2(n)}n^{-it_2}$. We recall that for $n \in T$, $h_1(n) = h_2(n) = 1$ and $n^{i(t_2-t_1)} = 1 + O(\varepsilon^2)$. For each $n \in T$ sufficiently large we have

$$\begin{aligned} F(n)\overline{G(Bn+1)} &= f(n)\overline{g(Bn+1)}n^{-it_1}(Bn+1)^{it_2} \\ &= f(Bn)\overline{g(Bn+1)} \cdot n^{i(t_2-t_1)}\overline{f(B)}B^{it_2} + O(\varepsilon^2) \\ &= f(Bn)\overline{g(Bn+1)} \cdot \overline{f(B)}B^{it_2} + O(\varepsilon^2). \end{aligned}$$

It follows that on some positive upper density subset of T, we have

$$|F(n)\overline{G(Bn+1)} - \bar{b}a\overline{f(B)}B^{it_2}| \geq \frac{\varepsilon}{3},$$

for ε sufficiently small. At this point, we will select B using the following remarks. First, we are supposing that for any $k \in \mathbb{N}$ and $t \in \mathbb{R}$ there is a prime p such that $f(p)^k \neq p^{it}$. As such, for any $M \geq 1$ we can find P sufficiently large in terms of M such that $f(P)P^{-it_2} = e(\alpha)$ where either $\alpha \notin \mathbb{Q}$ or else $\alpha \in \mathbb{Q}$ with denominator at least M. Taking $M \asymp \varepsilon^{-2}$, we can choose r such that αr lies in the same $O(\varepsilon^2)$-length arc of $\mathbb{T}$ as the argument of $(2q_1q_2)^{it_2}\overline{baf(2q_1q_2)}$. Consequently, with this choice of r we find that

$$f(P)^r P^{-irt_2} = b(2q_1q_2)^{it_2}\overline{af(2q_1q_2)} + O(\varepsilon^2).$$

Finally, we choose $B = 2q_1q_2P^r$, and with this choice

$$f(B)B^{-it_2} = b\bar{a} + O(\varepsilon^2).$$

It then follows that

$$|F(n)\overline{G(Bn+1)} - 1| \geq \frac{\varepsilon}{4}$$

for each n belonging to a positive upper density subset of $T' \subseteq T$. Now, with N large we select by Szemerédi's theorem (see [Gowers 2001]) a long arithmetic progression $S \subset T'$; in particular, $|S| \to \infty$ as $x \to \infty$. As in the proof of Theorem 2.1 in [Klurman and Mangerel 2018], we have that on one hand

$$\begin{aligned} \tfrac{1}{16}\varepsilon^2|S \cap [1,x]| &\leq \sum_{\substack{n \leq x \\ n \in S}} |F(n)\overline{G}(Bn+1) - 1|^2 \\ &= 2|S \cap [1,x]|\biggl(1 - \operatorname{Re}\biggl(\frac{1}{|S \cap [1,x]|}\sum_{\substack{n \leq x \\ n \in S}} F(n)\overline{G(Bn+1)}\biggr)\biggr). \end{aligned}$$

Now write $S = \{a + jd : 0 \leq j \leq |S| - 1\}$, for some $d \in \mathbb{N}$ and $a \geq 0$. Let L_1 and L_2 denote the integer-valued linear forms $L_1(j) := a + jd$ and $L_2(j) := Ba + 1 + Bjd$, so that if $n \in S$ then there is a

$n' \in \mathbb{N}$ for which $n = L_1(n')$ and $Bn + 1 = L_2(n')$. For every prime p define

$$M_p(F(L_1), \overline{G}(L_2)) := \lim_{X\to\infty} \frac{1}{X} \sum_{\nu_1,\nu_2\geq 0} F(p)^{\nu_1}\overline{G}(p)^{\nu_2} \sum_{\substack{n\leq X \\ p^{\nu_1} \,\|\, L_1(n),\, p^{\nu_2} \,\|\, L_2(n)}} 1.$$

Theorem 1.3 of [Klurman 2017] (see also Remark 2.12 in [Klurman and Mangerel 2018]) shows that as $x \to \infty$ along a suitable infinite subsequence,

$$\sum_{\substack{n\leq x \\ n\in S}} F(n)\overline{G(Bn+1)}$$
$$= |S\cap[1,x]|\Big(\prod_{p\leq x} M_p(F(L_1), \overline{G(L_2)}) + O\Big(\mathbb{D}(F,1;N,Bx) + \mathbb{D}(G,1;N,Bx) + \frac{1}{N} + \frac{1}{\log\log x}\Big)\Big).$$

A simple calculation shows that $M_p(F(L_1), G(L_2)) = 0$ if $p \leq N$ and for $p > N$ we have

$$M_p(F(L_1), \overline{G(L_2)}) = \begin{cases} 1 + (F(p)-1)/p + (\overline{G(p)}-1)/p + O(1/p^2) & \text{if } p \nmid Bd, \\ 1 + (F(p)-1)/p + O(1/p^2) & \text{if } p \mid B,\ p \nmid d, \\ 1_{p\nmid a(Ba+1)} + F(p)^{\min\{\nu_p(a),\nu_p(d)\}} + \overline{G(p)}^{\min\{\nu_p(Ba+1),\nu_p(d)\}} & \text{if } p \mid d. \end{cases}$$

whence it follows that

$$\prod_{p\leq x} M_p(F(L_1), \overline{G(L_2)}) = 1 + O(\mathbb{D}(F,1;N,x)^2 + \mathbb{D}(G,1;N,x)^2 + 1/N).$$

Since F and G are both 1-pretentious then, provided x and N are large enough in terms of ε this yields the inequality

$$\tfrac{1}{16}\varepsilon^2|S\cap[1,x]| \ll \varepsilon^3|S\cap[1,x]|,$$

which is patently false when ε is sufficiently small. Note that we can argue in precisely the same way with the roles of f and g reversed to conclude that case (ii) cannot occur provided that $g(n)^l \neq n^{it'}$ for some t' (looking at the forms $(g(n), f(n-1))$ instead). This contradiction completes the proof of Theorem 1.6 in those cases not covered in the previous section. □

Acknowledgments

The authors are grateful to Andrew Granville, John Friedlander and Imre Kátai for their interest in the results of the present paper. The authors are also grateful to the referee for a very careful reading of the manuscript and numerous comments, corrections and suggestions that improved the quality of the paper. The bulk of this work was completed while the Mangerel was a Ph.D student at the University of Toronto. He would like to thank University of Toronto for excellent working conditions while this project was being completed.

References

[Brüdern and Fouvry 1996] J. Brüdern and É. Fouvry, "Le crible à vecteurs", *Compos. Math.* **102**:3 (1996), 337–355. MR Zbl

[Daróczy and Kátai 1989] Z. Daróczy and I. Kátai, "Characterization of additive functions with values in the circle group", *Publ. Math. Debrecen* **36**:1-4 (1989), 1–7. MR Zbl

[Gowers 2001] W. T. Gowers, "A new proof of Szemerédi's theorem", *Geom. Funct. Anal.* **11**:3 (2001), 465–588. MR Zbl

[Granville and Soundararajan 2007] A. Granville and K. Soundararajan, "Large character sums: pretentious characters and the Pólya–Vinogradov theorem", *J. Amer. Math. Soc.* **20**:2 (2007), 357–384. MR Zbl

[Kátai 1989] I. Kátai, "Characterization of arithmetical functions, problems and results", pp. 544–555 in *Théorie des nombres* (Quebec, 1987), edited by J.-M. De Koninck and C. Levesque, de Gruyter, Berlin, 1989. MR Zbl

[Klurman 2017] O. Klurman, "Correlations of multiplicative functions and applications", *Compos. Math.* **153**:8 (2017), 1622–1657. MR Zbl

[Klurman and Mangerel 2018] O. Klurman and A. P. Mangerel, "Rigidity theorems for multiplicative functions", *Math. Ann.* **372**:1-2 (2018), 651–697. MR Zbl

[Matomäki and Radziwiłł 2016] K. Matomäki and M. Radziwiłł, "Multiplicative functions in short intervals", *Ann. of Math.* (2) **183**:3 (2016), 1015–1056. MR Zbl

[Matomäki and Radziwiłł 2019] K. Matomäki and M. Radziwiłł, "Multiplicative functions in short intervals, and correlations of multiplicative functions", pp. 321–343 in *Proc. Int. Congr. Math., I: Plenary lectures* (Rio de Janeiro, 2018), edited by B. Sirakov et al., World Sci., Singapore, 2019.

[Matomäki et al. 2015] K. Matomäki, M. Radziwiłł, and T. Tao, "An averaged form of Chowla's conjecture", *Algebra Number Theory* **9**:9 (2015), 2167–2196. MR Zbl

[Montgomery 1994] H. L. Montgomery, *Ten lectures on the interface between analytic number theory and harmonic analysis*, CBMS Region. Conf. Series Math. **84**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Niederreiter and Philipp 1973] H. Niederreiter and W. Philipp, "Berry–Esseen bounds and a theorem of Erdős and Turán on uniform distribution mod 1", *Duke Math. J.* **40**:3 (1973), 633–649. MR Zbl

[Tao 2016] T. Tao, "The logarithmically averaged Chowla and Elliott conjectures for two-point correlations", *Forum Math. Pi* **4** (2016), art. id. e8. MR Zbl

[Tao and Teräväinen 2019] T. Tao and J. Teräväinen, "The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures", *Duke Math. J.* **168**:11 (2019), 1977–2027. MR Zbl

[Tenenbaum 1995] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Stud. Adv. Math. **46**, Cambridge Univ. Press, 1995. MR Zbl

Communicated by Roger Heath-Brown
Received 2019-01-24 Revised 2019-07-03 Accepted 2019-08-05

lklurman@gmail.com *Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden*

smangerel@gmail.com *Centre Recherches Mathematiques, Université de Montréal, Canada*

ALGEBRA AND NUMBER THEORY 14:1 (2020)
dx.doi.org/10.2140/ant.2020.14.191

Birationally superrigid Fano 3-folds of codimension 4

Takuzo Okada

We determine birational superrigidity for a quasismooth prime Fano 3-fold of codimension 4 with no projection centers. In particular we prove birational superrigidity for Fano 3-folds of codimension 4 with no projection centers which were recently constructed by Coughlan and Ducat. We also pose some questions and a conjecture regarding the classification of birationally superrigid Fano 3-folds.

1. Introduction

A *prime Fano* 3*-fold* is a normal projective $\mathbb{Q}$-factorial 3-fold X with only terminal singularities such that $-K_X$ is ample and the class group $\mathrm{Cl}(X) \cong \mathbb{Z}$ is generated by $-K_X$. To such X there corresponds the anticanonical graded ring

$$R(X, -K_X) = \bigoplus_{m \in \mathbb{Z}_{\geq 0}} H^0(X, -mK_X),$$

and by choosing minimal generators we can embed X into a weighted projective space. By the *codimension* of X we mean the codimension of X in the weighted projective space. Based on analysis by Altınok, Brown, Iano-Fletcher, Kasprzyk, Prokhorov, Reid, and others (see for example [Altınok et al. 2002]) there is a database [Brown and Kasprzyk 2009] of numerical data (such as Hilbert series) coming from graded rings that can be the anticanonical graded ring of a prime Fano 3-fold. Currently it is not a classification, but it serves as a list, meaning that the anticanonical graded ring of a prime Fano 3-fold appears in the database.

The database contains a huge number of candidates, which suggests difficulty in the biregular classification of Fano 3-folds. The aim of this paper is to shed light on the classification of birationally superrigid Fano 3-folds. Here, a Fano 3-fold of Picard number 1 is said to be *birationally superrigid* if any birational map to a Mori fiber space is biregular. We remark that in [Ahmadinezhad and Okada 2018] a possible approach to achieving birational classification of Fano 3-folds is suggested by introducing notion of *solid Fano* 3*-folds*, which are Fano 3-folds not birational to either a conic bundles or a del Pezzo fibration.

Up to codimension 3, we have satisfactory results on the classification of quasismooth prime Fano 3-folds: the classification is complete in codimensions 1 and 2 [Iano-Fletcher 2000; Chen et al. 2011; Altınok 1998] and in codimension 3 the existence is known for all 70 numerical data in the database. Moreover birational superrigidity of quasismooth prime Fano 3-folds of codimension at most 3 has been well studied (see [Iskovskikh and Manin 1971; Corti et al. 2000; Cheltsov and Park 2017; Okada 2014a;

MSC2010: primary 14J45; secondary 14E07, 14E08.
Keywords: Fano variety, Birational rigidity.

Ahmadinezhad and Zucconi 2016; Ahmadinezhad and Okada 2018] and [Okada 2014b, 2018] for solid cases in codimension 2).

For quasismooth prime Fano 3-folds of codimension 4, there are 145 candidates of numerical data in [Brown and Kasprzyk 2009]. In [Brown et al. 2012], existence for 116 data is proved, where the construction is given by birationally modifying a known variety. This process is called unprojection and, as a consequence, a constructed Fano 3-fold corresponding to each of the 116 families admits a Sarkisov link to a Mori fiber space, hence it is not birationally superrigid. The 116 families of Fano 3-folds are characterized as those that possess a singular point which is so called a type I projection center (see [Brown et al. 2012] for details). There are other types of projection centers (such as types $\mathrm{II}_1, \dots, \mathrm{II}_7$, IV according to the database [Brown and Kasprzyk 2009]). Through the known results in codimensions 1, 2 and 3, we can expect that the existence of a projection center violates birational superrigidity. Therefore it is natural to consider prime Fano 3-folds without projection centers for the classification of birational superrigid Fano 3-folds (see also the discussion in Section 5).

According to the database [Brown and Kasprzyk 2009], there are 5 candidates of quasismooth prime Fano 3-folds of codimension 4 with no projection centers. Those are identified by database numbers #25, #166, #282, #308 and #29374. Among them, #29374 corresponds to smooth prime Fano 3-folds of degree 10 embedded in $\mathbb{P}^7$, and it is proved in [Debarre et al. 2012] that they are not birationally superrigid (not even birationally rigid, a weaker notion than superrigidity). Recently Coughlan and Ducat [2018] constructed many prime Fano 3-folds including those corresponding to #25 and #282 and we sometimes refer to these varieties as *cluster Fano* 3*-folds*. There are two constructions, $\mathsf{G}_2^{(4)}$ and C_2 formats (see [Coughlan and Ducat 2018, Section 5.6] for details and see page 205 for concrete descriptions) for #282 and they are likely to sit in different components of the Hilbert scheme.

Theorem 1.1. *Let X be a quasismooth prime Fano* 3*-fold of codimension* 4 *and of numerical type* #282 *which is constructed in either $\mathsf{G}_2^{(4)}$ format or C_2 format. If X is constructed in C_2 format*, *then we assume that X is general. Then X is birationally superrigid.*

For the remaining three candidates #25, #166 and #308, we can prove birational superrigidity in a stronger manner; we are able to prove birational superrigidity for these 3 candidates by utilizing only numerical data. Here, by numerical data for a candidate Fano 3-fold X, we mean the weights of the weighted projective space, degrees of the defining equations, the anticanonical degree $(-K_X)^3$ and the basket of singularities of X (see Section 3). Note that we do not know the existence of Fano 3-folds for #166 and #308.

Theorem 1.2. *Let X be a well-formed quasismooth prime Fano* 3*-fold of codimension* 4 *and of numerical type* #25, #166 *or* #308. *Then X is birationally superrigid.*

2. Birational superrigidity

Basic properties. Throughout this subsection we assume that X is a Fano 3-fold of Picard number 1, that is, X is a normal projective $\mathbb{Q}$-factorial 3-fold such that X has only terminal singularities, $-K_X$ is ample and $\operatorname{rank}\operatorname{Pic}(X) = 1$.

Definition 2.1. We say that X is *birationally superrigid* if any birational map $\sigma : X \dashrightarrow Y$ to a Mori fiber space $Y \to T$ is biregular.

By an *extremal divisorial extraction* $\varphi \colon (E \subset Y) \to (\Gamma \subset X)$, we mean an extremal divisorial contraction $\varphi \colon Y \to X$ from a normal projective $\mathbb{Q}$-factorial variety Y with only terminal singularities such that E is the φ-exceptional divisor and $\Gamma = \varphi(E)$.

Definition 2.2. Let $\mathcal{H} \sim_{\mathbb{Q}} -nK_X$ be a movable linear system, where n is a positive integer. A *maximal singularity* of $\mathcal{H}$ is an extremal extraction $\varphi \colon (E \subset Y) \to (\Gamma \subset X)$ such that

$$c(X, \mathcal{H}) = \frac{a_E(K_X)}{m_E(\mathcal{H})} < \frac{1}{n},$$

where

- $c(X, \mathcal{H}) := \max\{\lambda \mid K_X + \lambda\mathcal{H} \text{ is canonical}\}$ is the *canonical threshold* of $(X, \mathcal{H})$,
- $a_E(K_X)$ is the discrepancy of K_X along E, and
- $m_E(\mathcal{H})$ is the multiplicity along E of the proper transform of $\mathcal{H}$.

We say that an extremal divisorial extraction is a *maximal singularity* if there exists a movable linear system $\mathcal{H}$ such that the extraction is a maximal singularity of $\mathcal{H}$. A subvariety $\Gamma \subset X$ is called a *maximal center* if there is a maximal singularity $Y \to X$ whose center is Γ.

The following is the fundamental theorem in the study of birational superrigidity, which emerged in [Iskovskikh and Manin 1971] and has been simplified and generalized in [Pukhlikov 1998; Corti 1995].

Theorem 2.3 [Corti 1995, Theorem 4.10 and Proposition 2.10]. *If X admits no maximal center, then X is birationally superrigid.*

For the proof of birational superrigidity of a given Fano 3-fold X of Picard number 1 we need to exclude each subvariety of X as a maximal center. In the next subsection we will explain several methods of exclusion under a relatively concrete setting. Here we discuss methods of excluding terminal quotient singular points in a general setting.

For a terminal quotient singular point $\mathsf{p} \in X$ of type $\frac{1}{r}(1, a, r-a)$, where r is coprime to a and $0 < a < r$, there is a unique extremal divisorial extraction $\varphi \colon (E \subset Y) \to (\mathsf{p} \in X)$, which is the weighted blowup with weight $\frac{1}{r}(1, a, r-a)$, and we call it the *Kawamata blowup* (see [Kawamata 1996] for details). The integer $r > 1$ is called the *index* of $\mathsf{p} \in X$. For the Kawamata blowup $\varphi \colon (E \subset Y) \to (\mathsf{p} \in X)$, we have $K_Y = \varphi^* K_X + \frac{1}{r} E$ and

$$(E^3) = \frac{r^2}{a(r-a)}.$$

For a divisor D on X, the *order* of D along E, denoted by $\mathrm{ord}_E(D)$, is defined to be the coefficient of E in $\varphi^* D$.

We first explain the most basic method.

Lemma 2.4 [Corti et al. 2000, Lemma 5.2.1]. *Let* $\mathsf{p} \in X$ *a terminal quotient singular point and* $\varphi \colon (E \subset Y) \to (\mathsf{p} \in X)$ *the Kawamata blowup. If* $(-K_Y)^2 \notin \operatorname{Int} \overline{\mathrm{NE}}(Y)$, *then* p *is not a maximal center.*

For the application of the above lemma, we need to find a nef divisor on Y. The following result, which is a slight generalization of [Okada 2018, Lemma 6.6], is useful.

Lemma 2.5. *Let* $\mathsf{p} \in X$ *be a terminal quotient singular point and* $\varphi \colon (E \subset Y) \to (\mathsf{p} \in X)$ *the Kawamata blowup. Assume that there are effective Weil divisors* $D_1, \dots, D_k$ *such that the intersection* $\operatorname{Supp}(D_1) \cap \cdots \cap \operatorname{Supp}(D_k)$ *does not contain a curve through* p. *We set*

$$e := \min \left\{ \frac{\operatorname{ord}_E(D_i)}{n_i} \mid 1 \le i \le k \right\},$$

where n_i *is the positive rational number such that* $D_i \sim_{\mathbb{Q}} -n_i K_X$. *Then* $-\varphi^* K_X - \lambda E$ *is a nef divisor for* $0 \le \lambda \le e$.

Proof. We may assume $e > 0$, that is, $\operatorname{Supp}(D_i)$ passes through p for any i. For an effective divisor $D \sim_{\mathbb{Q}} -nK_X$, we call $\operatorname{ord}_E(D)/n$ the vanishing ratio of D along E. For $1 \le i \le k$, we choose a component of D_i, denoted D'_i, which has maximal vanishing ratio along E among the components of D_i. Clearly $D'_1 \cap \cdots \cap D'_k$ does not contain a curve through p and we have

$$e' := \min \left\{ \frac{\operatorname{ord}_E(D'_i)}{n'_i} \mid 1 \le i \le k \right\} \ge e,$$

where $n'_i \in \mathbb{Q}$ is such that $D'_i \sim_{\mathbb{Q}} -n'_i K_X$. Since $D'_1, \dots, D'_k$ are prime divisors, we can apply [Okada 2018, Lemma 6.6] and conclude that $-\varphi^* K_X - e'E$ is nef. Then so is $-\varphi^* K_X - \lambda E$ for any $0 \le \lambda \le e'$ since $-\varphi^* K_X$ is nef, and the proof is completed. □

We have another method of exclusion which can sometimes be effective when Lemma 2.4 is not applicable.

Lemma 2.6. *Let* $\mathsf{p} \in X$ *be a terminal quotient singular point and* $\varphi \colon (E \subset Y) \to (\mathsf{p} \in X)$ *the Kawamata blowup. Suppose that there exists an effective divisor* S *on* X *passing through* p *and a linear system* $\mathcal{L}$ *of divisors on* X *passing through* p *with the following properties*:

(1) $\operatorname{Supp}(S) \cap \operatorname{Bs} \mathcal{L}$ *does not contain a curve passing through* p.

(2) *For a general member* $L \in \mathcal{L}$, *we have*

$$(-K_Y \cdot \tilde{S} \cdot \tilde{L}) \le 0,$$

where $\tilde{S}$, $\tilde{L}$ *are the proper transforms of* S, L *on* Y, *respectively.*

Then p *is not a maximal center.*

Proof. According to [Okada 2018, Lemma 2.20], it suffices to show that there exist infinitely many distinct curves on Y which intersect $-K_Y$ nonpositively and E positively. For a curve or a divisor Δ on X, we denote by $\tilde{\Delta}$ its proper transform on Y.

We write $L \sim -nK_X$. Write $S = \sum m_i S_i + T$, where $m_i > 0$, S_i is a prime divisor and T is an effective divisor which does not pass through p. We have

$$(-K_Y \cdot \tilde{T} \cdot \tilde{L}) = nl(-K_X)^3 \geq 0,$$

where $T \sim -lK_X$ for some $l \geq 0$. Since

$$0 \geq (-K_Y \cdot \tilde{S} \cdot \tilde{L}) = \sum m_i(-K_Y \cdot \tilde{S}_i \cdot \tilde{L}) + (-K_Y \cdot \tilde{T} \cdot \tilde{L}),$$

there is a component S_i for which $(-K_Y \cdot \tilde{S}_i \cdot \tilde{L}) \leq 0$. Since $\mathsf{p} \in S_i \cap \operatorname{Bs} \mathcal{L} \subset \operatorname{Supp}(S) \cap \operatorname{Bs} \mathcal{L}$, we may assume that S is a prime divisor by replacing S by S_i.

Write $\mathcal{L} = \{L_\lambda \mid \lambda \in \mathbb{P}^l\}$. For $\lambda \in \mathbb{P}^l$, we write $S \cdot L_\lambda = \sum_i c_i C_{\lambda,i}$, where $c_i \geq 0$ and $C_{\lambda,i}$ is an irreducible and reduced curve on X. Then,

$$\tilde{S} \cdot \tilde{L}_\lambda = \sum_i c_i \tilde{C}_{\lambda,i} + \Xi,$$

where Ξ is an effective 1-cycle supported on E. Since any component of Ξ is contracted by φ and $-K_Y$ is φ-ample, we have $(-K_Y \cdot \Xi) \geq 0$. Thus, for a general $\lambda \in \mathbb{P}^l$, we have

$$0 \geq (-K_Y \cdot \tilde{S} \cdot \tilde{L}_\lambda) \geq \sum_i c_i(-K_Y \cdot \tilde{C}_{\lambda,i}).$$

It follows that $(-K_Y \cdot \tilde{C}_{\lambda,i}) \leq 0$ for some i. We choose such a $\tilde{C}_{\lambda,i}$ and denote it by $\tilde{C}^\circ_\lambda$. By assumption (1) the set

$$\{\tilde{C}^\circ_\lambda \mid \lambda \in \mathbb{P}^l \text{ is general}\}$$

consists of infinitely many distinct curves. We have $(-K_Y \cdot \tilde{C}^\circ_\lambda) \leq 0$ by the construction. We see that $(E \cdot \tilde{C}^\circ_\lambda) > 0$ since $\tilde{C}^\circ_\lambda$ is the proper transform of a curve passing through p. Therefore p is not a maximal center by [Okada 2018, Lemma 2.20]. □

Fano varieties in a weighted projective space. Let X be a prime Fano 3-fold. As in the introduction, we choose minimal generators of the anticanonical ring $R(X, -K_X)$ and let $X \subset \mathbb{P} := \mathbb{P}(a_0, \dots, a_n)$ be the corresponding embedding. We say that $X \subset \mathbb{P}$ is *anticanonically embedded*. We denote by $x_0, \dots, x_n$ the homogeneous coordinates of $\mathbb{P}$ with $\deg x_i = a_i$. Let

$$F_1 = F_2 = \cdots = F_N = 0$$

be defining equations of X inside $\mathbb{P}$, where $F_i \in \mathbb{C}[x_0, \dots, x_n]$ is a homogeneous polynomial of degree d_i with respect to the grading $\deg x_i = a_i$. We assume that $\mathbb{P}$ is *well-formed*, that is,

$$\gcd\{a_i \mid 0 \leq i \leq n, i \neq j\} = 1$$

for $j = 0, 1, \dots, n$. Note that X is not contained in a linear cone (i.e., a smaller weighted projective space in $\mathbb{P}$) by minimality of generators of $R(X, -K_X)$.

Definition 2.7. We say that X is *well-formed* if $\operatorname{codim}_X(X \cap \operatorname{Sing}(\mathbb{P})) \geq 2$. We say that X is *quasismooth* if the affine cone

$$(F_1 = F_2 = \cdots = F_N = 0) \subset \mathbb{A}^{n+1} = \operatorname{Spec} \mathbb{C}[x_0, \ldots, x_n]$$

is smooth outside the origin.

Remark 2.8. The description of $\operatorname{Sing}(\mathbb{P})$ is given in Remark 2.12 below. Under the assumption that $X \subset \mathbb{P}$ is an anticanonically embedded quasismooth prime Fano 3-fold, we believe that well-formedness is a very mild condition (or perhaps it is automatically satisfied). For example, a quasismooth weighted complete intersection $X \subset \mathbb{P}$ which is not contained in a linear cone is well-formed (see [Iano-Fletcher 2000, Theorem 6.17]).

In the following we assume that $X \subset \mathbb{P}$ is well-formed and quasismooth. For $0 \leq i \leq n$, we define $\mathsf{p}_{x_i} = (0 : \cdots : 1 : \cdots : 0) \in \mathbb{P}$, where the unique 1 is in the $(i+1)$-th position, and we define $D_i = (x_i = 0) \cap X$ which is a Weil divisor such that $D_i \sim -a_i K_X$.

Lemma 2.9. *If $(-K_X)^3 \leq 1$, then no curve on X is a maximal center.*

Proof. The same proof of [Ahmadinezhad and Okada 2018, Lemma 2.1] applies in this setting without any change. ☐

Lemma 2.10. *Assume that $a_0 \leq a_1 \leq \ldots \leq a_n$. If $a_{n-1} a_n (-K_X)^3 \leq 4$, then no nonsingular point of X is a maximal center.*

Proof. The proof is almost identical to that of [Ahmadinezhad and Okada 2018, Lemma 2.6]. ☐

Definition 2.11. Let $\mathcal{C} \subset \{x_0, \ldots, x_n\}$ be a nonempty set of homogeneous coordinates. We define

$$\Pi(\mathcal{C}) := \bigcap_{z \in \mathcal{C}} (z = 0) \subset \mathbb{P}, \quad \Pi_X(\mathcal{C}) := \Pi(\mathcal{C}) \cap X \subset X.$$

Sometimes we denote

$$\Pi(\mathcal{C}) = \Pi(x_{i_1}, \ldots, x_{i_m}), \quad \Pi_X(\mathcal{C}) = \Pi_X(x_{i_1}, \ldots, x_{i_m}),$$

when $\mathcal{C} = \{x_{i_1}, \ldots, x_{i_m}\}$. We also define

$$\gcd(\mathcal{C}) := \gcd\{\deg(x_i) \mid x_i \in \mathcal{C}\}.$$

Remark 2.12. We explain some consequences of well-formedness and quasismoothness, which will be frequently used. We keep the above notation and assumptions:

(1) For the singular locus, we have $\operatorname{Sing}(X) = X \cap \operatorname{Sing}(\mathbb{P})$. For the proof see [Dimca 1986, Proposition 8]. Note that $X \subset \mathbb{P}$ is additionally assumed to be a (weighted) complete intersection in [loc. cit.] but the same proof applies.

(2) The singular locus of $\mathbb{P}$ can be described as follows:

$$\mathrm{Sing}(\mathbb{P}) = \bigcup_{\substack{\mathcal{C} \subsetneq \{x_0, \dots, x_n\} \\ \gcd(\{x_0,\dots,x_n\} \setminus \mathcal{C}) > 1}} \Pi(\mathcal{C}).$$

By (1), we have

$$\mathrm{Sing}(X) = \bigcup_{\substack{\mathcal{C} \subsetneq \{x_0, \dots, x_n\} \\ \gcd(\{x_0,\dots,x_n\} \setminus \mathcal{C}) > 1}} \Pi_X(\mathcal{C}).$$

(3) For $\mathcal{C} \subset \{x_0, \dots, x_n\}$, we define

$$\Pi_X^*(\mathcal{C}) := \Pi_X(\mathcal{C}) \cap \Bigg(\bigcap_{z \in \{x_0, \dots, x_n\} \setminus \mathcal{C}} (z \neq 0) \Bigg).$$

Let $\mathcal{C} \subset \{x_0, \dots, x_n\}$ be a subset such that $r := \gcd(\{x_0, \dots, x_n\} \setminus \mathcal{C}) > 1$. Then any point $\mathsf{p} \in X$ which is contained in $\Pi_X^*(\mathcal{C})$ is a cyclic quotient singular point of index r and hence any point $\mathsf{p} \in X$ which is contained in $\Pi_X(\mathcal{C})$ is a cyclic quotient singular point of index divisible by r.

Lemma 2.13. *Let* $\mathsf{p} \in X$ *be a singular point of type* $\frac{1}{2}(1, 1, 1)$ *and let*

$$b := \max\{a_i \mid 0 \le i \le n, a_i \text{ is odd}\}.$$

If $2b(-K_X)^3 \le 1$, *then* p *is not a maximal center.*

Proof. Let $\mathcal{C} = \{x_{i_1}, \dots, x_{i_m}\}$ be the set of homogeneous coordinates of odd degree. The set $\Pi_X(\mathcal{C}) = D_{i_1} \cap \dots \cap D_{i_m}$ consists of singular points by Remark 2.12. In particular $\Pi_X(\mathcal{C})$ is a finite set of points since X has only terminal singular points. Let $\varphi \colon (E \subset Y) \to (\mathsf{p} \in X)$ be the Kawamata blowup. Then $\mathrm{ord}_E(D_{i_j}) \ge \frac{1}{2}$ since $2D_{i_j}$ is a Cartier divisor passing through p and thus $-b\varphi^* K_X - \frac{1}{2}E$ is nef by Lemma 2.5. We have

$$\left(-b\varphi^* K_X - \tfrac{1}{2}E\right) \cdot (-K_Y)^2 = b(-K_X)^3 - \tfrac{1}{2} \le 0.$$

This shows that $(-K_Y)^2 \notin \mathrm{Int}\,\overline{\mathrm{NE}}(Y)$ and p is not a maximal center by Lemma 2.4. □

Definition 2.14. Let $\mathsf{p} = \mathsf{p}_{x_k} \in X$ be a terminal quotient singular point of type $\frac{1}{a_k}(1, c, a_k - c)$ for some c with $1 \le c \le \frac{1}{2} a_k$. For a nonempty subset $\mathcal{C} = \{x_{i_1}, \dots, x_{i_m}\} \subset \{x_0, \dots, x_n\}$, we define

$$\mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) := \min_{1 \le j \le m} \left\{ \frac{\bar{a}_{i_j}}{a_{i_j} a_k} \right\},$$

where $\bar{a}_{i_j}$ is the integer such that $1 \le \bar{a}_{i_j} \le a_k$ and $\bar{a}_{i_j}$ is congruent to a_{i_j} modulo a_k, and call it the *initial vanishing ratio* of $\mathcal{C}$ at p.

Definition 2.15. For a terminal quotient singularity p of type $\frac{1}{r}(1, a, r - a)$, we define

$$\mathrm{wp}(\mathsf{p}) := a(r - a),$$

and call it the *weight product* of p.

Lemma 2.16. *Let* $\mathsf{p} = \mathsf{p}_{x_k} \in X$ *be a terminal quotient singular point. Suppose that there exists a subset* $\mathcal{C} \subset \{x_0, \dots, x_n\}$ *satisfying the following properties*:

(1) $\mathsf{p} \in \Pi_X(\mathcal{C})$, *or equivalently* $x_k \notin \mathcal{C}$.

(2) $\Pi_X(\mathcal{C} \cup \{x_k\}) = \varnothing$.

(3) $\mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) \geq \mathrm{wp}(\mathsf{p})(-K_X)^3$.

Then p *is not a maximal center.*

Proof. We write $\mathcal{C} = \{x_{i_1}, \dots, x_{i_m}\}$. We claim that $\Pi_X(\mathcal{C}) = D_{i_1} \cap \cdots \cap D_{i_m}$ is a finite set of points. Indeed, if $\Pi_X(\mathcal{C})$ contains a curve, then $\Pi_X(\mathcal{C} \cup \{x_k\}) = \Pi_X(\mathcal{C}) \cap D_k \neq \varnothing$ since D_k is an ample divisor on X. This is impossible by the assumption (2). Note that we have $\mathrm{ord}_E(D_{i_j}) \geq \bar{a}_{i_j}/a_k$ (see [Ahmadinezhad and Okada 2018, Section 3]) so that

$$e := \min\left\{ \frac{\mathrm{ord}_E(D_{i_j})}{a_{i_j}} \mid 1 \leq j \leq m \right\} \geq \mathrm{ivr}_{\mathsf{p}}(\mathcal{C}).$$

By Lemma 2.5, $-\varphi^* K_X - \mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) E$ is nef and we have

$$(-\varphi^* K_X - \mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) E)(-K_Y)^2 = (-K_X)^3 - \frac{\mathrm{ivr}_{\mathsf{p}}(\mathcal{C})}{\mathrm{wp}(\mathsf{p})} \leq 0$$

by the assumption (3). Therefore $(-K_Y)^2 \notin \mathrm{Int}\, \overline{\mathrm{NE}}(Y)$ and p is not a maximal center. $\square$

Let $\mathsf{p} \in X$ be a singular point such that it can be transformed to p_{x_k} by a change of coordinates. For simplicity of the description we assume $\mathsf{p} = \mathsf{p}_{x_0}$ and we set $r = a_0 > 1$. Let $\varphi\colon (E \subset Y) \to (\mathsf{p} \in X)$ be the Kawamata blowup. We explain a systematic way to estimate $\mathrm{ord}_E(D_i)$ for coordinates x_i and also an explicit description of φ. It is a consequence of the quasismoothness of X that after renumbering the defining equation we can write

$$F_l = \alpha_l x_0^{m_l} x_{i_l} + (\text{other terms}) \quad \text{for } 1 \leq l \leq n-3,$$

where $\alpha_l \in \mathbb{C} \setminus \{0\}$, m_l is a positive integer, $x_0, x_{i_1}, \dots, x_{i_{n-3}}$ are mutually distinct so that by denoting the other 3 coordinates as $x_{j_1}, x_{j_2}, x_{j_3}$ we have

$$\{x_0, x_{i_1}, \dots, x_{i_{n-3}}, x_{j_1}, x_{j_2}, x_{j_3}\} = \{x_0, \dots, x_n\},$$

and we can choose $x_{j_1}, x_{j_2}, x_{j_3}$ as local orbi-coordinates of X at p. In this case the singular point p is of type

$$\tfrac{1}{r}(a_{j_1}, a_{j_2}, a_{j_3}).$$

Definition 2.17 [Ahmadinezhad and Okada 2018, Definitions 3.6 and 3.7]. For an integer a, we denote by $\bar{a}$ the positive integer such that $\bar{a} \equiv a \pmod r$ and $0 < \bar{a} \leq r$. We say that

$$\mathbf{w}(x_1, \dots, x_n) = \tfrac{1}{r}(b_1, \dots, b_n)$$

is an *admissible weight* at p if $b_i \equiv a_i \pmod r$ for any i.

For an admissible weight $\mathbf{w}$ at p and a polynomial $f = f(x_0, \dots, x_n)$, we denote by $f^{\mathbf{w}}$ the lowest weight part of f, where we assume that $\mathbf{w}(x_0) = 0$.

We say that an admissible weight $\mathbf{w}$ at p satisfies the *KBL condition* if $x_0^{m_l} x_{i_l} \in F_l^{\mathbf{w}}$ for $1 \le l \le n-3$ and

$$(b_{j_1}, b_{j_2}, b_{j_3}) = (\bar{a}_{j_1}, \bar{a}_{j_2}, \bar{a}_{j_3}).$$

Let $\mathbf{w}(x_1, \dots, x_n) = \frac{1}{r}(b_1, \dots, b_n)$ be an admissible weight at p satisfying the KBL condition. We denote by $\Phi_{\mathbf{w}} \colon Q_{\mathbf{w}} \to \mathbb{P}$ the weighted blowup at p with weight $\mathbf{w}$, and by $Y_{\mathbf{w}}$ the proper transform of X via $\Phi_{\mathbf{w}}$. Then the induced morphism $\varphi_{\mathbf{w}} = \Phi_{\mathbf{w}}|_{Y_{\mathbf{w}}} \colon Y_{\mathbf{w}} \to X$ coincides with the Kawamata blowup at p. From this we see that the exceptional divisor E is isomorphic to

$$E_{\mathbf{w}} := (g_1 = \dots = g_{n-3} = 0) \subset \mathbb{P}(b_1, \dots, b_n),$$

where $g_l = F_l^{\mathbf{w}}(1, x_1, \dots, x_n)$. Note that the KBL condition implies that the equations defining $E_{\mathbf{w}}$ cut out a copy of $\mathbb{P}(b_{j_1}, b_{j_2}, b_{j_3})$. We refer readers to [Ahmadinezhad and Okada 2018, Section 3] for details.

Lemma 2.18 [Ahmadinezhad and Okada 2018, Lemma 3.9]. *Let $\mathbf{w}(x_1, \dots, x_n) = \frac{1}{r}(b_1, \dots, b_n)$ be an admissible weight at $\mathsf{p} \in X$ satisfying the KBL condition. Then the following assertions hold*:

(1) *We have $\operatorname{ord}_E(D_i) \ge b_i/r$ for any i.*

(2) *If $F_l^{\mathbf{w}} = \alpha_l x_0^{m_l} x_{i_l}$, where $\alpha_i \in \mathbb{C} \setminus \{0\}$, for some $1 \le l \le n-3$, then the weight*

$$\mathbf{w}'(x_1, \dots, x_n) = \tfrac{1}{r}(b'_1, \dots, b'_n),$$

where $b'_j = b_j$ for $j \ne l$ and $b'_l = b_l + r$, satisfies the KBL condition. In particular, $\operatorname{ord}_E(D_l) \ge (b_l + r)/r$.

We will use the following notation for a polynomial $f = f(x_0, \dots, x_n)$:

- For a monomial $p = x_0^{e_0} \cdots x_n^{e_n}$, we write $p \in f$ if p appears in f with nonzero (constant) coefficient.
- For a subset $\mathcal{C} \subset \{x_0, \dots, x_n\}$ and $\Pi = \Pi(\mathcal{C})$, we denote by $f|_{\Pi}$ the polynomial in variables $\{x_0, \dots, x_n\} \setminus \mathcal{C}$ obtained by putting $x_i = 0$ for $x_i \in \mathcal{C}$ in f.

3. Proof of birational superrigidity by numerical data

We prove birational superrigidity of codimension 4 quasismooth prime Fano 3-folds with no projections by utilizing only numerical data. The numerical data for each Fano 3-fold will be described in the beginning of the corresponding subsection. The Fano 3-folds are embedded in a weighted projective 7-space, denoted by $\mathbb{P}$, and we use the symbol p, q, r, s, t, u, v, w for the homogeneous coordinates of $\mathbb{P}$. We use the following terminologies: Let $X \subset \mathbb{P}$ be a codimension 4 quasismooth prime Fano 3-fold. For a homogeneous coordinate $z \in \{p, q, \dots, w\}$,

- $D_z := (z = 0) \cap X$ is the Weil divisor on X cut out by z, and
- $\mathsf{p}_z \in \mathbb{P}$ is the point at which only the coordinate z does not vanish.

Note that Theorem 1.2 will follow from Theorems 3.1, 3.2 and 3.4.

Fano 3-folds of numerical type #25. Let X be a well-formed quasismooth prime Fano 3-fold of numerical type #25, whose data consist of the following:

- $X \subset \mathbb{P}(2_p, 5_q, 6_r, 7_s, 8_t, 9_u, 10_v, 11_w)$.
- $(-K_X)^3 = \frac{1}{70}$.
- $\deg(F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9) = (16, 17, 18, 18, 19, 20, 20, 21, 22)$.
- $\mathcal{B}_X = \left\{7 \times \frac{1}{2}(1,1,1), \frac{1}{5}(1,1,4), \frac{1}{7}(1,2,5)\right\}$.

Here the subscripts $p, q, \dots, w$ of the weights means that they are the homogeneous coordinates of the indicated degrees, and $\mathcal{B}_X$ indicates the numbers and the types of singular points of X.

Theorem 3.1. *Let X be a well-formed quasismooth prime codimension 4 Fano 3-fold of numerical type #25. Then X is birationally superrigid.*

Proof. By Lemmas 2.9 and 2.10, no curve and no nonsingular point on X is a maximal center. By Lemma 2.13, singular points of type $\frac{1}{2}(1,1,1)$ are not maximal centers.

Let p be the singular point of type $\frac{1}{5}(1,1,4)$. Replacing the coordinate v if necessary, we may assume $\mathsf{p} = \mathsf{p}_q$. We set $\mathcal{C} = \{p, s, u, v\}$. We have

$$\mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) = \tfrac{2}{35} = \mathrm{wp}(\mathsf{p})(-K_X)^3.$$

By Lemma 2.16, it remains to show that $\Pi_X := \Pi_X(\mathcal{C} \cup \{q\}) = \varnothing$. We set $\Pi := \Pi(\mathcal{C} \cup \{q\}) \subset \mathbb{P}$ so that $\Pi_X = \Pi \cap X$. Since $\mathsf{p}_t \notin X$, one of the defining polynomials contain a power of t. By looking at the degrees of $F_1, \dots, F_9$, we have $t^2 \in F_1$. Similarly, we have $r^3 \in F_3$ and $w^2 \in F_9$ after possibly interchanging F_3 and F_4. The monomial t^2 (resp. r^3) is the only monomial of degree 16 (resp. 18) consisting of the variables r, t, w. The monomials w^2 and $t^2 r$ are the only monomials of degree 22 consisting of the variables r, t, w. Hence, rescaling r, t, w, we can write

$$F_1|_\Pi = t^2, \quad F_3|_\Pi = r^3, \quad F_9|_\Pi = w^2 + \alpha t^2 r,$$

for some $\alpha \in \mathbb{C}$. The set Π_X is contained in the common zero locus of the above 3 polynomials inside Π. The equations have only trivial solution and this shows that $\Pi_X = \varnothing$. Thus p is not a maximal center.

Let $\mathsf{p} = \mathsf{p}_s$ be the singular point of type $\frac{1}{7}(1,2,5)$ and set $\mathcal{C} = \{p, q, r\}$. We have

$$\mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) = \tfrac{1}{7} = \mathrm{wp}(\mathsf{p})(-K_X)^3.$$

By Lemma 2.16, it remains to show that $\Pi_X := \Pi_X(\mathcal{C} \cup \{s\}) = \varnothing$. We set $\Pi := \Pi(\mathcal{C} \cup \{s\}) \subset \mathbb{P}$ so that $\Pi_X = \Pi \cap X$. Since $\mathsf{p}_t, \mathsf{p}_u, \mathsf{p}_v, \mathsf{p}_w \notin X$, we may assume $t^2 \in F_1$, $u^2 \in F_3$, $v^2 \in F_6$ and $w^2 \in F_9$ after possibly interchanging defining polynomials of the same degree. Then we can write

$$F_1|_\Pi = t^2, \quad F_3|_\Pi = u^2 + \alpha v t, \quad F_6|_\Pi = v^2 + \beta w u, \quad F_9|_\Pi = w^2 + \gamma t^2 r,$$

for some $\alpha, \beta, \gamma \in \mathbb{C}$. This shows that $\Pi_X = \varnothing$ and thus p is not a maximal center. This completes the proof. $\square$

***Fano 3-folds of numerical type* #166.** Let X be a well-formed quasismooth prime Fano 3-fold of numerical type #166, whose data consist of the following:

- $X \subset \mathbb{P}(2_p, 2_q, 3_r, 3_s, 4_t, 4_u, 5_v, 5_w)$.
- $(-K_X)^3 = \frac{1}{6}$.
- $\deg(F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9) = (8, 8, 8, 9, 9, 9, 10, 10, 10)$.
- $\mathcal{B}_X = \left\{11 \times \frac{1}{2}(1,1,1), \frac{1}{3}(1,1,2)\right\}$.

Theorem 3.2. *Let X be a well-formed quasismooth prime codimension* 4 *Fano* 3*-fold of numerical type* #166*. Then X is birationally superrigid.*

Proof. By Lemma 2.9 no curve is a maximal center.

Let $\mathsf{p} = (\alpha_p : \alpha_q : \cdots : \alpha_w) \in X$ be a nonsingular point where $\alpha_p, \alpha_q, \dots, \alpha_w \in \mathbb{C}$. By Remark 2.12, we have $\Pi_X(p, q, r, s, t, u) = \varnothing$ since X does not have a singular point of index 5. Then we can take a coordinate $x \in \{p, q, r, s, t, u\}$ such that $\mathsf{p} \in (x \neq 0)$, i.e., $\alpha_x \neq 0$. The common zero locus of the homogeneous polynomials in the set

$$\{\alpha_y^{\deg x} x^{\deg y} - \alpha_x^{\deg y} y^{\deg x} \mid y \in \{p, q, r, s, t, u, v, w\} \setminus \{x\}\}$$

is a finite set of points including p. Any polynomial in the above set is of degree at most 20 since $x \notin \{v, w\}$. It follows that the base locus of $|\mathcal{I}_{\mathsf{p}}^m(-mlK_X)|$ is a finite set of points, that is, $-lK_X$ isolates p (see [Corti et al. 2000, Definition 5.2.4 and Lemma 5.6.4]), where $l \le 20$. By the argument in [loc. cit., Section 5.3], we conclude that p is not a maximal center since $20 < 4/(-K_X)^3$.

Let p be a singular point of type $\frac{1}{2}(1,1,1)$. After a change of coordinates, we may assume $\mathsf{p} = \mathsf{p}_p$. We set $\mathcal{C} = \{q, r, s, t, u\}$. We have

$$\mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) = \tfrac{1}{6} = \mathrm{wp}(\mathsf{p})(-K_X)^3.$$

Moreover we have $\Pi_X(\mathcal{C} \cup \{p\}) = \varnothing$ because X is quasismooth and it does not have a singular point of index 5. Thus, by Lemma 2.16, p is not a maximal center.

Let p be the singular point of type $\frac{1}{3}(1,1,2)$. After a change of coordinates, we may assume $\mathsf{p} = \mathsf{p}_s$. We set $\mathcal{C} = \{p, q, r\}$. Then we have

$$\mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) = \tfrac{1}{3} = \mathrm{wp}(\mathsf{p})(-K_X)^3.$$

By Lemma 2.16, it remains to show that $\Pi_X := \Pi_X(\mathcal{C} \cup \{s\}) = \varnothing$. We set $\Pi := \Pi(\mathcal{C} \cup \{s\}) \subset \mathbb{P}$ so that $\Pi_X = \Pi \cap X$. We have

$$\Pi_X = (F_1|_\Pi = F_2|_\Pi = F_3|_\Pi = F_7|_\Pi = F_8|_\Pi = F_9|_\Pi = 0) \cap \Pi.$$

We see that $F_1|_\Pi, F_2|_\Pi, F_3|_\Pi$ consist only of monomials in variables t, u and that $F_7|_\Pi, F_8|_\Pi, F_9|_\Pi$ consist only of monomials in variables v, w. It follows that

$$\Pi_X(p,q,r,s,v,w) = \Pi_X \cap \Pi(v,w) = (F_1|_\Pi = F_2|_\Pi = F_3|_\Pi = 0) \cap \Pi(p,q,r,s,v,w).$$

We have $\Pi_X(p,q,r,s,v,w) = \varnothing$ since X is well-formed, quasismooth and X has no singular point of index 4 (see Remark 2.12). Hence the equations

$$F_1|_\Pi = F_2|_\Pi = F_3|_\Pi = 0$$

imply $t = u = 0$. Similarly, by considering $\Pi_X(p,q,r,s,t,u) = \varnothing$, we see that the equations

$$F_7|_\Pi = F_8|_\Pi = F_9|_\Pi = 0$$

imply $v = w = 0$. It follows that $\Pi_X = \varnothing$ and p is not a maximal center. Therefore X is birationally superrigid. $\square$

***Fano* 3*-folds of numerical type* #282.** Let X be a well-formed quasismooth prime Fano 3-fold of numerical type #282, whose data consist of the following:

- $X \subset \mathbb{P}(1_p, 6_q, 6_r, 7_s, 8_t, 9_u, 10_v, 11_w)$.
- $(-K_X)^3 = \frac{1}{42}$.
- $\deg(F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9) = (16, 17, 18, 18, 19, 20, 20, 21, 22)$.
- $\mathcal{B} = \left\{2 \times \frac{1}{2}(1,1,1), 2 \times \frac{1}{3}(1,1,2), \frac{1}{6}(1,1,5), \frac{1}{7}(1,1,6)\right\}$.

Proposition 3.3. *Let X be a well-formed quasismooth prime codimension* 4 *Fano* 3*-fold of numerical type* #282*. Then no curve and no point is a maximal center except possibly for the singular point of type* $\frac{1}{6}(1,1,5)$.

Proof. By Lemmas 2.9, 2.10 and 2.13, it remains to exclude singular points of type $\frac{1}{3}(1,1,2)$ and $\frac{1}{7}(1,1,6)$ as maximal centers.

Let p be a singular point of type $\frac{1}{3}(1,1,2)$ and let $\varphi\colon (E \subset Y) \to (\mathsf{p} \in X)$ be the Kawamata blowup. We claim that $\Pi_X(p,s,t,w) = D_p \cap D_s \cap D_t \cap D_w$ is a finite set of points (containing p). Since X does not contain a singular point of index 10, we may assume that $v^2 \in F_6$. Then, by rescaling v, we have

$$F_6(0,q,r,0,0,u,v,0) = v^2$$

and this shows that $\Pi_X(p,s,t,w) = \Pi_X(p,s,t,v,w)$. The latter set consists of singular points $\{2 \times \frac{1}{3}(1,1,2), \frac{1}{6}(1,1,5)\}$ (see Remark 2.12) and thus $\Pi_X(p,s,t,w)$ is a finite set of points. We have

$$\operatorname{ord}_E(D_p), \operatorname{ord}_E(D_s) \geq \tfrac{1}{3}, \quad \operatorname{ord}_E(D_t), \operatorname{ord}_E(D_w) \geq \tfrac{2}{3}.$$

By Lemma 2.5, $N := -\varphi^* K_X - \frac{1}{21}E$ is a nef divisor on Y and we have $(N \cdot (-K_Y)^2) = 0$. Thus p is not a maximal center by Lemma 2.4.

Let $\mathsf{p} = \mathsf{p}_s$ be the singular point of type $\frac{1}{7}(1, 1, 6)$ and set $\mathcal{C} = \{p, q, r\}$. We have

$$\mathrm{ivr}_{\mathsf{p}}(\mathcal{C}) = \tfrac{1}{7} = \mathrm{wp}(\mathsf{p})(-K_X)^3.$$

We set $\Pi := \Pi(\mathcal{C} \cup \{s\})$. We see that $\mathsf{p}_t, \mathsf{p}_u, \mathsf{p}_v, \mathsf{p}_w \notin X$ since X does not have a singular point of index 8, 9, 10, 11. It follows that $t^2 \in F_1$, $w^2 \in F_9$ and we may assume $u^2 \in F_3$, $v^2 \in F_6$. Then, by rescaling t, u, v, w, we can write

$$F_1|_\Pi = t^2, \quad F_3|_\Pi = \alpha vt + u^2, \quad F_6|_\Pi = \beta wu + v^2, \quad F_9|_\Pi = w^2,$$

where $\alpha, \beta \in \mathbb{C}$. This shows that $\Pi_X(\mathcal{C} \cup \{s\}) = \Pi \cap X = \varnothing$. Thus p is not a maximal center by Lemma 2.16 and the proof is completed. □

***Fano* 3*-folds of numerical type* #308.** Let X be a well-formed quasismooth prime Fano 3-fold of numerical type #308, whose data consist of the following:

- $X \subset \mathbb{P}(1_p, 5_q, 6_r, 6_s, 7_t, 8_u, 9_v, 10_w)$.
- $(-K_X)^3 = \frac{1}{30}$.
- $\deg(F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9) = (14, 15, 16, 16, 17, 18, 18, 19, 20)$.
- $\mathcal{B}_X = \left\{\frac{1}{2}(1, 1, 1), \frac{1}{3}(1, 1, 2), \frac{1}{5}(1, 2, 3), 2 \times \frac{1}{6}(1, 1, 5)\right\}$.

Theorem 3.4. *Let X be a well-formed quasismooth prime Fano* 3*-fold of numerical type* #308. *Then X is birationally superrigid.*

Proof. By Lemmas 2.9, 2.10 and 2.13 no curve and no nonsingular point is a maximal center and the singular point of type $\frac{1}{2}(1, 1, 1)$ is not a maximal center.

Let p be the singular point of type $\frac{1}{3}(1, 1, 2)$, which is necessarily contained in $(p = q = t = u = w = 0)$, and let $\varphi\colon (E \subset Y) \to (\mathsf{p} \in X)$ be the Kawamata blowup. We set $\mathcal{C} = \{p, q, u\}$ and $\Pi = \Pi(\mathcal{C}) \subset \mathbb{P}$. Since $\mathsf{p}_t, \mathsf{p}_w \notin X$, we have $t^2 \in F_1$, $w^2 \in F_9$ and we can write

$$F_1|_\Pi = t^2, \quad F_9|_\Pi = w^2 + \alpha t^2 r + \beta t^2 s,$$

where $\alpha, \beta \in \mathbb{C}$. Thus,

$$\Pi_X(\mathcal{C}) = \Pi \cap X = \Pi_X(p, q, t, u, w),$$

and this consists of two $\frac{1}{6}(1, 1, 5)$ points and p. In particular $D_p \cap D_q \cap D_u = \Pi_X(\mathcal{C})$ is a finite set of points. We have

$$\mathrm{ord}_E(D_p) \geq \tfrac{1}{3}, \quad \mathrm{ord}_E(D_q) \geq \tfrac{2}{3}, \quad \mathrm{ord}_E(D_u) \geq \tfrac{2}{3},$$

hence $N := -8\varphi^* K_X - \frac{2}{3}E$ is a nef divisor on Y by Lemma 2.5. We have

$$(N \cdot (-K_Y)^2) = 8(-K_X)^3 - \tfrac{2}{3^3} \cdot \tfrac{3^2}{2} = -\tfrac{1}{15} < 0.$$

By Lemma 2.4, p is not a maximal center.

Let p be a singular point of type $\frac{1}{6}(1, 1, 5)$. After replacing r and s, we may assume $\mathsf{p} = \mathsf{p}_s$. We set $\mathcal{C} = \{p, q, r\}$. We have

$$\operatorname{ivr}_{\mathsf{p}}(\mathcal{C}) = \tfrac{1}{6} = \operatorname{wp}(\mathsf{p})(-K_X)^3.$$

Since $\mathsf{p}_t, \mathsf{p}_u, \mathsf{p}_v, \mathsf{p}_w \notin X$, we may assume $t^2 \in F_1$, $u^2 \in F_3$, $v^2 \in F_6$, $w^2 \in F_9$ after possibly interchanging F_3 with F_4 and F_6 with F_7. Then, by setting $\Pi = \Pi(\mathcal{C} \cup \{s\})$ and by rescaling t, u, v, w, we have

$$F_1|_\Pi = t^2, \quad F_3|_\Pi = u^2 + \alpha v t, \quad F_6|_\Pi = v^2 + \beta w u, \quad F_9|_\Pi = w^2,$$

where $\alpha, \beta \in \mathbb{C}$. This shows that $\Pi_X(\mathcal{C} \cup \{s\}) = \varnothing$ and p is not a maximal center by Lemma 2.16.

Finally, let p be a singular point of type $\frac{1}{5}(1, 2, 3)$ and let $\varphi \colon (E \subset Y) \to (\mathsf{p} \in X)$ be the Kawamata blowup. Replacing the coordinate w, we may assume $\mathsf{p} = \mathsf{p}_q$. We write

$$F_3 = \lambda q^3 p + \mu q^2 r + \nu q^2 s + q f_{11} + f_{16}, \quad F_4 = \lambda' q^3 p + \mu' q^2 r + \nu' q^2 s + q g_{11} + g_{16},$$

where $\lambda, \mu, \nu, \lambda', \mu', \nu' \in \mathbb{C}$ and $f_{11}, f_{16}, g_{11}, g_{16} \in \mathbb{C}[p, r, s, t, u, v, w]$ are homogeneous polynomials of the indicated degrees. Since X is quasismooth at $\mathsf{p} = \mathsf{p}_q$ and is of type $\frac{1}{5}(1, 2, 3)$, the matrix

$$\begin{pmatrix} \frac{\partial F_3}{\partial p}(\mathsf{p}) & \frac{\partial F_3}{\partial r}(\mathsf{p}) & \frac{\partial F_3}{\partial s}(\mathsf{p}) \\ \frac{\partial F_4}{\partial p}(\mathsf{p}) & \frac{\partial F_4}{\partial r}(\mathsf{p}) & \frac{\partial F_4}{\partial s}(\mathsf{p}) \end{pmatrix} = \begin{pmatrix} \lambda & \mu & \nu \\ \lambda' & \mu' & \nu' \end{pmatrix}$$

is of rank 2.

We first consider the case where $\mu\nu' - \nu\mu' \ne 0$. By replacing r and s, we may assume that $\mu = \nu' = 1$ and $\lambda = \nu = \lambda' = \mu' = 0$. We consider the weight at p

$$\mathbf{w}(p, r, s, t, u, v, w) = \tfrac{1}{5}(1, 1, 1, 2, 3, 4, 5),$$

which is an admissible weight satisfying the KBL condition. Then $F_3^{\mathbf{w}} = q^2 r$ and $F_4^{\mathbf{w}} = q^2 s$, and this implies $\operatorname{ord}_E(D_r), \operatorname{ord}_E(D_s) \ge \frac{6}{5}$ by Lemma 2.18. Note that $\operatorname{ord}_E(D_p) \ge \mathbf{w}(p) = \frac{1}{5}$ by Lemma 2.18. We set $\mathcal{C} = \{p, r, s\}$ and $\Pi = \Pi(\mathcal{C} \cup \{q\})$. By rescaling t, u, v, w, we can write

$$F_1|_\Pi = t^2, \quad F_3|_\Pi = u^2 + \alpha v t, \quad F_6|_\Pi = v^2 + \beta w u, \quad F_9|_\Pi = w^2,$$

where $\alpha, \beta \in \mathbb{C}$. Hence $\Pi_X(\mathcal{C} \cup \{q\}) = \varnothing$. Since D_q is an ample divisor, this implies that $D_p \cap D_r \cap D_s$ is a finite set of points (including p). By Lemma 2.5, $N := -\varphi^* K_X - \frac{1}{5} E$ is a nef divisor on Y. We have

$$(N \cdot (-K_Y)^2) = (-K_X)^3 - \tfrac{1}{5^3}(E^3) = \tfrac{1}{30} - \tfrac{1}{30} = 0,$$

and this shows that p is not a maximal center.

Next we consider the case where $\mu\nu' - \nu\mu' = 0$. By replacing r and s suitably and by possibly interchanging F_3 and F_4, we may assume that

$$F_3 = q^3 p + q f_{11} + f_{16}, \quad F_4 = q^2 s + q g_{11} + g_{16}.$$

Let $\mathbf{w}$ be the same weight at p as in the previous case, which is again an admissible weight satisfying the KBL condition. It is straightforward to see that $F_3^{\mathbf{w}} = q^3 p$, so that $\operatorname{ord}_E(D_p) \ge \frac{6}{5}$. Let $\mathcal{L} \subset |-6K_X|$ be the pencil generated by the sections r and s. Since $\mathrm{ord}_E(D_r) = \frac{1}{5}$ and $\mathrm{ord}_E(D_s) \geq \frac{1}{5}$, a general member $L \in \mathcal{L}$ vanishes along E to order $\frac{1}{5}$ so that $\tilde{L} \sim -6\varphi^* K_X - \frac{1}{5}E$. We have

$$(-K_Y \cdot \tilde{D}_p \cdot \tilde{L}) = 6(-K_X)^3 - \frac{\mathrm{ord}_E(D_p)}{5^2} \cdot (E^3) = \frac{1}{5} - \frac{\mathrm{ord}_E(D_p)}{6} \leq 0$$

since $\mathrm{ord}_E(D_p) \geq \frac{6}{5}$. By Lemma 2.6, p is not a maximal center and the proof is complete. □

4. Birational superrigidity of cluster Fano 3-folds

In this section we prove Theorem 1.1 which follows from Theorems 4.2 and 4.4 below.

#282 *by* $\mathbf{G}_2^{(4)}$ *format.* Let X be a quasismooth codimension 4 prime Fano 3-fold of numerical type #282 constructed in $\mathsf{G}_2^{(4)}$ format. Then, by [Coughlan and Ducat 2018, Example 5.5], X is defined by the following polynomials in $\mathbb{P} := \mathbb{P}(1_p, 6_q, 6_r, 7_s, 8_t, 9_u, 10_v, 11_w)$:

$$\begin{aligned}
F_1 &= t^2 - qv + sQ_9,\\
F_2 &= ut - qw + s(v + p^2 t),\\
F_3 &= t(v + p^2 t) - uQ_9 + q(qr + p^4 t),\\
F_4 &= (w + p^4 s)s - P_{12} q + u(u + p^2 s),\\
F_5 &= tw - uv + s(qr + p^4 t),\\
F_6 &= (qr + p^4 t)t - Q_9 w + v(v + p^2 t),\\
F_7 &= rs^2 - wu + tP_{12},\\
F_8 &= P_{12} Q_9 - (vw + p^4 qw + p^2 uv + uqr + str - stp^2),\\
F_9 &= rs(u + p^2 s) - vP_{12} + w(w + p^4 s).
\end{aligned}$$

Here $P_{12}, Q_9 \in \mathbb{C}[p, q, r, s, t, u, v, w]$ are homogeneous polynomials of the indicated degree. Recall that $(-K_X)^3 = \frac{1}{42}$.

Lemma 4.1. *The following assertions hold*:

(1) $r^2 \in P_{12}$ *and* $u \in Q_9$.

(2) $X \subset \mathbb{P}$ *is well-formed.*

Proof. It is straightforward to check that X is quasismooth at $\mathsf{p}_r \in X$ if and only if $r^2 \in P_{12}$ and $u \in Q_9$, and this proves (1).

We prove (2). We set

$$\Pi_2 := \Pi_X(p, s, u, w), \quad \Pi_3 := \Pi_X(p, s, t, v, w).$$

It is enough to show that neither Π_2 nor Π_3 contain a surface (note here that $P_{12}|_{\Pi_2} \neq 0$ by (1)). We see that Π_2 is isomorphic to the closed subscheme in $\mathbb{P}(6_q, 6_r, 8_t, 10_v)$ defined by the equations

$$t^2 - qv = tv + q^2 r = qP_{12}|_{\Pi_2} = qrt + v^2 = tP_{12}|_{\Pi_2} = vP_{12}|_{\Pi_2} = 0.$$

We leave the readers to check that Π_2 does not contain a surface. We see that Π_3 is isomorphic to the closed subscheme in $\mathbb{P}(6_q, 6_r, 9_u)$ defined by the equations

$$-uQ_9|_{\Pi_3} + q^2 r = -qP_{12}|_{\Pi_3} + u^2 = P_{12}|_{\Pi_3} Q_9|_{\Pi_3} - uqr = 0.$$

Hence Π_3 does not contain a surface since it is clearly a proper closed subset of the surface $\mathbb{P}(6, 6, 9)$. Thus $X \subset \mathbb{P}$ is well-formed. $\square$

Theorem 4.2. *Let X be a codimension 4 Fano 3-fold of numerical type #282 constructed in $\mathsf{G}_2^{(4)}$ format. Then X is birationally superrigid.*

Proof. By Lemma 4.1, $X \subset \mathbb{P}$ is well-formed. We can apply Proposition 3.3 and it remains to exclude the singular point $\mathsf{p} \in X$ of type $\frac{1}{6}(1, 1, 5)$ as a maximal center. We have $\mathsf{p} = \mathsf{p}_r$ since $\mathsf{p}_r \in X$ and X has a unique singular point of index 6. We set $\mathcal{C} = \{p, q\}$, $\Pi = \Pi(\mathcal{C})$ and $\Gamma := \Pi_X(\mathcal{C}) = \Pi \cap X$.

We will show that Γ is an irreducible and reduced curve. By Lemma 4.1, we can write

$$P_{12}|_{\Pi} = \lambda r^2, \quad Q_9|_{\Pi} = \mu u,$$

where $\lambda, \mu \in \mathbb{C} \setminus \{0\}$. Then we have

$$\begin{aligned}
&F_1|_{\Pi} = t^2 + \mu su, && F_4|_{\Pi} = ws + u^2, && F_7|_{\Pi} = rs^2 - wu + \lambda tr^2,\\
&F_2|_{\Pi} = ut + sv, && F_5|_{\Pi} = tw - uv, && F_8|_{\Pi} = \lambda\mu r^2 u - (vw + str),\\
&F_3|_{\Pi} = tv - \mu u^2, && F_6|_{\Pi} = -\mu uw + v^2, && F_9|_{\Pi} = rsu - \lambda vr^2 + w^2.
\end{aligned}$$

We work on the open subset U on which $w \neq 0$. Then $\Gamma \cap U$ is isomorphic to the $\mathbb{Z}/11\mathbb{Z}$-quotient of the affine curve

$$(\lambda r^2 v + \mu^3 r v^6 - 1 = 0) \subset \mathbb{A}^2_{r,v}.$$

It is straightforward to check that the polynomial $\lambda r^2 v + \mu^3 r v^6 - 1$ is irreducible. Thus $\Gamma \cap U$ is an irreducible and reduced affine curve. It is also straightforward to check that

$$\Gamma \cap (w = 0) = X \cap (p = q = w = 0) = \{\mathsf{p}_r, \mathsf{p}_s\}.$$

This shows that Γ is an irreducible and reduced curve.

Let $\varphi\colon (E \subset Y) \to (\mathsf{p} \in X)$ be the Kawamata blowup and let $\tilde{\Delta}$ be the proper transform via φ of a divisor or curve Δ on X. We show that $\tilde{D}_p \cap \tilde{D}_q \cap E$ does not contain a curve. Consider the weight

$$\mathbf{w}(p, q, s, t, u, v, w) = \tfrac{1}{6}(1, 6, 1, 2, 3, 4, 5),$$

which is clearly an admissible weight satisfying the KBL condition. We set $g_i = F_i^{\mathbf{w}}(p, q, 1, s, t, u, v, w)$. We have

$$\begin{aligned} g_4 &= (w + p^4)s - \lambda q + u(u + p^2 s), \\ g_7 &= s^2 + \lambda t, \\ g_8 &= \lambda\mu u - st, \\ g_9 &= s(u + p^2 s) - \lambda v. \end{aligned}$$

Since E is isomorphic to the subvariety

$$(g_4 = g_7 = g_8 = g_9 = 0) \subset \mathbb{P}(1_p, 6_q, 1_s, 2_t, 3_u, 4_v, 5_w),$$

it is straightforward to check that $\tilde{D}_p \cap \tilde{D}_q \cap E$ consists of a finite set of points (in fact, 2 points). Thus we have $\tilde{D}_p \cdot \tilde{D}_q = \tilde{\Gamma}$ since $D_p \cdot D_q = \Gamma$.

We have

$$\tilde{D}_p \sim -\varphi^* K_X - \tfrac{1}{6} E, \quad \tilde{D}_q \sim -6\varphi^* K_X - \tfrac{e}{6} E,$$

for some integer $e \geq 6$ and hence

$$(\tilde{D}_p \cdot \tilde{\Gamma}) = (\tilde{D}_p^2 \cdot \tilde{D}_q) = \tfrac{1}{7} - \tfrac{e}{30} < 0.$$

By [Okada 2018, Lemma 2.18], p is not a maximal center. □

#282 *by* C_2 *format.* Let X be a quasismooth codimension 4 prime Fano 3-fold of numerical type #282 constructed in C_2 format. Then, by [Coughlan and Ducat 2018, Example 5.5], X is defined by the following polynomials in $\mathbb{P}(1_p, 6_q, 6_r, 7_s, 8_t, 9_u, 10_v, 11_w)$:

$$\begin{aligned} F_1 &= tR_8 - S_6 Q_{10} + su, \\ F_2 &= tu - wS_6 + sv, \\ F_3 &= rS_6^2 - vR_8 + u^2, \\ F_4 &= tQ_{10} - S_6 P_{12} + sw, \\ F_5 &= rsS_6 - wR_8 + uQ_{10}, \\ F_6 &= rs^2 - P_{12}R_8 + Q_{10}^2, \\ F_7 &= rtS_6 - vQ_{10} + uw, \\ F_8 &= rst - wQ_{10} + uP_{12}, \\ F_9 &= rt^2 - vP_{12} + w^2. \end{aligned}$$

Here $P_{12}, Q_{10}, R_8, S_6 \in \mathbb{C}[p, q, r, s, t, u, v, w]$ are homogeneous polynomials of the indicated degree. In the following we assume that $q \in S_6$, and then, we assume that $S_6 = q$ by a change of coordinates.

Lemma 4.3. *Under the above setting, the following assertions hold*:

(1) $r^2 \in P_{12}$, $v \in Q_{10}$ *and* $t \in R_8$.

(2) $X \subset \mathbb{P}$ *is well-formed.*

Proof. We have $\mathsf{p}_r \in X$ and X is quasismooth at p_r if and only if $r^2 \in P_{12}$. Similarly, it is easy to check that if $v \notin Q_{10}$ (resp. $t \notin R_8$), then X is not quasismooth at p_v (resp. p_r). This proves (1). We leave the readers to check that neither Π_2 nor Π_3 contain a surface, where Π_2, Π_3 are those given in the proof of Lemma 4.1, and this proves (2). $\square$

Theorem 4.4. *Let X be a quasismooth prime codimension* 4 *Fano* 3*-fold of numerical type* #282 *constructed by* C_2 *format. We assume that $q \in S_6$. Then X is birationally superrigid.*

Proof. By Lemma 4.3, we can apply Proposition 3.3 and it remains to exclude the singular point p of type $\frac{1}{6}(1, 1, 5)$ as a maximal center.

The singular point p corresponds to the solution of the equation

$$p = s = t = u = v = w = S_6 = 0,$$

and thus $\mathsf{p} = \mathsf{p}_r$ since $S_6 = q$ by our setting. We set $\mathcal{C} = \{p, q\}$ and $\Pi = \Pi(\mathcal{C})$.

We will show that $\Gamma := \Pi \cap X$ is an irreducible and reduced curve. We have $\Pi_X(\{p, q, r, s\}) = \varnothing$ (see the proof of Proposition 3.3). Hence $\Gamma \cap (s = 0) = \Pi_X(\{p, q, s\})$ does not contain a curve and it remains to show that $\Gamma \cap U_s$ is irreducible and reduced, where $U_s := (s \neq 0) \subset \mathbb{P}$ is the open subset. By Lemma 4.3 we can write

$$P_{12}|_{\Pi} = \lambda r^2, \quad Q_{10}|_{\Pi} = \mu v, \quad R_8|_{\Pi} = \nu t,$$

for some $\lambda, \mu, \nu \in \mathbb{C} \setminus \{0\}$, and we have $S_6|_{\Pi} = 0$. Note that $F_i|_{\Pi} = F_i|_{\Pi}(r, s, t, u, v, w)$ is a polynomial in variables r, s, t, u, v, w and we set $f_i = F_i|_{\Pi}(r, 1, t, u, v, w)$. Let $C \subset \mathbb{A}^5_{r,t,u,v,w}$ be the affine scheme defined by the equations

$$f_1 = f_2 = \cdots = f_9 = 0.$$

Then $\Gamma \cap U_s$ is isomorphic to the quotient of C by the natural $\mathbb{Z}/7\mathbb{Z}$-action. We have

$$\begin{aligned}
f_1 &= \nu t^2 + u, & f_4 &= \mu t v + w, & f_7 &= -\mu v^2 + u w,\\
f_2 &= t u + v, & f_5 &= -\nu t w + \mu u v, & f_8 &= r t - \mu v w + \lambda r^2 u,\\
f_3 &= -\nu t v + u^2, & f_6 &= r - \lambda \nu r^2 t + \mu^2 v^2, & f_9 &= r t^2 - \lambda r^2 v + w^2.
\end{aligned}$$

By the equations $f_1 = 0$, $f_2 = 0$ and $f_4 = 0$, we have

$$u = -\nu t^2, \quad v = -t u = \nu t^3, \quad w = -\mu t v = -\mu \nu t^4.$$

By eliminating the variables u, v, w and cleaning up the equations, C is isomorphic to the hypersurface in $\mathbb{A}^2_{r,t}$ defined by

$$r - \lambda \nu r^2 t + \mu^2 \nu^2 t^6 = 0,$$

which is an irreducible and reduced curve since $\mu\nu \neq 0$, and so is $\Gamma \cap U_s$. Thus Γ is an irreducible and reduced curve.

Let $\varphi\colon (E \subset Y) \to (\mathsf{p} \in X)$ be the Kawamata blowup. We have $e := \operatorname{ord}_E(D_q) \geq \frac{6}{6}$ and $\operatorname{ord}_E(D_p) = \frac{1}{6}$ so that we have

$$\tilde{D}_q \sim -6\varphi^* K_X - \tfrac{e}{6}E = -6K_Y + \tfrac{6-e}{6}E, \quad \tilde{D}_p \sim -\varphi^* K_X - \tfrac{1}{6}E = -K_Y.$$

We show that $\tilde{D}_q \cap \tilde{D}_p \cap E$ does not contain a curve. The Kawamata blowup φ is realized as the weighted blowup at p with the weight

$$\mathbf{w}(p, q, s, t, u, v, w) = \tfrac{1}{6}(1, 6, 1, 2, 3, 4, 5),$$

which is an admissible weight satisfying the KBL condition. We have

$$\begin{aligned}
F_4^{\mathbf{w}} &= -\lambda q r^2 + t(\mu v + h) + sw,\\
F_6^{\mathbf{w}} &= -\lambda \mu t r^2 + r s^2,\\
F_8^{\mathbf{w}} &= \lambda u r^2 + rst,\\
F_9^{\mathbf{w}} &= -\lambda v r^2 + r t^2,
\end{aligned}$$

where we define $h := Q_{10}^{\mathbf{w}} - \mu v$. Note that h is a linear combination of up, tp^2, sp^3, rp^4 and thus h is divisible by p. It follows that E is isomorphic to the subscheme in $\mathbb{P}(1_p, 6_q, 1_s, 2_t, 3_u, 4_v, 5_w)$ defined by the equations

$$\lambda q - t(\mu v + h) - sw = \lambda \mu t - s^2 = \lambda u + st = -\lambda v + t^2 = 0.$$

It is now straightforward to check that $\tilde{D}_q \cap \tilde{D}_p \cap E = (p = q = 0) \cap E$ is a finite set of points (in fact, it consists of 2 points). This shows that $\tilde{D}_q \cdot \tilde{D}_p = \tilde{\Gamma}$ since $D_q \cdot D_p = \Gamma$. We have

$$(\tilde{D}_p \cdot \tilde{\Gamma}) = (\tilde{D}_p^2 \cdot \tilde{D}_q) = 6(-K_X)^3 - \tfrac{e}{6^3}(E^3) = \tfrac{1}{7} - \tfrac{e}{30} < 0$$

since $e \geq 6$. By [Okada 2018, Lemma 2.18] p is not a maximal center. □

5. On further problems

Prime Fano 3-folds with no projection centers. We further investigate birational superrigidity of prime Fano 3-folds of codimension c with no projection centers for $5 \leq c \leq 9$. There are only a few such candidates, which can be summarized as follows.

- In codimension $c \in \{5, 7\}$ there is a unique candidate and it corresponds to smooth prime Fano 3-folds of degree $2c+2$. All of these Fano 3-folds are rational (see [Iskovskikh and Prokhorov 1999, Corollary 4.3.5 or Section 12.2]) and are not birationally superrigid.
- In codimension 6 there are 2 candidates; one candidate corresponds to smooth prime Fano 3-folds of degree 14 which are birational to smooth cubic 3-folds (see [Takeuchi 1989; Iskovskikh 1979]) and are not birationally superrigid, and the existence is not known for the other candidate which is #78 in the database.

- In codimension 8 there are 2 candidates; one corresponds to smooth prime Fano 3-folds of degree 18 which are rational (see [Iskovskikh and Prokhorov 1999, Corollary 4.3.5 or Section 12.2]), and the existence is not known for the other candidate which is #33 in the database.
- In codimension 9 there is a unique candidate of smooth prime Fano 3-folds of degree 20. However, according to the classification of smooth Fano 3-folds there is no such Fano 3-fold (see e.g., [Takeuchi 1989, Theorem 0.1]).

It follows that, in codimension up to 9, #33 and #78 are the only remaining unknown cases for birational superrigidity (of general members).

Question 5.1. Do there exist prime Fano 3-folds which correspond to #33 or #78? If yes, then is a (general) such Fano 3-fold birationally superrigid?

In codimension 10 and higher there are a lot of candidates of Fano 3-folds with no projection centers. We expect that many of them are nonexistence cases and that there are only a few birationally superrigid Fano 3-folds in higher codimensions.

Question 5.2. Is there a numerical type (in other words, graded ring database ID) #i in codimension greater than 9 such that a (general) quasismooth prime Fano 3-fold of numerical type #i is birationally superrigid?

Classification of birationally superrigid Fano 3-folds. There are many difficulties in the complete classification of birationally superrigid Fano 3-folds. For example, we need to consider Fano 3-folds which are not necessarily quasismooth or not necessarily prime. We also need to understand subtle behaviors of birational superrigidity in a family.

Question 5.3. Is there a birationally superrigid Fano 3-fold which is either of Fano index greater than 1 or has a nonquotient singularity?

Remark 5.4. By recent developments [Pukhlikov 2019; Suzuki 2017; Liu and Zhuang 2019], it is known that there exist birationally superrigid Fano varieties which have nonquotient singularities at least in very high dimensions. On the other hand, only a little is known for Fano varieties of index greater than 1 (see [Pukhlikov 2016]) and there is no example of birationally superrigid Fano varieties of index greater than 1.

We concentrate on quasismooth prime Fano 3-folds. Even in that case, it is necessary to consider those with a projection center, which are not treated in this paper. Let X be a general quasismooth prime Fano 3-fold of codimension c. Then the following are known:

- When $c = 1$, X is birationally superrigid if and only if X does not admit a type I projection center (see [Iskovskikh and Manin 1971; Corti et al. 2000; Cheltsov and Park 2017]).
- When $c = 2, 3$, X is birationally superrigid if and only if X is singular and admits no projection center (see [Iskovskikh and Pukhlikov 1996; Okada 2014a; Ahmadinezhad and Zucconi 2016; Ahmadinezhad and Okada 2018]).

With this evidence we expect the following.

Conjecture 5.5. Let X be a general quasismooth prime Fano 3-fold of codimension at least 2. Then X is birationally superrigid if and only if X is singular and admits no projection centers.

Acknowledgements

Okada would like to thank Stephen Coughlan for fruitful information on cluster Fano 3-folds. He also would like to thank the referees for valuable suggestions. He is partially supported by JSPS KAKENHI Grant Number JP18K03216.

References

[Ahmadinezhad and Okada 2018] H. Ahmadinezhad and T. Okada, "Birationally rigid Pfaffian Fano 3-folds", *Algebr. Geom.* **5**:2 (2018), 160–199. MR Zbl

[Ahmadinezhad and Zucconi 2016] H. Ahmadinezhad and F. Zucconi, "Mori dream spaces and birational rigidity of Fano 3-folds", *Adv. Math.* **292** (2016), 410–445. MR Zbl

[Altınok 1998] S. Altınok, *Graded rings corresponding to polarised K3 surfaces and $\mathbb{Q}$-Fano 3-folds*, Ph.D. thesis, University of Warwick, 1998.

[Altınok et al. 2002] S. Altınok, G. Brown, and M. Reid, "Fano 3-folds, $K3$ surfaces and graded rings", pp. 25–53 in *Topology and geometry: commemorating SISTAG* (Singapore, 2001), edited by A. J. Berrick et al., Contemp. Math. **314**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Brown and Kasprzyk 2009] G. Brown and A. Kasprzyk, "The graded ring database", online database, 2009, Available at www.grdb.co.uk.

[Brown et al. 2012] G. Brown, M. Kerber, and M. Reid, "Fano 3-folds in codimension 4, Tom and Jerry, I", *Compos. Math.* **148**:4 (2012), 1171–1194. MR Zbl

[Cheltsov and Park 2017] I. Cheltsov and J. Park, *Birationally rigid Fano threefold hypersurfaces*, Mem. Amer. Math. Soc. **1167**, Amer. Math. Soc., Providence, RI, 2017. MR Zbl

[Chen et al. 2011] J.-J. Chen, J. A. Chen, and M. Chen, "On quasismooth weighted complete intersections", *J. Algebraic Geom.* **20**:2 (2011), 239–262. MR Zbl

[Corti 1995] A. Corti, "Factoring birational maps of threefolds after Sarkisov", *J. Algebraic Geom.* **4**:2 (1995), 223–254. MR Zbl

[Corti et al. 2000] A. Corti, A. Pukhlikov, and M. Reid, "Fano 3-fold hypersurfaces", pp. 175–258 in *Explicit birational geometry of 3-folds*, edited by A. Corti and M. Reid, Lond. Math. Soc. Lecture Note Ser. **281**, Cambridge Univ. Press, 2000. MR Zbl

[Coughlan and Ducat 2018] S. Coughlan and T. Ducat, "Constructing Fano 3-folds from cluster varieties of rank 2", preprint, 2018. arXiv

[Debarre et al. 2012] O. Debarre, A. Iliev, and L. Manivel, "On the period map for prime Fano threefolds of degree 10", *J. Algebraic Geom.* **21**:1 (2012), 21–59. MR Zbl

[Dimca 1986] A. Dimca, "Singularities and coverings of weighted complete intersections", *J. Reine Angew. Math.* **366** (1986), 184–193. MR Zbl

[Iano-Fletcher 2000] A. R. Iano-Fletcher, "Working with weighted complete intersections", pp. 101–174 in *Explicit birational geometry of 3-folds*, edited by A. Corti, Lond. Math. Soc. Lecture Note Ser. **281**, Cambridge Univ. Press, 2000. MR Zbl

[Iskovskikh 1979] V. A. Iskovskikh, "Birational automorphisms of three-dimensional algebraic varieties", pp. 159–236 in *Current problems in mathematics, XII*, edited by R. V. Gamkrelidze, VINITI, Moscow, 1979. In Russian; translated in *J. Soviet Math.* **13**:6 (1980), 815–868. MR Zbl

[Iskovskikh and Manin 1971] V. A. Iskovskikh and Y. I. Manin, "Three-dimensional quartics and counterexamples to the Lüroth problem", *Mat. Sb.* (*N.S.*) **86(128)**:1(9) (1971), 140–166. In Russian; translated in *Math. USSR-Sb.* **15**:1 (1971), 141–166. MR Zbl

[Iskovskikh and Prokhorov 1999] V. A. Iskovskikh and Y. G. Prokhorov, *Algebraic geometry, V: Fano varieties*, Encycl. Math. Sci. **47**, Springer, 1999. MR Zbl

[Iskovskikh and Pukhlikov 1996] V. A. Iskovskikh and A. V. Pukhlikov, "Birational automorphisms of multidimensional algebraic manifolds", *J. Math. Sci.* **82**:4 (1996), 3528–3613. MR Zbl

[Kawamata 1996] Y. Kawamata, "Divisorial contractions to 3-dimensional terminal quotient singularities", pp. 241–246 in *Higher-dimensional complex varieties* (Trento, Italy, 1994), edited by M. Andreatta and T. Peternell, de Gruyter, Berlin, 1996. MR Zbl

[Liu and Zhuang 2019] Y. Liu and Z. Zhuang, "Birational superrigidity and K-stability of singular Fano complete intersections", *Int. Math. Res. Not.* (online publication August 2019).

[Okada 2014a] T. Okada, "Birational Mori fiber structures of $\mathbb{Q}$-Fano 3-fold weighted complete intersections", *Proc. Lond. Math. Soc.* (3) **109**:6 (2014), 1549–1600. MR Zbl

[Okada 2014b] T. Okada, "Birational Mori fiber structures of $\mathbb{Q}$-Fano 3-fold weighted complete intersections, III", 2014. To appear in *Kyoto J. Math.* arXiv

[Okada 2018] T. Okada, "Birational Mori fiber structures of $\mathbb{Q}$-Fano 3-fold weighted complete intersections, II", *J. Reine Angew. Math.* **738** (2018), 73–129. MR Zbl

[Pukhlikov 1998] A. V. Pukhlikov, "Birational automorphisms of Fano hypersurfaces", *Invent. Math.* **134**:2 (1998), 401–426. MR Zbl

[Pukhlikov 2016] A. V. Pukhlikov, "Birational geometry of Fano hypersurfaces of index two", *Math. Ann.* **366**:1-2 (2016), 721–782. MR Zbl

[Pukhlikov 2019] A. V. Pukhlikov, "Birationally rigid complete intersections with a singular point of high multiplicity", *Proc. Edinb. Math. Soc.* (2) **62**:1 (2019), 221–239. MR Zbl

[Suzuki 2017] F. Suzuki, "Birational rigidity of complete intersections", *Math. Z.* **285**:1-2 (2017), 479–492. MR Zbl

[Takeuchi 1989] K. Takeuchi, "Some birational maps of Fano 3-folds", *Compos. Math.* **71**:3 (1989), 265–283. MR Zbl

Communicated by János Kollár
Received 2019-01-31 Revised 2019-06-08 Accepted 2019-08-26

okada@cc.saga-u.ac.jp *Department of Mathematics, Saga University, Saga, Japan*

ALGEBRA AND NUMBER THEORY 14:1 (2020)
dx.doi.org/10.2140/ant.2020.14.213

Coble fourfold, $\mathfrak{S}_6$-invariant quartic threefolds, and Wiman–Edge sextics

Ivan Cheltsov, Alexander Kuznetsov and Konstantin Shramov

To Arnaud Beauville, on the occasion of his 70th birthday

We construct two small resolutions of singularities of the Coble fourfold (the double cover of the four-dimensional projective space branched over the Igusa quartic). We use them to show that all $\mathfrak{S}_6$-invariant three-dimensional quartics are birational to conic bundles over the quintic del Pezzo surface with the discriminant curves from the Wiman–Edge pencil. As an application, we check that $\mathfrak{S}_6$-invariant three-dimensional quartics are unirational, obtain new proofs of rationality of four special quartics among them and irrationality of the others, and describe their Weil divisor class groups as $\mathfrak{S}_6$-representations.

1. Introduction

Consider the projectivization $\mathbb{P}^5$ of the standard permutation representation of the symmetric group $\mathfrak{S}_6$ over an algebraically closed field $\Bbbk$ of characteristic zero, and the invariant hyperplane $\mathbb{P}^4$ given by the equation

$$x_1+x_2+x_3+x_4+x_5+x_6=0 \tag{1.1}$$

therein, where $x_1,\dots,x_6$ are homogeneous coordinates in $\mathbb{P}^5$. Consider the classical family of $\mathfrak{S}_6$-invariant quartics X_t, $t\in\Bbbk\cup\{\infty\}$, in this hyperplane defined by the equations

$$(x_1^4+x_2^4+x_3^4+x_4^4+x_5^4+x_6^4)-t(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2)^2=0; \tag{1.2}$$

MSC2010: primary 14E08; secondary 14E05, 14J30, 14J35, 14J45.
Keywords: Fano varieties, Igusa quartic, conic bundle, del Pezzo surface, Wiman–Edge pencil.

studied in [Beauville 2013]. Every $\mathfrak{S}_6$-invariant quartic in $\mathbb{P}^4$ is one of the quartics X_t; moreover, most of these quartics have automorphism groups isomorphic to $\mathfrak{S}_6$, and every quartic threefold with a faithful $\mathfrak{S}_6$-action is isomorphic to some X_t (see Lemma 3.4). We refer to these quartics as $\mathfrak{S}_6$-*invariant quartics*.

Every quartic X_t is singular along a certain 30-point orbit $\Sigma_{30} \subset \mathbb{P}^4$ of the group $\mathfrak{S}_6$ (see Section 3.1), and Σ_{30} coincides with $\mathrm{Sing}(X_t)$ unless $t = \infty$ or t is in the finite *discriminant set*

$$\mathfrak{D} := \left\{\tfrac{1}{4}, \tfrac{1}{2}, \tfrac{1}{6}, \tfrac{7}{10}\right\}. \tag{1.3}$$

For these special values of t the singular locus of X_t is even larger (see Theorem 3.3 for its detailed description).

The quartic $X_{1/4}$ that corresponds to the parameter $t = \frac{1}{4}$ is particularly interesting. Its equation can be written as

$$(x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_5^4 + x_6^4) - \tfrac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2)^2 = 0 \tag{1.4}$$

inside the hyperplane (1.1). It is called the *Igusa quartic*. The Igusa quartic is singular along a union of 15 lines (that itself forms an interesting configuration CR, called the *Cremona–Richmond configuration*). In this sense, $X_{1/4}$ is the most singular of all $\mathfrak{S}_6$-invariant quartics, except for X_∞ (which is a double quadric, i.e., a quadric with an everywhere nonreduced scheme structure).

The quartic $X_{1/2}$ is known as the *Burkhardt quartic*. It has the largest symmetry group among the other quartics in this family (with the exception of X_∞); see [Coble 1906] and Lemma 3.4. It also has many other interesting properties; see for instance [Todd 1936; de Jong et al. 1990; Hunt 1996, Section 5].

The quartics $X_{1/6}$ and $X_{7/10}$ have been studied in [Cheltsov and Shramov 2016b], compare [Todd 1933; 1935; Cheltsov and Shramov 2014].

The double cover of $\mathbb{P}^4$ branched over the Igusa quartic is called the *Coble fourfold*. We denote it by $\mathscr{Y}$ and write

$$\pi : \mathscr{Y} \to \mathbb{P}^4$$

for the double covering morphism. The Coble fourfold can be written as a complete intersection in the weighted projective space $\mathbb{P}(2, 1^6)$ of the hyperplane (1.1) with the hypersurface

$$x_0^2 = (x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_5^4 + x_6^4) - \tfrac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2)^2, \tag{1.5}$$

where x_0 is the coordinate of weight 2. The Coble fourfold $\mathscr{Y}$ is singular along the Cremona–Richmond configuration CR, because so is the Igusa quartic. Moreover, it has a big group of symmetries: it carries an action of the symmetric group $\mathfrak{S}_6$ by permutation of coordinates

$$g \cdot (x_0 : x_1 : x_2 : x_3 : x_4 : x_5 : x_6) := (x_0 : x_{g(1)} : x_{g(2)} : x_{g(3)} : x_{g(4)} : x_{g(5)} : x_{g(6)}), \tag{1.6}$$

and also the Galois involution $\sigma : \mathscr{Y} \to \mathscr{Y}$ of the double cover

$$\sigma(x_0 : x_1 : x_2 : x_3 : x_4 : x_5 : x_6) := (-x_0 : x_1 : x_2 : x_3 : x_4 : x_5 : x_6), \tag{1.7}$$

commuting with the symmetric group action. One can check (see Corollary 3.5) that they generate the whole automorphism group

$$\mathrm{Aut}(\mathscr{Y}) \cong \mathfrak{S}_6 \times \boldsymbol{\mu}_2,$$

where $\boldsymbol{\mu}_2$ denotes the group of order 2. Sometimes it is convenient to twist the action of the symmetric group by the Galois involution. The obtained action

$$g \diamond (x_0 : x_1 : x_2 : x_3 : x_4 : x_5 : x_6) := (\epsilon(g)x_0 : x_{g(1)} : x_{g(2)} : x_{g(3)} : x_{g(4)} : x_{g(5)} : x_{g(6)}), \tag{1.8}$$

where $g \in \mathfrak{S}_6$ and $\epsilon(g)$ is the sign of the permutation g, is called *the twisted action*. In contrast, the action (1.6) is called *the natural action*. It is important not to confuse between these two actions, so we strongly recommend the reader to keep an eye on them. Note however, that the actions agree on the alternating group $\mathfrak{A}_6 \subset \mathfrak{S}_6$. Similarly, if G is a subgroup of $\mathfrak{S}_6$, by the natural and the twisted action of G on $\mathscr{Y}$ we mean the restrictions to G of the natural and the twisted actions of $\mathfrak{S}_6$, respectively.

Recall that the group $\mathfrak{S}_6$ has outer automorphisms (in fact, the group $\mathrm{Out}(\mathfrak{S}_6)$ is of order 2; see for instance [Howard et al. 2008]) characterized by the property that they take a transposition in $\mathfrak{S}_6$ to a permutation of cycle type [2, 2, 2]; see Lemma 5.12 for other information about outer automorphisms. If the image of a subgroup $G \subset \mathfrak{S}_6$ under an outer automorphism is not conjugate to G, we call this image a *nonstandard* embedding of G. For instance, we have nonstandard embeddings of $\mathfrak{S}_5$, $\mathfrak{A}_5$, $\mathfrak{S}_4 \times \mathfrak{S}_2$, etc.

The first main result of this paper is a construction of two small resolutions of singularities of the Coble fourfold that are equivariant with respect to maximal proper subgroups of $\mathfrak{S}_6$; note that the rank of the $\mathfrak{S}_6$-invariant Weil divisor class group of $\mathscr{Y}$ (with respect both to the natural and the twisted action of $\mathfrak{S}_6$) equals 1; see Corollary 5.4, hence there are no small resolutions of singularities of $\mathscr{Y}$ equivariant with respect to the entire group $\mathfrak{S}_6$. The varieties $\mathscr{Y}_{4,2}$ and $\mathscr{Y}_{5,1}$ discussed below already appeared in [Farkas and Verra 2016] in a slightly different context. A smooth quintic del Pezzo surface S is unique up to isomorphism, and $\mathrm{Aut}(S) \cong \mathfrak{S}_5$; see for instance [Dolgachev 2012, Section 8.5]; we fix such an isomorphism.

Theorem 1.9. *Consider the twisted $\mathfrak{S}_6$-action* (1.8) *on the Coble fourfold $\mathscr{Y}$*:

(i) *For every nonstandard embedding $\mathfrak{S}_4 \times \mathfrak{S}_2 \hookrightarrow \mathfrak{S}_6$ there is an $\mathfrak{S}_4 \times \mathfrak{S}_2$-equivariant small resolution of singularities*

$$\rho_{4,2} \colon \mathscr{Y}_{4,2} = \mathrm{Bl}_{\boldsymbol{P}_0,\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{P}_3}(\mathbb{P}^2 \times \mathbb{P}^2) \to \mathscr{Y},$$

where $\mathrm{Bl}_{\boldsymbol{P}_0,\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{P}_3}(\mathbb{P}^2 \times \mathbb{P}^2)$ is the blow up of $\mathbb{P}^2 \times \mathbb{P}^2$ at a general quadruple of points $\boldsymbol{P}_0$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{P}_3$ in $\mathbb{P}^2 \times \mathbb{P}^2$.

(ii) *For every nonstandard embedding $\mathfrak{S}_5 \hookrightarrow \mathfrak{S}_6$ there is an $\mathfrak{S}_5$-equivariant small resolution of singularities*

$$\rho_{5,1} \colon \mathscr{Y}_{5,1} = \mathbb{P}_S(\mathscr{U}_3) \to \mathscr{Y},$$

where S is the quintic del Pezzo surface and $\mathscr{U}_3$ is a vector bundle of rank 3 *on S.*

(iii) *The maps $\rho_{4,2}$ and $\rho_{5,1}$ are isomorphisms over the complement of the Cremona–Richmond configuration* $\mathrm{CR} \subset \mathscr{Y}$ *and are uniquely defined up to the Galois involution σ of $\mathscr{Y}$ over $\mathbb{P}^4$ by the above properties.*

(iv) *For every nonstandard embedding $\mathfrak{S}_5 \hookrightarrow \mathfrak{S}_6$ and every subgroup $\mathfrak{S}_4 \subset \mathfrak{S}_5$ there is a unique $\mathfrak{S}_4$-equivariant small birational map $\theta_1 \colon \mathscr{Y}_{5,1} \dashrightarrow \mathscr{Y}_{4,2}$ such that the diagram*

$$\begin{tikzcd} \mathscr{Y}_{5,1} \arrow[rr, dashed, "\theta_1"] \arrow[dr, "\rho_{5,1}"] \arrow[dd, "p"'] & & \mathscr{Y}_{4,2} \arrow[dl, "\rho_{4,2}"'] \arrow[dd, "p_1"] \\ & \mathscr{Y} & \\ S \arrow[rr, "\varphi"] & & \mathbb{P}^2 \end{tikzcd} \tag{1.10}$$

commutes, where $p\colon \mathscr{Y}_{5,1} = \mathbb{P}_S(\mathscr{U}_3) \to S$ is the natural projection, $p_1 : \mathscr{Y}_{4,2} \to \mathbb{P}^2 \times \mathbb{P}^2 \to \mathbb{P}^2$ is the composition of the blow up with the first projection, and φ is the unique $\mathfrak{S}_4$-equivariant birational contraction $S \to \mathbb{P}^2$.

The vector bundle $\mathscr{U}_3$ is described explicitly in Section 2.2.

The Coble fourfold is constructed from the Igusa quartic $X_{1/4}$, but it turns out that it has a very interesting property with respect to all $\mathfrak{S}_6$-invariant quartics. Since the pencil $\{X_t\}$ is generated by $X_{1/4}$ and the double quadric X_∞, we have

$$X_{1/4} \cap X_t = X_\infty \cap X_t \quad \text{for any } t \notin \left\{\tfrac{1}{4}, \infty\right\}.$$

Hence the restriction of $X_{1/4}$ to X_t has multiplicity 2, so that the double cover $\pi : \mathscr{Y} \to \mathbb{P}^4$ splits over X_t. In other words, $\pi^{-1}(X_t)$ is the union of two irreducible components that are isomorphic to X_t and are swapped by the Galois involution (1.7). It is natural here to replace the parameter t in the pencil with the new parameter τ defined by

$$t = \frac{\tau^2+1}{4}, \tag{1.11}$$

and define the subvarieties $\mathscr{X}_\tau \subset \mathscr{Y} \subset \mathbb{P}(2, 1^6)$ by (1.1), (1.5), and the formula

$$x_0 + \frac{\tau}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2) = 0. \tag{1.12}$$

Note that $\mathscr{X}_\tau \subset \mathscr{Y}$ is fixed by the natural action of $\mathfrak{S}_6$, but *is not fixed* by the twisted action. This trivial observation leads to various reductions of groups of symmetries.

With this definition of $\mathscr{X}_\tau$ we have an equality (see Lemma 3.12)

$$\pi^{-1}(X_{(\tau^2+1)/4}) = \mathscr{X}_\tau \cup \mathscr{X}_{-\tau}.$$

The map $\sigma\colon \mathscr{X}_\tau \to \mathscr{X}_{-\tau}$ is an isomorphism and the map $\pi\colon \mathscr{X}_\tau \to X_{(\tau^2+1)/4}$ is an isomorphism for all $\tau \neq \infty$. The map $\pi\colon \mathscr{X}_\infty \to (X_\infty)_{\mathrm{red}}$ is the double covering branched over $(X_\infty)_{\mathrm{red}} \cap X_{1/4}$. Thus, the threefolds $\mathscr{X}_\tau$ have the same singularities as the quartics X_t (except for $\mathscr{X}_\infty$ which becomes smooth away from the $\mathfrak{S}_6$-orbit Σ_{30}; see Remark 3.13).

We consider the preimages of the divisors $\mathscr{X}_\tau$ in the small resolutions $\mathscr{Y}_{5,1}$ and $\mathscr{Y}_{4,2}$:

$$\mathscr{X}_\tau^{5,1} := \rho_{5,1}^{-1}(\mathscr{X}_\tau), \quad \mathscr{X}_\tau^{4,2} := \rho_{4,2}^{-1}(\mathscr{X}_\tau). \tag{1.13}$$

Because of the mixture of the natural and the twisted action, the natural groups of symmetries of the maps $\rho_{5,1}\colon \mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau$ and $\rho_{4,2}\colon \mathscr{X}_\tau^{4,2} \to \mathscr{X}_\tau$ (that is, the groups with respect to which these maps are equivariant) get smaller. In particular, for $\tau \neq 0, \infty$ the first of them reduces to $\mathfrak{A}_5$ and the other to

$$\mathfrak{A}_{4,2} := (\mathfrak{S}_4 \times \mathfrak{S}_2) \cap \mathfrak{A}_6 \cong \mathfrak{S}_4.$$

Our second main result is the following. Recall the discriminant set $\mathfrak{D}$ defined in (1.3).

Theorem 1.14. *The maps*

$$\rho_{5,1}\colon \mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau \quad \textit{and} \quad \rho_{4,2}\colon \mathscr{X}_\tau^{4,2} \to \mathscr{X}_\tau$$

are birational contractions for all τ, and are small for $\tau \neq 0$. Similarly, the maps

$$\pi \circ \rho_{5,1}\colon \mathscr{X}_\tau^{5,1} \to X_{(\tau^2+1)/4} \quad \textit{and} \quad \pi \circ \rho_{4,2}\colon \mathscr{X}_\tau^{4,2} \to X_{(\tau^2+1)/4}$$

are birational contractions for all $\tau \neq \infty$, and are small for $\tau \neq 0, \infty$. Moreover, $\mathscr{X}_\tau^{5,1}$ is smooth (and thus $\rho_{5,1}$ is a small resolution of singularities of $\mathscr{X}_\tau$) unless

$$t = \frac{\tau^2+1}{4} \in \mathfrak{D}.$$

The above maps are equivariant with respect to the following group actions:

	$\rho_{5,1}$ *or* $\pi \circ \rho_{5,1}$	$\rho_{4,2}$ *or* $\pi \circ \rho_{4,2}$
$\tau \neq 0, \infty$	$\mathfrak{A}_5$	$\mathfrak{A}_{4,2}$
$\tau = 0$ *or* $\tau = \infty$	$\mathfrak{S}_5$	$\mathfrak{S}_{4,2}$

where all subgroups of $\mathfrak{S}_6$ are nonstandard and the action is twisted.

We use the above results to construct an interesting (birational) conic bundle structure on the quartics X_t as follows. The fourfold $\mathscr{Y}_{5,1} = \mathbb{P}_S(\mathscr{U}_3)$ by definition comes with a $\mathbb{P}^2$-fibration $p\colon \mathscr{Y}_{5,1} \to S$ over the quintic del Pezzo surface S. We consider its restriction to the threefolds $\mathscr{X}_\tau^{5,1} \subset \mathscr{Y}_{5,1}$. We show that the maps

$$p\colon \mathscr{X}_\tau^{5,1} \to S$$

are $\mathfrak{A}_5$-equivariant conic bundles (and for $\tau = 0, \infty$ they are $\mathfrak{S}_5$-equivariant). We also discuss their properties, and identify their discriminant curves in S with the Wiman–Edge pencil (see Section 3.2 for its definition and the choice of parametrization) of $\mathfrak{A}_5$-invariant divisors from the linear system $|-2K_S|$.

All this is combined in our third main result. Recall that a flat conic bundle $\mathscr{X} \to S$ is called *standard* if both $\mathscr{X}$ and S are smooth and the relative Picard rank $\rho(\mathscr{X}/S)$ equals 1.

Theorem 1.15. *The map $p\colon \mathscr{X}_\tau^{5,1} \to S$ is a flat conic bundle, equivariant with respect to the group $\mathfrak{A}_5$ (for $\tau = 0, \infty$ it is $\mathfrak{S}_5$-equivariant). It is a standard conic bundle unless*

$$t = \frac{\tau^2+1}{4} \in \mathfrak{D}.$$

Its discriminant locus is the curve $\Delta_{s(\tau)} \subset S$ from the Wiman–Edge pencil, where

$$s(\tau) = \frac{\tau^3-\tau}{5\tau^2+3} \tag{1.16}$$

for an appropriate choice of the resolution $\rho_{5,1}$.

We apply the above results in several ways. First, we prove unirationality of $\mathfrak{S}_6$-invariant quartics X_t (see Corollary 4.2). Further, we give a new and uniform proof of rationality and irrationality of the quartics X_t. For $t \notin \mathfrak{D}$ irrationality follows from the description of the intermediate Jacobian of a resolution of singularities of X_t via the Prym variety arising from the conic bundle; see Theorem 4.4. For $t \in \mathfrak{D}$ we show that the conic bundle can be transformed birationally into the product $S \times \mathbb{P}^1$, hence X_t is rational; see Theorem 4.6. Finally, we describe the class groups $\mathrm{Cl}(X_t)$ of Weil divisors of the quartics X_t as $\mathfrak{S}_6$-representations (see Theorem 5.1), and discuss G-Sarkisov links centered at these quartics for some subgroups $G \subset \mathfrak{S}_6$. We also prove unirationality and irrationality of the threefold $\mathscr{X}_\infty$, and describe its class group as an $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$-representation.

The plan of our paper is the following. In Section 2 we construct the resolutions of the Coble fourfold $\mathscr{Y}$ and prove Theorem 1.9. In Section 3 we discuss the conic bundle structures on the $\mathfrak{S}_6$-invariant quartics induced by the resolutions of the Coble fourfold, and prove Theorems 1.14 and 1.15. In Section 4 we prove rationality and irrationality of the quartics X_t, and in Section 5 we describe the $\mathfrak{S}_6$-action on their class groups. In the Appendix we discuss the Cremona–Richmond configuration $\mathrm{CR} = \mathrm{Sing}(X_{1/4})$ of 15 lines in $\mathbb{P}^4$ and show that such configuration is unique up to a projective transformation of $\mathbb{P}^4$.

Throughout the paper $\Bbbk$ denotes an algebraically closed field of characteristic zero; however, many constructions do not use the assumption that the field is algebraically closed. By $\boldsymbol{\mu}_n$ we denote the cyclic group of order n. Furthermore, we denote by

$$\mathfrak{S}_{n_1,n_2} \cong \mathfrak{S}_{n_1} \times \mathfrak{S}_{n_2} \subset \mathfrak{S}_{n_1+n_2} \quad \text{and} \quad \mathfrak{A}_{n_1,n_2} = \mathfrak{A}_{n_1+n_2} \cap \mathfrak{S}_{n_1,n_2} \subset \mathfrak{A}_{n_1+n_2} \tag{1.17}$$

the subgroup of $\mathfrak{S}_{n_1+n_2}$ that consists of permutations preserving the subsets of the first n_1 and the last n_2 indices, and its intersection with the alternating group $\mathfrak{A}_{n_1+n_2} \subset \mathfrak{S}_{n_1+n_2}$. Note that $\mathfrak{A}_{n-2,2} \cong \mathfrak{S}_{n-2}$.

2. Small resolutions of the Coble fourfold

Recall that the fourfold $\mathscr{Y}$ is defined by (1.5) as the double cover of $\mathbb{P}^4$ (considered as the hyperplane (1.1) in $\mathbb{P}^5$) branched over the Igusa quartic (1.4). It comes with the natural and the twisted actions of the symmetric group $\mathfrak{S}_6$; see (1.6) and (1.8), the double covering $\pi\colon \mathscr{Y} \to \mathbb{P}^4$ and its Galois involution $\sigma\colon \mathscr{Y} \to \mathscr{Y}$; see (1.7), commuting with both actions of $\mathfrak{S}_6$.

The fourfold $\mathscr{Y}$ has been studied by Coble [1915; 1916; 1917]. He showed that $\mathscr{Y}$ is a compactification of the moduli space of ordered sets of 6 points in the projective plane. A modern treatment of $\mathscr{Y}$ has been given in [Dolgachev and Ortland 1988; Matsumoto et al. 1992; Hunt 1996; Howard et al. 2008]; see also [Bauer and Verra 2010]. In particular, Dolgachev and Ortland [1988] proved that $\mathscr{Y}$ can be obtained as the GIT-quotient $(\mathbb{P}^2)^6 /\!\!/ \mathrm{SL}_3(\Bbbk)$ with respect to the diagonal action of $\mathrm{SL}_3(\Bbbk)$. In [Clingher et al. 2019] the variety $\mathscr{Y}$ came up in the study of moduli spaces of K3 surfaces. Hunt [1996] called it the Coble variety (he also denoted it by $\mathscr{Y}$). In the current paper we prefer to call $\mathscr{Y}$ the *Coble fourfold*.

Since the Coble fourfold $\mathscr{Y}$ is singular, it is interesting to construct its resolution of singularities that would be natural from the geometric point of view. One interesting resolution was provided by Naruki [1982]; see also [Hacking et al. 2009; Dolgachev et al. 2005, Section 2]. It has plenty of important properties due to its interpretation as a moduli space of cubic surfaces. However, it is quite big (it has a horde of exceptional divisors). On the other hand, one can observe that the variety $\mathscr{Y}$ has non-$\mathbb{Q}$-factorial singularities, so we can hope to have a nice *small* resolution (i.e., with exceptional locus of codimension 2).

In this section we construct two small resolutions of singularities of $\mathscr{Y}$; one is equivariant with respect to the subgroup $\mathfrak{S}_{4,2} \subset \mathfrak{S}_6$ and another is equivariant with respect to the subgroup $\mathfrak{S}_5 \subset \mathfrak{S}_6$. Note that in both cases a *nonstandard* embedding of the subgroup is used (equivalently, a standard embedding is composed with an outer automorphism of $\mathfrak{S}_6$) and in both cases we consider the *twisted* action of $\mathfrak{S}_6$ on $\mathscr{Y}$.

2.1. ***Blow up of*** $\mathbb{P}^2 \times \mathbb{P}^2$**.** Let W_3 be the irreducible three-dimensional representation of the symmetric group $\mathfrak{S}_4$ with the nontrivial determinant, i.e., a summand of the four-dimensional permutation representation. Explicitly, $\mathrm{W}_3 \cong \mathrm{R}(3, 1)$ in the notation of [Fulton and Harris 1991, Section 4.1]. Choose a $\mathfrak{S}_4$-orbit of length 4

$$\{P_0, P_1, P_2, P_3\} \subset \mathbb{P}(\mathrm{W}_3) \cong \mathbb{P}^2.$$

In appropriate coordinates such quadruple can be written as

$$P_0 = (1:1:1), \quad P_1 = (1:0:0), \quad P_2 = (0:1:0), \quad P_3 = (0:0:1). \tag{2.1}$$

Denote by

$$\overline{P_i P_j} \subset \mathbb{P}(\mathrm{W}_3), \quad 0 \leqslant i < j \leqslant 3,$$

the line passing through the points P_i and P_j.

Consider the diagonal action of $\mathfrak{S}_4$ on $\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ and the diagonal quadruple

$$\boldsymbol{P} = \{\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3\} \subset \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3), \quad \boldsymbol{P}_i = (P_i, P_i).$$

Note that $\boldsymbol{P}$ is an $\mathfrak{S}_4$-orbit. The vector space $\mathrm{W}_3 \otimes \mathrm{W}_3$ can be regarded as a representation of the group $\mathfrak{S}_{4,2}$; see (1.17), where $\mathfrak{S}_4$ acts diagonally and the nontrivial element of $\mathfrak{S}_2$ interchanges the factors. The linear span of the points $\boldsymbol{P}_i$ in $\mathbb{P}(\mathrm{W}_3 \otimes \mathrm{W}_3)$ induces an embedding of the permutation representation $\Bbbk^4$ of $\mathfrak{S}_4$ (with the trivial action of $\mathfrak{S}_2$) into $\mathrm{W}_3 \otimes \mathrm{W}_3$. We denote by

$$\mathrm{W}_5 := (\mathrm{W}_3 \otimes \mathrm{W}_3)/\Bbbk^4 \tag{2.2}$$

the quotient five-dimensional representation of $\mathfrak{S}_{4,2}$. Note that as a representation of $\mathfrak{S}_4$ it is the direct sum $\mathrm{W}_5|_{\mathfrak{S}_4} \cong \mathrm{R}(2, 2) \oplus \mathrm{R}(2, 1, 1)$; here we again use the (standard) notation oi [Fulton and Harris 1991, Section 4.1].

The linear projection $\mathrm{W}_3 \otimes \mathrm{W}_3 \to \mathrm{W}_5$ induces a rational map

$$\bar{\pi}_{4,2}\colon \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \hookrightarrow \mathbb{P}(\mathrm{W}_3 \otimes \mathrm{W}_3) \dashrightarrow \mathbb{P}(\mathrm{W}_5).$$

Note that the center of this projection is the linear span of the orbit $\boldsymbol{P}$ in $\mathbb{P}(\mathrm{W}_3 \otimes \mathrm{W}_3)$, which intersects $\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ exactly by $\boldsymbol{P}$. Therefore, to regularize the map $\bar{\pi}_{4,2}$ we should consider the blow up $\mathscr{Y}_{4,2}$ of $\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ in the quadruple $\boldsymbol{P}$

$$\mathscr{Y}_{4,2} := \mathrm{Bl}_{\boldsymbol{P}_0,\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{P}_3}(\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)) \xrightarrow{\beta} \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \tag{2.3}$$

with β being the blow up morphism. This induces a commutative diagram:

$$\begin{tikzcd} & \mathscr{Y}_{4,2} \arrow[dl, "\beta"'] \arrow[dr, "\pi_{4,2}"] & \\ \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \arrow[rr, dashed, "\bar{\pi}_{4,2}"] & & \mathbb{P}(\mathrm{W}_5) \end{tikzcd} \tag{2.4}$$

By construction the fourfold $\mathscr{Y}_{4,2}$ is smooth and carries a faithful action of $\mathfrak{S}_{4,2}$. The above diagram is $\mathfrak{S}_{4,2}$-equivariant.

We are going to show that the map $\pi_{4,2}\colon \mathscr{Y}_{4,2} \to \mathbb{P}(\mathrm{W}_5)$ defined by the diagram (2.4) factors through the Coble fourfold; more precisely, $\pi_{4,2}$ factors as a composition

$$\mathscr{Y}_{4,2} \xrightarrow{\rho_{4,2}} \mathscr{Y} \xrightarrow{\pi} \mathbb{P}(\mathrm{W}_5),$$

with $\rho_{4,2}$ being a small $\mathfrak{S}_{4,2}$-equivariant resolution of singularities. We accomplish this in two steps.

First, consider the linear projection

$$\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \hookrightarrow \mathbb{P}(\mathrm{W}_3 \otimes \mathrm{W}_3) \dashrightarrow \mathbb{P}^5$$

from the linear span of the points $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, and $\boldsymbol{P}_3$; as before, the latter linear span intersects $\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ exactly by the triple $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{P}_3$. If $(u_1 : u_2 : u_3)$ and $(v_1 : v_2 : v_3)$ are homogeneous coordinates on the first and the second factors of $\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ such that (2.1) holds, this map is given by

$$((u_1 : u_2 : u_3), (v_1 : v_2 : v_3)) \mapsto (u_2v_3 : u_3v_1 : u_1v_2 : u_3v_2 : u_1v_3 : u_2v_1), \tag{2.5}$$

and it is easy to describe its structure. We denote by y_1, y_2, y_3, z_1, z_2, and z_3 the homogeneous coordinates on $\mathbb{P}^5$, so that the right-hand side of (2.5) is the point $(y_1 : y_2 : y_3 : z_1 : z_2 : z_3)$.

Lemma 2.6. *The linear projection $\mathbb{P}(\mathrm{W}_3)\times\mathbb{P}(\mathrm{W}_3) \dashrightarrow \mathbb{P}^5$ with center in the span of the points $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{P}_3$ induces an $\mathfrak{S}_{3,2}$-equivariant commutative diagram*

$$\begin{tikzcd}[column sep=small]
 & \mathrm{Bl}_{\boldsymbol{P}_1,\boldsymbol{P}_2,\boldsymbol{P}_3}(\mathbb{P}(\mathrm{W}_3)\times\mathbb{P}(\mathrm{W}_3)) \arrow[dl, "\beta'"'] \arrow[dr, "\rho'_{4,2}"] & & \\
\mathbb{P}(\mathrm{W}_3)\times\mathbb{P}(\mathrm{W}_3) \arrow[rr, dashed] & & \mathscr{Y}'_{4,2} \arrow[r, hook] & \mathbb{P}^5
\end{tikzcd}$$

where β' is the blow up, $\mathscr{Y}'_{4,2}\subset\mathbb{P}^5$ is a singular cubic hypersurface given by the equation

$$y_1y_2y_3 = z_1z_2z_3, \tag{2.7}$$

and $\rho'_{4,2}$ is a small birational contraction. The map $\rho'_{4,2}$ contracts

- *the proper transforms of the six planes $\mathbb{P}(\mathrm{W}_3)\times P_i$ and $P_i\times\mathbb{P}(\mathrm{W}_3)$, $1\leqslant i\leqslant 3$, and*
- *the proper transforms of the three quadrics $\overline{P_iP_j}\times\overline{P_iP_j}$, $1\leqslant i<j\leqslant 3$,*

onto nine lines L_{ij}, $1\leqslant i,\ j\leqslant 3$, given in $\mathbb{P}^5$ by the equations

$$y_k = z_l = 0,\quad k\neq i,\ l\neq j.$$

Moreover, $\rho'_{4,2}$ is an isomorphism over the complement of the lines L_{ij}. Finally, the map $\rho'_{4,2}\circ(\beta')^{-1}$ takes the point $\boldsymbol{P}_0$ to the point $\boldsymbol{P}'_0=(1:1:1:1:1:1)\in\mathscr{Y}'_{4,2}$.

Proof. The map is toric, so everything is easy to describe. We skip the actual computation which is straightforward but tedious. □

The cubic fourfold (2.7) is known as *Perazzo primal*, [Dolgachev 2012, Exercise 9.16; Looijenga 2009, Section 6].

Using the equation (2.7) one can easily check that the union of the nine lines L_{ij} is the singular locus of the cubic $\mathscr{Y}'_{4,2}$.

The second step is to project the cubic $\mathscr{Y}'_{4,2}$ from the point $\boldsymbol{P}'_0$.

Lemma 2.8. *The linear projection $\bar{\pi}'_{4,2}\colon\mathscr{Y}'_{4,2}\dashrightarrow\mathbb{P}(\mathrm{W}_5)$ from the point $\boldsymbol{P}'_0$ defines a regular map $\pi''_{4,2}\colon\ \mathrm{Bl}_{\boldsymbol{P}'_0}(\mathscr{Y}'_{4,2})\to\mathbb{P}(\mathrm{W}_5)$ that fits into a commutative diagram*

$$\begin{tikzcd}
 & \mathrm{Bl}_{\boldsymbol{P}'_0}(\mathscr{Y}'_{4,2}) \arrow[dl, "\beta''"'] \arrow[d, "\pi''_{4,2}"] \arrow[r, "\rho''_{4,2}"] & \mathscr{Y} \arrow[dl, "\pi"] \\
\mathscr{Y}'_{4,2} \arrow[r, dashed, "\bar{\pi}'_{4,2}"] & \mathbb{P}(\mathrm{W}_5) &
\end{tikzcd} \tag{2.9}$$

where $\mathscr{Y}$ is the Coble fourfold, $\pi:\mathscr{Y}\to\mathbb{P}(\mathrm{W}_5)$ is the double covering, and $\rho''_{4,2}$ is a small birational morphism. Furthermore, the exceptional locus of $\rho''_{4,2}$ is the union of proper transforms of the six planes $\Pi_w\subset\mathscr{Y}'_{4,2}$ given by the equations

$$z_i = y_{w(i)},\quad 1\leqslant i\leqslant 3,$$

indexed by all bijections $w\colon \{1,2,3\} \to \{1,2,3\}$; *the map* $\rho''_{4,2}$ *contracts them onto six lines in* $\mathscr{Y}$ (*i.e., rational curves that are isomorphically projected to lines in* $\mathbb{P}(\mathrm{W}_5)$), *and is an isomorphism over the complement of those.*

Proof. Note that the point $\boldsymbol{P}'_0$ is a smooth point of the cubic $\mathscr{Y}'_{4,2}$, so the projection from it factors through a double covering of $\mathbb{P}(\mathrm{W}_5)$; in fact, this is the Stein factorization for the morphism $\pi''_{4,2}$. We have to identify its branch divisor with the Igusa quartic.

Take a point

$$(y_i : z_i) = (y_1 : y_2 : y_3 : z_1 : z_2 : z_3)$$

in $\mathbb{P}^5$ which is different from $\boldsymbol{P}'_0$. The line $M_{(y_i:z_i)}$ in $\mathbb{P}(\mathrm{W}_5)$ passing through the point $(y_i : z_i)$ and the point $\boldsymbol{P}'_0$ can be parametrized as

$$M_{(y_i:z_i)} = \{(\lambda+\mu y_1 : \lambda+\mu y_2 : \lambda+\mu y_3 : \lambda+\mu z_1 : \lambda+\mu z_2 : \lambda+\mu z_3)\}, \tag{2.10}$$

where λ and μ are considered as homogeneous coordinates on this line. Substituting this parametrization into (2.7), we see that the intersection of $M_{(y_i:z_i)}$ with the cubic $\mathscr{Y}'_{4,2}$ is given by the equation

$$(\lambda+\mu y_1)(\lambda+\mu y_2)(\lambda+\mu y_3) = (\lambda+\mu z_1)(\lambda+\mu z_2)(\lambda+\mu z_3).$$

Expanding both sides and canceling the factor μ that corresponds to the intersection point $\boldsymbol{P}'_0$, we can rewrite the above equation as

$$(\boldsymbol{s}_1(y)-\boldsymbol{s}_1(z))\lambda^2 + (\boldsymbol{s}_2(y)-\boldsymbol{s}_2(z))\lambda\mu + (\boldsymbol{s}_3(y)-\boldsymbol{s}_3(z))\mu^2 = 0, \tag{2.11}$$

where $\boldsymbol{s}_d$ denotes the elementary symmetric polynomial of degree d. Restricting (2.11) to the hyperplane

$$y_1+y_2+y_3+z_1+z_2+z_3 = 0, \tag{2.12}$$

which is identified by the linear projection $\bar{\pi}'_{4,2}$ from the point $\boldsymbol{P}'_0$ with the space $\mathbb{P}(\mathrm{W}_5)$, we obtain the equation of the double cover over $\mathbb{P}(\mathrm{W}_5)$ we are interested in (embedded into the projectivization of the vector bundle $\mathscr{O}_{\mathbb{P}(\mathrm{W}_5)} \oplus \mathscr{O}_{\mathbb{P}(\mathrm{W}_5)}(-1)$ over $\mathbb{P}(\mathrm{W}_5)$). The branch divisor of $\bar{\pi}'_{4,2}$ is given in the hyperplane (2.12) by the discriminant of the quadratic (2.11)

$$(\boldsymbol{s}_2(y)-\boldsymbol{s}_2(z))^2 - 4(\boldsymbol{s}_1(y)-\boldsymbol{s}_1(z))(\boldsymbol{s}_3(y)-\boldsymbol{s}_3(z)) = 0. \tag{2.13}$$

Let us show that the quartic $X'' \subset \mathbb{P}^4$ defined by equations (2.12) and (2.13) is isomorphic to the Igusa quartic; this will identify the double covering with the Coble fourfold in a way respecting the projection to $\mathbb{P}^4$, that is, ensuring that the upper right triangle in diagram (2.9) is commutative.

To do this we use the following substitutions:

$$\begin{aligned}
x_1 &= y_1 - \tfrac{2}{3}\boldsymbol{s}_1(y) + \tfrac{1}{3}\boldsymbol{s}_1(z), & x_4 &= z_1 + \tfrac{1}{3}\boldsymbol{s}_1(y) - \tfrac{2}{3}\boldsymbol{s}_1(z),\\
x_2 &= y_2 - \tfrac{2}{3}\boldsymbol{s}_1(y) + \tfrac{1}{3}\boldsymbol{s}_1(z), & x_5 &= z_2 + \tfrac{1}{3}\boldsymbol{s}_1(y) - \tfrac{2}{3}\boldsymbol{s}_1(z),\\
x_3 &= y_3 - \tfrac{2}{3}\boldsymbol{s}_1(y) + \tfrac{1}{3}\boldsymbol{s}_1(z), & x_6 &= z_3 + \tfrac{1}{3}\boldsymbol{s}_1(y) - \tfrac{2}{3}\boldsymbol{s}_1(z).
\end{aligned} \tag{2.14}$$

They express the composition of the projection $\bar{\pi}'_{4,2}$ with a particular identification of its target space $\mathbb{P}(\mathrm{W}_5)$ with the hyperplane (1.1) in $\mathbb{P}^5$. A direct verification shows that substituting these expressions into (1.4) of the Igusa quartic we get (2.13). This proves that (2.13) is isomorphic to the cone over the Igusa quartic with the vertex at the point $\boldsymbol{P}'_0$, hence its intersection with (2.12) is isomorphic to the Igusa quartic.

Finally, we describe the exceptional locus of the projection $\pi''_{4,2}$. Clearly, it is the union of those lines $M_{(y_i:z_i)}$ that are contained in the cubic $\mathscr{Y}'_{4,2}$, i.e., the subvariety of those points $(y_i : z_i)$ for which (2.11) is identically zero. This condition can be rewritten as

$$\boldsymbol{s}_1(y) - \boldsymbol{s}_1(z) = \boldsymbol{s}_2(y) - \boldsymbol{s}_2(z) = \boldsymbol{s}_3(y) - \boldsymbol{s}_3(z) = 0$$

Of course, this is equivalent to $(y_i : z_i) \in \Pi_w$ for some permutation w. Thus the exceptional locus is the union of the proper transforms of the planes Π_w. Each of these planes passes through $\boldsymbol{P}'_0$, hence is contracted onto a line in $\mathbb{P}^4 \cong \mathbb{P}(\mathrm{W}_5)$. □

Remark 2.15. There is also a computation-free way to identify the branch divisor X'' of the map $\pi''_{4,2}$ with the Igusa quartic. Indeed, note that the singular locus of X'' contains 15 lines (the images of the 9 singular lines L_{ij} of $\mathscr{Y}'_{4,2}$ and the images of the 6 planes Π_w), then check that they form a Cremona–Richmond configuration (e.g., by using Theorem A.8), and then apply Corollary A.14.

Remark 2.16. Using (2.11) it is easy to write the (birational) involution of the double covering $\mathscr{Y}'_{4,2} \dashrightarrow \mathbb{P}^4$ explicitly. Indeed, choose a point $(y_i : z_i) = (y_1 : y_2 : y_3 : z_1 : z_2 : z_3)$ on the cubic $\mathscr{Y}'_{4,2} \subset \mathbb{P}^5$ different from $\boldsymbol{P}'_0$. Using the parametrization (2.10), we see that the point $(y_i : z_i)$ corresponds to $\lambda = 0$. Keeping in mind that $\boldsymbol{s}_3(y) = \boldsymbol{s}_3(z)$ at our point $(y_i : z_i)$, and finding the second root of the (2.11) in λ/μ, we conclude that the involution of the double covering $\mathscr{Y}'_{4,2} \dashrightarrow \mathbb{P}^4$ is given by

$$(y_i : z_i) \mapsto ((\boldsymbol{s}_1(y) - \boldsymbol{s}_1(z))y_i - (\boldsymbol{s}_2(y) - \boldsymbol{s}_2(z)) : (\boldsymbol{s}_1(y) - \boldsymbol{s}_1(z))z_i - (\boldsymbol{s}_2(y) - \boldsymbol{s}_2(z))). \tag{2.17}$$

Furthermore, the induced birational involution of $\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ can be written as

$$\begin{aligned}\bar{\sigma}_{4,2}\colon &((u_1 : u_2 : u_3), (v_1 : v_2 : v_3)) \\ &\mapsto \left(\left(\frac{v_2 - v_3}{\det\left(\begin{smallmatrix} u_2 & u_3 \\ v_2 & v_3 \end{smallmatrix}\right)} : \frac{v_3 - v_1}{\det\left(\begin{smallmatrix} u_3 & u_1 \\ v_3 & v_1 \end{smallmatrix}\right)} : \frac{v_1 - v_2}{\det\left(\begin{smallmatrix} u_1 & u_2 \\ v_1 & v_2 \end{smallmatrix}\right)}\right), \left(\frac{u_2 - u_3}{\det\left(\begin{smallmatrix} u_2 & u_3 \\ v_2 & v_3 \end{smallmatrix}\right)} : \frac{u_3 - u_1}{\det\left(\begin{smallmatrix} u_3 & u_1 \\ v_3 & v_1 \end{smallmatrix}\right)} : \frac{u_1 - u_2}{\det\left(\begin{smallmatrix} u_1 & u_2 \\ v_1 & v_2 \end{smallmatrix}\right)}\right)\right);\end{aligned} \tag{2.18}$$

to see this one can just compose (2.5) with (2.17) and observe that it gives the same result as a composition of (2.18) with (2.5). Similarly, we deduce from (2.13) that the ramification divisor of the map $\bar{\pi}_{4,2}$ is given by the equation

$$\boldsymbol{s}_2(u_2v_3, u_3v_1, u_1v_2) = \boldsymbol{s}_2(u_3v_2, u_1v_3, u_2v_1),$$

that can be compactly rewritten as

$$\det\left(\begin{smallmatrix} u_1v_1 & u_2v_2 & u_3v_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{smallmatrix}\right) = 0. \tag{2.19}$$

This gives a determinantal representation of a threefold birational to the Igusa quartic.

Combining the results of Lemmas 2.6 and 2.8 we obtain a commutative diagram

$$
\begin{tikzcd}
\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \arrow[dd, dashed, "\bar{\pi}_{4,2}"'] & \mathrm{Bl}_{\boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3}(\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)) \arrow[l, "\beta'"'] \arrow[d, "\rho'_{4,2}"'] & \mathscr{Y}_{4,2} \arrow[l] \arrow[d] \\
 & \mathscr{Y}'_{4,2} \arrow[dl, dashed, "\bar{\pi}'_{4,2}"'] & \mathrm{Bl}_{\boldsymbol{P}_0}(\mathscr{Y}'_{4,2}) \arrow[l, "\beta''"'] \arrow[d, "\rho''_{4,2}"] \\
\mathbb{P}(\mathrm{W}_5) & & \mathscr{Y} \arrow[ll, "\pi"']
\end{tikzcd}
\tag{2.20}
$$

where the upper right square is Cartesian and the composition $\mathscr{Y}_{4,2} \to \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ of the upper horizontal arrows is the blow up map β.

Proposition 2.21. *The linear projection $\bar{\pi}_{4,2} \colon \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \dashrightarrow \mathbb{P}(\mathrm{W}_5)$ with center in the span of the points $\boldsymbol{P}_0$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, and $\boldsymbol{P}_3$ gives rise to a commutative diagram*

$$
\begin{tikzcd}
 & \mathscr{Y}_{4,2} \arrow[rr, "\rho_{4,2}"] \arrow[dl, "\beta"'] \arrow[dr, "\pi_{4,2}"] & & \mathscr{Y} \arrow[dl, "\pi"] \\
\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \arrow[rr, dashed, "\bar{\pi}_{4,2}"] & & \mathbb{P}(\mathrm{W}_5) &
\end{tikzcd}
\tag{2.22}
$$

where $\rho_{4,2}$ is a small resolution of singularities defined uniquely up to a composition with the Galois involution $\sigma \colon \mathscr{Y} \to \mathscr{Y}$. The map $\rho_{4,2}$ contracts

- *the proper transforms of the eight planes $\mathbb{P}(\mathrm{W}_3) \times P_i$ and $P_i \times \mathbb{P}(\mathrm{W}_3)$, $0 \leqslant i \leqslant 3$,*
- *the proper transforms of the six quadrics $\overline{P_i P_j} \times \overline{P_i P_j}$, $0 \leqslant i < j \leqslant 3$, and*
- *the proper transform of the diagonal $\mathbb{P}(\mathrm{W}_3) \hookrightarrow \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$,*

and is an isomorphisms on the complement of those. Moreover, the morphism $\pi_{4,2}$ induces a nonstandard embedding $\mathfrak{S}_{4,2} \to \mathfrak{S}_6$ such that $\rho_{4,2}$ is $\mathfrak{S}_{4,2}$-equivariant with respect to the twisted action of $\mathfrak{S}_{4,2}$ on $\mathscr{Y}$.

Proof. We define the map $\rho_{4,2}$ as the composition of the right vertical arrows in (2.20). Its uniqueness up to σ is evident. We note that the composition

$$\rho_{4,2} \circ \beta^{-1} \colon \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \dashrightarrow \mathscr{Y} \subset \mathbb{P}(2, 1^6)$$

can be defined by explicit formulas:

$$
\begin{aligned}
x_0 &= -u_1u_3v_1v_2 - u_1u_2v_2v_3 - u_2u_3v_1v_3 + u_1u_2v_1v_3 + u_2u_3v_1v_2 + u_1u_3v_2v_3, \\
x_1 &= \tfrac{1}{3}(u_2v_3 - 2u_3v_1 - 2u_1v_2 + u_3v_2 + u_1v_3 + u_2v_1), \\
x_2 &= \tfrac{1}{3}(-2u_2v_3 + u_3v_1 - 2u_1v_2 + u_3v_2 + u_1v_3 + u_2v_1), \\
x_3 &= \tfrac{1}{3}(-2u_2v_3 - 2u_3v_1 + u_1v_2 + u_3v_2 + u_1v_3 + u_2v_1), \\
x_4 &= \tfrac{1}{3}(u_2v_3 + u_3v_1 + u_1v_2 + u_3v_2 - 2u_1v_3 - 2u_2v_1), \\
x_5 &= \tfrac{1}{3}(u_2v_3 + u_3v_1 + u_1v_2 - 2u_3v_2 + u_1v_3 - 2u_2v_1), \\
x_6 &= \tfrac{1}{3}(u_2v_3 + u_3v_1 + u_1v_2 - 2u_3v_2 - 2u_1v_3 + u_2v_1).
\end{aligned}
\tag{2.23}
$$

Indeed, x_0 defines in $\mathscr{Y}$ the ramification divisor of the map π, hence its pullback to $\mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$ coincides (up to a scalar) with the equation (2.19) of the ramification divisor of $\bar{\pi}_{4,2}$. The pullbacks of $x_1, \dots, x_6$ are given by the composition of (2.14) and (2.5), which gives the required formulas. Substituting those into (1.5), we see that the scalar in the formula for x_0 is ± 1. So, (2.23) gives one of the two maps $\rho_{4,2}$, while the other sign choice gives $\sigma \circ \rho_{4,2}$.

For the description of the exceptional locus of $\rho_{4,2}$ we combine the results of Lemmas 2.6 and 2.8 with the simple observation (using (2.5)) that the map $\rho'_{4,2} \circ \beta'^{-1}$ from (2.20) takes the two planes $\mathbb{P}(\mathrm{W}_3) \times P_0$ and $P_0 \times \mathbb{P}(\mathrm{W}_3)$ to the planes Π_w, where w are cycles of length 3; takes the three quadrics $\overline{P_0 P_i} \times \overline{P_0 P_i}$ to the planes Π_w, where w are transpositions; and takes the diagonal to Π_w, where w is the identity permutation.

The space W_5 by definition (2.2) comes with an $\mathfrak{S}_{4,2}$ action, such that the map $\pi_{4,2} \colon \mathscr{Y}_{4,2} \to \mathbb{P}(\mathrm{W}_5)$ obtained by resolving the indeterminacy of the linear projection $\bar{\pi}_{4,2}$ is $\mathfrak{S}_{4,2}$-equivariant. It follows that its branch divisor, which was shown to be the Igusa quartic $X_{1/4}$, is invariant under this action. On the other hand, it is well known that $\mathrm{Aut}(X_{1/4}) \cong \mathfrak{S}_6$ (this follows for instance from [Finkelnberg 1987, Section 3; Hunt 1996, Proposition 3.3.1]; see also Lemma 3.4 below). Thus, we obtain an embedding $\mathfrak{S}_{4,2} \hookrightarrow \mathfrak{S}_6$.

Moreover, for every element $g \in \mathfrak{S}_{4,2}$ the conjugation of the diagram (2.22) by g gives a diagram of the same form. Since $\rho_{4,2}$ is uniquely defined up to σ, we obtain an equality

$$g \circ \rho_{4,2} \circ g^{-1} = \sigma^{k(g)} \circ \rho_{4,2},$$

where $k \colon \mathfrak{S}_{4,2} \to \mathbb{Z}/2\mathbb{Z}$ is a group homomorphism. Using the explicit expression for x_0 provided by (2.19) it is easy to see that transpositions in the group $\mathfrak{S}_{4,2}$ change the sign of x_0. This means that k is the homomorphism of parity $\mathfrak{S}_6 \to \mathbb{Z}/2\mathbb{Z}$ restricted to $\mathfrak{S}_{4,2}$, which means that the map $\rho_{4,2}$ is equivariant with respect to the twisted action (1.8) of $\mathfrak{S}_6$ on $\mathscr{Y}$.

Finally, to show that the embedding $\mathfrak{S}_{4,2} \hookrightarrow \mathfrak{S}_6$ is nonstandard, we use (2.23) to observe that transpositions in $\mathfrak{S}_{4,2}$ go to permutations of cycle type [2, 2, 2] in $\mathfrak{S}_6$. Alternatively, we could notice that the restriction of the representation (1.1) with respect to a standard embedding $\mathfrak{S}_4 \hookrightarrow \mathfrak{S}_6$ decomposes as a direct sum of three irreducible representations of $\mathfrak{S}_4$ (see (5.11) and Lemma 5.12), while (2.2) is the sum of two irreducibles. □

Let us emphasize again that there are exactly two maps $\rho_{4,2}$ that fit into commutative diagram (2.22): the first is given by (2.23) and the second is obtained by its composition with σ, i.e., by the change of sign of x_0. The particular choice (2.23) will lead us to a particular choice of the map $\rho_{5,1}$ in the next subsection.

We write down here a simple consequence of Proposition 2.21 concerning the Weil divisor class group of the Coble fourfold.

Corollary 2.24. *One has* $\operatorname{rk} \mathrm{Cl}(\mathscr{Y}) = 6$.

Proof. Since the map $\rho_{4,2}\colon \mathscr{Y}_{4,2} \to \mathscr{Y}$ is a small resolution of singularities, it induces an isomorphism $\mathrm{Cl}(\mathscr{Y}) \cong \mathrm{Pic}(\mathscr{Y}_{4,2})$, and since $\mathscr{Y}_{4,2}$ is the blow up of $\mathbb{P}^2 \times \mathbb{P}^2$ in 4 points, its Picard rank equals 6. □

In Theorem 5.1 we will describe the action of the group $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$ on $\mathrm{Cl}(\mathscr{Y}) \otimes \mathbb{Q}$.

Remark 2.25. For each three-element subset $I \subset \{1, \dots, 6\}$ denote by $\bar{I} \subset \{1, \dots, 6\}$ its complement. Consider the hyperplane $H_I \subset \mathbb{P}^4$ defined in (1.1) by the equation

$$\sum_{i\in I} x_i = 0. \tag{2.26}$$

Note that $H_I = H_{\bar{I}}$. In the terminology of the Appendix these are the ten jail hyperplanes (A.5) of the Cremona–Richmond configuration. The preimage of H_I on $\mathscr{Y}$ splits as the union of two irreducible components. Indeed, consider the subvariety $\mathscr{H}_I \subset \mathscr{Y}$ defined by the (2.26) together with the equation

$$x_0 + \tfrac{1}{2}\Bigl(2\sum_{i\in I} x_i^2 - \sum_{i\in \bar{I}} x_i^2\Bigr) = 0. \tag{2.27}$$

Then it is easy to check that

$$\pi^{-1}(H_I) = \pi^{-1}(H_{\bar{I}}) = \mathscr{H}_I \cup \mathscr{H}_{\bar{I}}.$$

An even easier way to see this splitting is provided by the morphism $\rho_{4,2}$. Indeed, using formulas (2.23) one can check that the preimages on $\mathbb{P}^2 \times \mathbb{P}^2$ of the six hyperplanes H_{124}, H_{125}, H_{134}, H_{136}, H_{235}, and H_{236} are divisors given by equations

$$\begin{aligned}
&(u_1 - u_3)v_2 = 0, \quad u_1(v_2 - v_3) = 0, \quad u_3(v_1 - v_2) = 0,\\
&(u_2 - u_3)v_1 = 0, \quad (u_1 - u_2)v_3 = 0, \quad u_2(v_1 - v_3) = 0,
\end{aligned}$$

respectively. Each of these divisors is a union of two irreducible components, and each component is the product $\overline{P_iP_j} \times \mathbb{P}^2$ or $\mathbb{P}^2 \times \overline{P_iP_j}$ for appropriate i and j. Note that the action of $\mathfrak{S}_{4,2}$ on the set of all twelve of these irreducible components is transitive. For each I denote

$$\mathscr{H}_I^{4,2} := \rho_{4,2}^{-1}(\mathscr{H}_I).$$

Therefore, if I is one of the above six triples or one of their complements, then $\beta(\mathscr{H}_I^{4,2})$ is one of the above twelve components, hence these divisors $\mathscr{H}_I^{4,2}$ form a single $\mathfrak{S}_{4,2}$-orbit.

Similarly, formulas (2.23) show that the preimages on $\mathbb{P}^2 \times \mathbb{P}^2$ of the remaining four hyperplanes H_{123}, H_{156}, H_{246}, and H_{345} are irreducible divisors singular at the points $\boldsymbol{P}_0$, $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, and $\boldsymbol{P}_3$, respectively. This means that for each of the above four triples I the preimage $\pi_{4,2}^{-1}(H_I)$ of H_I on $\mathscr{Y}_{4,2}$ consists of two irreducible components, one of them being the exceptional divisor of the blow up β over the corresponding point $\boldsymbol{P}_r$. A straightforward computation shows that

$$\mathscr{H}_{123}^{4,2}, \quad \mathscr{H}_{156}^{4,2}, \quad \mathscr{H}_{246}^{4,2}, \quad \mathscr{H}_{345}^{4,2}$$

are the exceptional divisors, while

$$\mathscr{H}_{456}^{4,2}, \quad \mathscr{H}_{234}^{4,2}, \quad \mathscr{H}_{135}^{4,2}, \quad \mathscr{H}_{126}^{4,2}$$

are the proper transforms of irreducible divisors from $\mathbb{P}^2 \times \mathbb{P}^2$.

Using the above observations we can write down the resolution $\rho_{4,2}$ as a blow up. Set

$$\mathscr{H}_{+}^{4,2} = \mathscr{H}_{123}^{4,2} + \mathscr{H}_{156}^{4,2} + \mathscr{H}_{246}^{4,2} + \mathscr{H}_{345}^{4,2}, \quad \text{and} \quad \mathscr{H}_{-}^{4,2} = \mathscr{H}_{456}^{4,2} + \mathscr{H}_{234}^{4,2} + \mathscr{H}_{135}^{4,2} + \mathscr{H}_{126}^{4,2}.$$

Then the divisor $-\mathscr{H}_{+}^{4,2}$ is β-ample. Since $\operatorname{rk}\operatorname{Pic}(\mathscr{Y}_{4,2})^{\mathfrak{S}_{4,2}} = 2$ by definition (2.3) (indeed, the group $\mathfrak{S}_{4,2}$ acts transitively on the set $\{\boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3, \boldsymbol{P}_4\}$ and swaps the factors of $\mathbb{P}(W_3) \times \mathbb{P}(W_3)$) the divisor $\mathscr{H}_{+}^{4,2}$ is $\rho_{4,2}$-ample, so that the divisor $-\mathscr{H}_{-}^{4,2}$ is also $\rho_{4,2}$-ample. We conclude that the small birational morphism $\rho_{4,2}$ is the blow up of the Weil divisor $\mathscr{H}_{456} + \mathscr{H}_{234} + \mathscr{H}_{135} + \mathscr{H}_{126}$ on $\mathscr{Y}$. Note that the other choice of an $\mathfrak{S}_{4,2}$-equivariant small resolution of singularities of $\mathscr{Y}$, that is, the morphism $\sigma \circ \rho_{4,2}$, is the blow up of the Weil divisor $\mathscr{H}_{123} + \mathscr{H}_{156} + \mathscr{H}_{246} + \mathscr{H}_{345}$ on $\mathscr{Y}$.

2.2. ***$\mathbb{P}^2$-bundle over the quintic del Pezzo surface.*** In this section we construct another resolution of the Coble fourfold, using geometry of the quintic del Pezzo surface. Before explaining the construction, we start with recalling this geometry (we refer the reader to [Dolgachev 2012, Section 8.5; Cheltsov and Shramov 2016a, Section 6.2] for more details).

Let S be the (smooth) del Pezzo surface of degree 5. Recall that S can be represented as the blow up of $\mathbb{P}^2$ in four points (in five different ways), and one has $\operatorname{Aut}(S) \cong \mathfrak{S}_5$. The vector space $H^0(S, \omega_S^{-1})$ is the unique irreducible six-dimensional representation of $\mathfrak{S}_5$ (corresponding to the partition (3, 1, 1) in the notation of [Fulton and Harris 1991, Section 4.1]); see [Shepherd-Barron 1989, Lemma 1]; in particular, this representation is invariant under the sign twist. Moreover, the anticanonical line bundle ω_S^{-1} is very ample and defines an $\mathfrak{S}_5$-equivariant embedding

$$S \hookrightarrow \mathbb{P}^5 = \mathbb{P}(H^0(S, \omega_S^{-1})^\vee)$$

such that S is an intersection of five quadrics in $\mathbb{P}^5$. The five-dimensional space of quadrics passing through S in $\mathbb{P}^5$ is an irreducible representation of $\mathfrak{S}_5$; see [Shepherd-Barron 1989, Proposition 2]. We denote by

$$\mathrm{W}_5 := H^0(\mathbb{P}^5, I_S(2))^\vee \tag{2.28}$$

its dual space. Later, we will identify this space with the space defined by (2.2).

Below we consider the Grassmannian $\operatorname{Gr}(2, \mathrm{W}_5^\vee) \cong \operatorname{Gr}(3, \mathrm{W}_5)$ of two-dimensional vector subspaces in $\mathrm{W}_5^\vee$ (respectively, three-dimensional subspaces in W_5) and denote by $\mathscr{U}_2$ and $\mathscr{U}_3$ the tautological rank 2 and rank 3 subbundles in the trivial vector bundles on this Grassmannian with fibers $\mathrm{W}_5^\vee$ and W_5, respectively.

The following result is well known.

Lemma 2.29. *There is an $\mathfrak{S}_5$-equivariant linear embedding $\mathbb{P}^5 \subset \mathbb{P}(\Lambda^3 \mathrm{W}_5)$ such that*

$$S = \mathrm{Gr}(3, \mathrm{W}_5) \cap \mathbb{P}^5 \subset \mathbb{P}(\Lambda^3 \mathrm{W}_5)$$

is a complete intersection of the Grassmannian $\mathrm{Gr}(3, \mathrm{W}_5)$ *with* $\mathbb{P}^5$.

Proof. We use the technique of excess conormal bundles developed in [Debarre and Kuznetsov 2018, Appendix A]. Since S is an intersection of quadrics, the composition

$$\mathrm{W}_5^\vee \otimes \mathscr{O}_{\mathbb{P}^5} \to I_S(2) \to (I_S/I_S^2)(2)$$

is surjective. The conormal sheaf $\mathscr{N}^\vee_{S/\mathbb{P}^5} \cong I_S/I_S^2$ is locally free of rank 3 on S, hence the above surjection induces an $\mathfrak{S}_5$-equivariant map $S \to \mathrm{Gr}(3, \mathrm{W}_5)$ such that the pullback of the dual tautological bundle $\mathscr{U}_3^\vee$ from $\mathrm{Gr}(3, \mathrm{W}_5)$ to S is isomorphic to $(I_S/I_S^2)(2)$. By adjunction formula we have

$$\det(I_S/I_S^2) \cong \omega_{\mathbb{P}^5|_S} \otimes \omega_S^{-1} \cong \det(\mathrm{W}_5^\vee) \otimes \omega_S^6 \otimes \omega_S^{-1},$$

hence

$$\det((I_S/I_S^2)(2)) \cong \det(\mathrm{W}_5^\vee) \otimes \omega_S^{-1},$$

hence the pullback of $\mathscr{O}_{\mathrm{Gr}(3,\mathrm{W}_5)}(1) \cong \det(\mathscr{U}_3^\vee)$ to S is isomorphic to $\det(\mathrm{W}_5^\vee) \otimes \omega_S^{-1}$. The induced map

$$\Lambda^3 \mathrm{W}_5^\vee \cong H^0(\mathrm{Gr}(3, \mathrm{W}_5), \mathscr{O}_{\mathrm{Gr}(3,\mathrm{W}_5)}(1)) \to H^0(S, \det(\mathrm{W}_5^\vee) \otimes \omega_S^{-1}) \cong \det(\mathrm{W}_5^\vee) \otimes H^0(S, \omega_S^{-1})$$

is $\mathfrak{S}_5$-equivariant and surjective (since the target space is an irreducible $\mathfrak{S}_5$-representation). Moreover, since the $\mathfrak{S}_5$-representation $H^0(S, \omega_S^{-1})$ is invariant under the sign twist, the above composition defines an embedding

$$\mathbb{P}^5 = \mathbb{P}(H^0(S, \omega_S^{-1})^\vee) \hookrightarrow \mathbb{P}(\Lambda^3 \mathrm{W}_5)$$

such that $S \subset \mathrm{Gr}(3, \mathrm{W}_5) \cap \mathbb{P}^5$. It remains to show that this embedding of S is an equality.

Since $\mathrm{Gr}(3, \mathrm{W}_5) \subset \mathbb{P}(\Lambda^3 \mathrm{W}_5)$ is cut out by Plücker quadrics that are parametrized by the space $\mathrm{W}_5^\vee \otimes \det(\mathrm{W}_5^\vee)$, we obtain a map (where the first isomorphism takes place by [Debarre and Kuznetsov 2018, Proposition A.7])

$$\mathrm{W}_5^\vee \otimes \det(\mathrm{W}_5^\vee) \cong H^0(\mathbb{P}(\Lambda^3 \mathrm{W}_5), I_{\mathrm{Gr}(3,\mathrm{W}_5)}(2)) \to H^0(\mathbb{P}^5, I_S(2)) \cong \mathrm{W}_5^\vee \tag{2.30}$$

which by construction commutes with the natural $\mathfrak{S}_5$-action. It is nonzero since $\mathrm{Gr}(3, \mathrm{W}_5)$ does not contain $\mathbb{P}^5$, hence it is an isomorphism by irreducibility of W_5. Since S is an intersection of quadrics, it follows that $S = \mathrm{Gr}(3, \mathrm{W}_5) \cap \mathbb{P}^5$. □

Remark 2.31 (cf. [Shepherd-Barron 1989, Corollary 3]). In (2.30) we obtained an $\mathfrak{S}_5$-equivariant isomorphism $\mathrm{W}_5^\vee \otimes \det(\mathrm{W}_5^\vee) \cong \mathrm{W}_5^\vee$. This allows to identify W_5 as the (unique) irreducible five-dimensional representation of $\mathfrak{S}_5$ with $\det(\mathrm{W}_5)$ being trivial. It corresponds to the Young diagram of the partition $(3, 2)$ in the notation of [Fulton and Harris 1991, Section 4.1].

We denote the restriction of the tautological bundles $\mathscr{U}_2$ and $\mathscr{U}_3$ to S also by $\mathscr{U}_2$ and $\mathscr{U}_3$. The tautological embeddings $\mathscr{U}_2 \hookrightarrow \mathrm{W}_5^\vee \otimes \mathscr{O}_S$ and $\mathscr{U}_3 \hookrightarrow \mathrm{W}_5 \otimes \mathscr{O}_S$ induce $\mathfrak{S}_5$-equivariant maps

$$\mathbb{P}_S(\mathscr{U}_2) \to \mathbb{P}(\mathrm{W}_5^\vee) \quad \text{and} \quad \mathbb{P}_S(\mathscr{U}_3) \to \mathbb{P}(\mathrm{W}_5).$$

Below we describe these maps explicitly. We start with the first of them.

Lemma 2.32. *The image of the map $\varpi : \mathbb{P}_S(\mathscr{U}_2) \to \mathbb{P}(\mathrm{W}_5^\vee)$ is the Segre cubic hypersurface in $\mathbb{P}(\mathrm{W}_5^\vee) \cong \mathbb{P}^4$, and $\mathbb{P}_S(\mathscr{U}_2)$ provides its small $\mathfrak{S}_5$-equivariant resolution of singularities.*

Proof. Let us describe the fiber of ϖ over a point of $\mathbb{P}(\mathrm{W}_5^\vee)$. Thinking of such a point as of a four-dimensional subspace $U_4 \subset \mathrm{W}_5$, we conclude that

$$\varpi^{-1}([U_4]) = \mathrm{Gr}(3, U_4) \cap \mathbb{P}^5 \subset \mathrm{Gr}(3, \mathrm{W}_5) \cap \mathbb{P}^5 = S.$$

Since $\mathrm{Gr}(3, U_4) \cong \mathbb{P}^3$, this intersection is a linear space contained in S, hence either is empty, or is a point, or is a line. Conversely, if $L \subset S$ is a line, then

$$\mathscr{U}_2|_L \cong \mathscr{O}_L \oplus \mathscr{O}_L(-1)$$

because $\mathscr{U}_2^\vee$ is globally generated with $\det(\mathscr{U}_2^\vee) \cong \omega_S^{-1}$. Moreover, the section

$$L = \mathbb{P}_L(\mathscr{O}_L) \hookrightarrow \mathbb{P}_L(\mathscr{U}_2|_L) \hookrightarrow \mathbb{P}_S(\mathscr{U}_2)$$

of the projection $\mathbb{P}_L(\mathscr{U}_2|_L) \to L$ is contracted by the map ϖ. This proves that ϖ contracts precisely the exceptional sections over the ten lines of S, hence the image

$$Z := \varpi(\mathbb{P}_S(\mathscr{U}_2)) \subset \mathbb{P}(\mathrm{W}_5^\vee)$$

is a hypersurface with ten isolated singular points and the map $\mathbb{P}_S(\mathscr{U}_2) \to Z$ is a small resolution of singularities. On the other hand, since $\det(\mathscr{U}_2) \cong \omega_S$, it follows that

$$\omega_{\mathbb{P}_S(\mathscr{U}_2)} \cong \varpi^* \mathscr{O}_{\mathbb{P}(\mathrm{W}_5^\vee)}(-2).$$

Since the map $\mathbb{P}_S(\mathscr{U}_2) \to Z$ is small, we have $\omega_Z \cong \mathscr{O}_{\mathbb{P}(\mathrm{W}_5^\vee)}(-2)|_Z$, so that Z is a cubic hypersurface. It remains to notice that the only three-dimensional cubic with ten isolated singular points is the Segre cubic; see e.g., [Dolgachev 2016, Proposition 2.1]; alternatively, one can deduce this from the fact that the group $\mathfrak{S}_5$ acting in the irreducible five-dimensional representation $\mathrm{W}_5^\vee$ has a unique cubic invariant, which must thus define the Segre cubic. $\square$

Remark 2.33 [Dolgachev 2016, Section 2; Prokhorov 2010, Proposition 4.6]. The relation of the quintic del Pezzo surface S and the Segre cubic threefold Z extends to an $\mathfrak{S}_5$-equivariant diagram

$$\begin{tikzcd}
 & & \overline{\mathcal{M}}_{0,6} \arrow[dl] \arrow[dr] & & \\
 & \mathrm{Bl}_{5\,\mathrm{pt}}(\mathbb{P}^3) \arrow[dl] \arrow[dr] \arrow[rr, dashed, leftrightarrow] & & \mathbb{P}_S(\mathscr{U}_2) \arrow[dl] \arrow[dr] & \\
\mathbb{P}^3 & & Z & & S \cong \overline{\mathcal{M}}_{0,5}
\end{tikzcd}$$

Here $\overline{\mathcal{M}}_{0,n}$ is the moduli spaces of stable rational curves with n marked points, the left outer diagonal arrows provide its Kapranov's representation (the lower left arrow is the blow up of five general points on $\mathbb{P}^3$), the right outer diagonal arrows compose to the forgetful map $\overline{\mathcal{M}}_{0,6} \to \overline{\mathcal{M}}_{0,5}$, the inner diagonal arrows contract ten smooth rational curves each (and provide two $\mathfrak{S}_5$-equivariant small resolutions of Z), and the dashed arrow is a flop in these curves.

The above diagram can be thought of as an $\mathfrak{S}_5$-Sarkisov link from the Mori fiber space $\mathbb{P}_S(\mathscr{U}_2) \to S$ to $\mathbb{P}^3$ centered at Z; see Section 5.1 below for explanation of terminology. It is natural to ask what is the $\mathfrak{S}_5$-Sarkisov link starting from $\mathbb{P}_S(\mathscr{U}_3) \to S$. We will see in diagram (2.48) below that it is a symmetric link centered at the Coble fourfold $\mathscr{Y}$.

So, we consider the projectivization $\mathbb{P}_S(\mathscr{U}_3)$ of the rank 3 bundle $\mathscr{U}_3$ and denote it by

$$\mathscr{Y}_{5,1} := \mathbb{P}_S(\mathscr{U}_3).$$

The embedding $\mathscr{U}_3 \hookrightarrow \mathrm{W}_5 \otimes \mathscr{O}_S$ induces an $\mathfrak{S}_5$-equivariant diagram

$$\begin{tikzcd}
 & \mathscr{Y}_{5,1} \arrow[dl, "p"'] \arrow[dr, "\pi_{5,1}"] & \\
S & & \mathbb{P}(\mathrm{W}_5)
\end{tikzcd} \tag{2.34}$$

where p is the natural projection $\mathscr{Y}_{5,1} = \mathbb{P}_S(\mathscr{U}_3) \to S$, and $\pi_{5,1}$ is the composition of the embedding $\mathscr{Y}_{5,1} \hookrightarrow S \times \mathbb{P}(\mathrm{W}_5)$ with the projection to the second factor. In particular, the restriction of the map p to any fiber of $\pi_{5,1}$ is an isomorphism to its image. This allows to consider every fiber

$$S_w := \pi_{5,1}^{-1}(w)$$

of the map $\pi_{5,1}$ as a closed subscheme of S. In the next lemma we describe these subschemes.

For each point $w \in \mathbb{P}(\mathrm{W}_5)$ denote by W_5/w the four-dimensional quotient of the space W_5 by the line in W_5 that corresponds to w. Every two-dimensional subspace in W_5/w gives (by taking preimage) a three-dimensional subspace in W_5 containing w. This allows to consider $\mathrm{Gr}(2, \mathrm{W}_5/w)$ as a subvariety of $\mathrm{Gr}(3, \mathrm{W}_5)$.

Lemma 2.35. *The fiber S_w of the map $\pi_{5,1}$ over a point $w \in \mathbb{P}(\mathrm{W}_5)$ can be described as*

$$S_w = \mathrm{Gr}(2, \mathrm{W}_5/w) \cap \mathbb{P}^5 \subset \mathrm{Gr}(3, \mathrm{W}_5) \cap \mathbb{P}^5 = S.$$

In particular, S_w is either a zero-dimensional scheme of length 2, or a line, or a conic.

Proof. The first equality is obvious. Consequently, S_w is a linear section of the four-dimensional quadric $\mathrm{Gr}(2, \mathrm{W}_5/w)$ of codimension at most 4. So, if S_w is zero-dimensional, it is a scheme of length 2. Furthermore, if S_w is one-dimensional, it is either a line or a conic. It remains to notice that $\dim S_w < \dim S = 2$ since S is irreducible. □

Our goal is to describe the map $\pi_{5,1}$ in (2.34). We start by presenting some surfaces in $\mathscr{Y}_{5,1}$ contracted by it. Recall that S contains 10 lines. Recall also that $\mathscr{U}_3$ is a subbundle in the trivial vector bundle with fiber W_5 over S, so that $\mathscr{Y}_{5,1}$ is a subvariety in $S \times \mathbb{P}(\mathrm{W}_5)$.

Lemma 2.36. *For every line $L \subset S$ there is a unique line $L' \subset \mathbb{P}(\mathrm{W}_5)$ such that for the surface $R_L = L \times L'$ one has*

$$R_L \subset \mathscr{Y}_{5,1} \subset S \times \mathbb{P}(\mathrm{W}_5). \tag{2.37}$$

In particular, the map $\pi_{5,1}$ contracts R_L onto the line L'. Moreover, if $L_1 \neq L_2$ are distinct lines on S then the corresponding lines $L'_1, L'_2 \subset \mathbb{P}(\mathrm{W}_5)$ are distinct as well.

Proof. Since L is a line on $\mathrm{Gr}(3, \mathrm{W}_5)$, there is a unique two-dimensional subspace $U_2 \subset \mathrm{W}_5$ such that $L \subset \mathbb{P}(\mathrm{W}_5/U_2) \subset \mathrm{Gr}(3, \mathrm{W}_5)$. Then for every point $[U_3]$ of L we have $U_2 \subset U_3$, that is, $U_2 \otimes \mathscr{O}_L \subset \mathscr{U}_3|_L$, hence

$$L \times \mathbb{P}(U_2) = \mathbb{P}_L(U_2 \otimes \mathscr{O}_L) \subset \mathbb{P}_S(\mathscr{U}_3) = \mathscr{Y}_{5,1}.$$

Thus, the line $L' = \mathbb{P}(U_2) \subset \mathbb{P}(\mathrm{W}_5)$ has the required property.

Furthermore, for any two-dimensional subspace $U_2 \subset \mathrm{W}_5$ the intersection

$$\mathbb{P}(\mathrm{W}_5/U_2) \cap S = \mathbb{P}(\mathrm{W}_5/U_2) \cap \mathbb{P}^5$$

is a linear space contained in S, hence is either empty, or a point, or a line. In particular, two distinct lines L_1 and L_2 on S cannot correspond to the same subspace $U_2 \subset \mathrm{W}_5$, hence the corresponding lines L'_1 and L'_2 in $\mathbb{P}(\mathrm{W}_5)$ are distinct. □

As we already mentioned, a quintic del Pezzo surface is classically represented as the blow up of $\mathbb{P}^2$ in four general points. Let $\varphi\colon S \to \mathbb{P}^2$ be one of such blow up representations with exceptional divisors E_0, E_1, E_2, and E_3. Denote by e_i their classes in $\mathrm{Pic}(S)$, and by ℓ the pullback of the line class from $\mathbb{P}^2$ to S, so that

$$K_S \sim -3\ell + e_0 + e_1 + e_2 + e_3.$$

The line bundle $\mathscr{O}_S(\ell)$ defines the contraction $\varphi\colon S \to \mathbb{P}^2$ and the line bundle $\mathscr{O}_S(2\ell - e_0 - e_1 - e_2 - e_3)$ defines a conic bundle $\bar{\varphi}\colon S \to \mathbb{P}^1$. The combination of φ and $\bar{\varphi}$ defines an embedding

$$\varphi \times \bar{\varphi}\colon S \hookrightarrow \mathbb{P}^2 \times \mathbb{P}^1,$$

whose image is a divisor of bidegree (2, 1). Moreover, the composition of $\varphi \times \bar{\varphi}$ with the Segre embedding $\mathbb{P}^2 \times \mathbb{P}^1 \hookrightarrow \mathbb{P}^5$ is the anticanonical embedding of S, therefore we have an exact sequence of normal bundles

$$0 \to \mathcal{N}_{S/\mathbb{P}^2\times\mathbb{P}^1} \to \mathcal{N}_{S/\mathbb{P}^5} \to \mathcal{N}_{\mathbb{P}^2\times\mathbb{P}^1/\mathbb{P}^5}|_S \to 0. \tag{2.38}$$

The first of these bundles is isomorphic to

$$\varphi^*\mathscr{O}_{\mathbb{P}^2}(2) \otimes \bar{\varphi}^*\mathscr{O}_{\mathbb{P}^1}(1) \cong \mathscr{O}_S(4\ell - e_0 - e_1 - e_2 - e_3),$$

and the second is isomorphic to $\mathscr{U}_3(6\ell - 2e_0 - 2e_1 - 2e_2 - 2e_3)$ by the proof of Lemma 2.29. The third vector bundle in (2.38) can be computed as follows: We denote by $\mathscr{T}_{\mathbb{P}^n}$ the tangent bundle of $\mathbb{P}^n$.

Lemma 2.39. *For any positive integers m, n we have $\mathcal{N}_{\mathbb{P}^m\times\mathbb{P}^n/\mathbb{P}^{mn+m+n}} \cong \mathscr{T}_{\mathbb{P}^m} \boxtimes \mathscr{T}_{\mathbb{P}^n}$.*

Proof. Let A and B be vector spaces of dimension $m+1$ and $n+1$ respectively. Tensoring pullbacks to $\mathbb{P}(A) \times \mathbb{P}(B)$ of the Euler sequences

$$0 \to \mathscr{O}_{\mathbb{P}(A)} \to A \otimes \mathscr{O}_{\mathbb{P}(A)}(1) \to \mathscr{T}_{\mathbb{P}(A)} \to 0 \quad \text{and} \quad 0 \to \mathscr{O}_{\mathbb{P}(B)} \to B \otimes \mathscr{O}_{\mathbb{P}(B)}(1) \to \mathscr{T}_{\mathbb{P}(B)} \to 0,$$

we obtain an exact sequence

$$\begin{aligned} 0 \to \mathscr{O}_{\mathbb{P}(A)\times\mathbb{P}(B)} \to A \otimes \mathscr{O}_{\mathbb{P}(A)\times\mathbb{P}(B)}(1,0) \oplus B \otimes \mathscr{O}_{\mathbb{P}(A)\times\mathbb{P}(B)}(0,1) \\ \to A \otimes B \otimes \mathscr{O}_{\mathbb{P}(A)\times\mathbb{P}(B)}(1,1) \to \mathscr{T}_{\mathbb{P}(A)} \boxtimes \mathscr{T}_{\mathbb{P}(B)} \to 0. \end{aligned}$$

Comparing it with the restriction to $\mathbb{P}(A) \times \mathbb{P}(B)$ of the Euler sequence

$$0 \to \mathscr{O}_{\mathbb{P}(A)\times\mathbb{P}(B)} \to A \otimes B \otimes \mathscr{O}_{\mathbb{P}(A)\times\mathbb{P}(B)}(1,1) \to \mathscr{T}_{\mathbb{P}(A\otimes B)}|_{\mathbb{P}(A)\times\mathbb{P}(B)} \to 0$$

of $\mathbb{P}(A \otimes B)$ and with the pullbacks of the Euler sequences of $\mathbb{P}(A)$ and $\mathbb{P}(B)$, we obtain an exact sequence

$$0 \to \mathrm{pr}^*_{\mathbb{P}(A)} \mathscr{T}_{\mathbb{P}(A)} \oplus \mathrm{pr}^*_{\mathbb{P}(B)} \mathscr{T}_{\mathbb{P}(B)} \to \mathscr{T}_{\mathbb{P}(A\otimes B)}|_{\mathbb{P}(A)\times\mathbb{P}(B)} \to \mathscr{T}_{\mathbb{P}(A)} \boxtimes \mathscr{T}_{\mathbb{P}(B)} \to 0,$$

where $\mathrm{pr}_{\mathbb{P}(A)}$ and $\mathrm{pr}_{\mathbb{P}(B)}$ are the projections, which proves the lemma. □

Applying Lemma 2.39 in the case $m = 2$, $n = 1$, we see that the third bundle in (2.38) is isomorphic to

$$\varphi^*(\mathscr{T}_{\mathbb{P}^2}) \otimes \bar{\varphi}^*(\mathscr{T}_{\mathbb{P}^1}) \cong \varphi^*\mathscr{T}_{\mathbb{P}^2} \otimes \mathscr{O}_S(4\ell - 2e_0 - 2e_1 - 2e_2 - 2e_3).$$

So, twisting the normal bundle sequence (2.38) by the line bundle $\mathscr{O}_S(-6\ell + 2e_0 + 2e_1 + 2e_2 + 2e_3)$ we obtain

$$0 \to \mathscr{O}_S(-2\ell + e_0 + e_1 + e_2 + e_3) \to \mathscr{U}_3 \to \varphi^*(\mathscr{T}_{\mathbb{P}^2}(-2)) \to 0 \tag{2.40}$$

Denote by $r_\varphi \colon S \to \mathbb{P}_S(\mathscr{U}_3)$ the section of the projection p induced by the first map in (2.40).

Lemma 2.41. *There is a line $\Gamma_\varphi \subset \mathbb{P}(\mathrm{W}_5)$ and a commutative diagram*

$$\begin{tikzcd} S \arrow[r, "r_\varphi"] \arrow[d, "\bar{\varphi}"'] & \mathbb{P}_S(\mathscr{U}_3) \arrow[d, "\pi_{5,1}"] \\ \Gamma_\varphi \arrow[r, hook] & \mathbb{P}(\mathrm{W}_5) \end{tikzcd}$$

that identifies the line Γ_φ with the base of the conic bundle $\bar{\varphi}$. In particular, for any $w \in \Gamma_\varphi$ the fiber $S_w = \pi_{5,1}^{-1}(w)$ is a conic from the pencil $\bar{\varphi}$.

Proof. By definition of r_φ the composition

$$\pi_{5,1} \circ r_\varphi \colon S \to \mathbb{P}(\mathrm{W}_5)$$

is given by the line bundle $\mathscr{O}_S(2\ell - e_0 - e_1 - e_2 - e_3)$ on S, hence factors as the projection $\bar{\varphi}$ followed by a linear embedding. This proves that we have the required diagram. Moreover, it follows that for every $w \in \Gamma$ the fiber $\pi_{5,1}^{-1}(w)$ contains a conic from the pencil $\bar{\varphi}$. By Lemma 2.35 the fiber coincides with this conic. □

For each contraction $\varphi \colon S \to \mathbb{P}^2$ (recall that for a quintic del Pezzo surface S there are five such contractions), define the surface

$$R_\varphi = r_\varphi(S) \subset \mathbb{P}_S(\mathscr{U}_3), \tag{2.42}$$

so that the map $\pi_{5,1}$ contracts it onto the line $\Gamma_\varphi \subset \mathbb{P}(\mathrm{W}_5)$.

Lemma 2.43. *The five lines $\Gamma_\varphi \subset \mathbb{P}(\mathrm{W}_5)$ corresponding to the contractions $\varphi \colon S \to \mathbb{P}^2$ are pairwise disjoint. Moreover, for each φ the line Γ_φ is distinct from the lines $L' \subset \mathbb{P}(\mathrm{W}_5)$ associated with lines L on S in Lemma 2.36.*

Proof. If w is a common point of the curves Γ_φ and $\Gamma_{\varphi'}$, then by Lemma 2.41 the fiber S_w is a conic that belongs to the corresponding pencils $\bar{\varphi}$ and $\bar{\varphi}'$, hence the pencils coincide, hence $\varphi = \varphi'$.

Assume that $\Gamma_\varphi = L'$, where L' is associated with some line $L \subset S$ as in Lemma 2.36. By Lemma 2.36 we have $L \subset S_w$ for each $w \in L' = \Gamma_\varphi$, and by Lemma 2.41 when w runs over Γ_φ the curves S_w run over the corresponding pencil of conics $\bar{\varphi}$. So, the assumption we made implies that every conic in the pencil contains the line L, which is absurd. □

Now we are ready to prove the main result of this subsection.

Proposition 2.44. *The $\mathfrak{S}_5$-equivariant morphism $\pi_{5,1} \colon \mathscr{Y}_{5,1} \to \mathbb{P}(\mathrm{W}_5)$ gives rise to a commutative diagram*

$$\begin{tikzcd} & \mathscr{Y}_{5,1} \arrow[rr, "\rho_{5,1}"] \arrow[dl, "p"'] \arrow[dr, "\pi_{5,1}"'] & & \mathscr{Y} \arrow[dl, "\pi"] \\ S & & \mathbb{P}(\mathrm{W}_5) & \end{tikzcd} \tag{2.45}$$

where $\mathscr{Y}$ is the Coble fourfold, $\pi : \mathscr{Y} \to \mathbb{P}(\mathrm{W}_5)$ is the double covering, and $\rho_{5,1}$ is a small resolution of singularities, defined uniquely up to composition with the Galois involution $\sigma : \mathscr{Y} \to \mathscr{Y}$. Furthermore, the exceptional locus of $\rho_{5,1}$ is the union of 15 irreducible rational surfaces $\{R_L\}_{L\subset S} \cup \{R_\varphi\}_{\varphi\colon S\to\mathbb{P}^2}$, such that:

- *$R_L \cong \mathbb{P}^1 \times \mathbb{P}^1$; each of these surfaces is contracted by p onto the line $L \subset S$ and by $\pi_{5,1}$ onto the line $L' \subset \mathbb{P}(\mathrm{W}_5)$.*
- *$R_\varphi \cong S$ with the map $p\colon R_\varphi \to S$ being an isomorphism and with the map $\pi_{5,1}|_{R_\varphi}$ being the conic bundle $\bar{\varphi}\colon R_\varphi \twoheadrightarrow \Gamma_\varphi$ over the line $\Gamma_\varphi \subset \mathbb{P}(\mathrm{W}_5)$.*

Moreover, the morphism $\pi_{5,1}$ induces a nonstandard embedding $\mathfrak{S}_5 \to \mathfrak{S}_6$ such that $\rho_{5,1}$ is $\mathfrak{S}_5$-equivariant with respect either to the natural or to the twisted action of $\mathfrak{S}_5$ on $\mathscr{Y}$.

Using a compatibility result from Proposition 2.50, we will show in Section 2.4 that $\rho_{5,1}$ is $\mathfrak{S}_5$-equivariant with respect to the twisted action of a nonstandard $\mathfrak{S}_5$.

Proof. By Lemma 2.35 the map $\pi_{5,1}$ is generically finite of degree 2. Denote by $R \subset \mathscr{Y}_{5,1}$ the ramification locus of the morphism $\pi_{5,1}\colon \mathscr{Y}_{5,1} \to \mathbb{P}(\mathrm{W}_5)$ and by $B = \pi_{5,1}(R) \subset \mathbb{P}(\mathrm{W}_5)$ its image. Let us show that B is the Igusa quartic. For this we show that B is projectively dual to the Segre cubic $Z = \varpi(\mathbb{P}_S(\mathscr{U}_2))$; see Lemma 2.32.

Indeed, by Lemma 2.35 we know that B is the locus of $w \in \mathbb{P}(\mathrm{W}_5)$ such that S_w is either a double point or a curve. On the other hand, w defines a hyperplane $\mathbb{P}(w^\perp) \subset \mathbb{P}(\mathrm{W}_5^\vee)$ in the dual projective space, and

$$\varpi^{-1}(Z \cap \mathbb{P}(w^\perp)) = \mathbb{P}_S(\mathscr{U}_2) \times_{\mathbb{P}(\mathrm{W}_5^\vee)} \mathbb{P}(w^\perp)$$

is a relative hyperplane in the $\mathbb{P}^1$-bundle $\mathbb{P}_S(\mathscr{U}_2) \to S$. Moreover, the zero locus of the corresponding section of $\mathscr{U}_2^\vee$ is precisely the scheme S_w. If S_w is zero-dimensional then by [Kuznetsov 2016, Lemma 2.1] we have

$$\varpi^{-1}(Z \cap \mathbb{P}(w^\perp)) = \mathrm{Bl}_{S_w}(S),$$

and if it is one-dimensional, then $\varpi^{-1}(Z \cap \mathbb{P}(w^\perp))$ contains the surface $\mathbb{P}_{S_w}(\mathscr{U}_2|_{S_w})$, hence is reducible. Thus, $\varpi^{-1}(Z \cap \mathbb{P}(w^\perp))$ is singular if and only if $w \in B$. Since the singular points of Z are nodes, and ϖ resolves them, it follows that B is the projective dual of Z. Hence $B = X_{1/4}$ is the Igusa quartic (see [Hunt 1996, Proposition 3.3.1]).

It follows from Lemma 2.35 that the map $\pi_{5,1}$ is an étale double cover over $\mathbb{P}(\mathrm{W}_5) \setminus B$, and that the Stein factorization of the map $\pi_{5,1}$ provides a (unique up to σ) decomposition

$$\mathscr{Y}_{5,1} \xrightarrow{\rho_{5,1}} \mathscr{Y} \xrightarrow{\pi} \mathbb{P}(\mathrm{W}_5),$$

where $\rho_{5,1}$ is a birational map.

Let us show that $\rho_{5,1}$ is small. Indeed, since $\det(\mathscr{U}_3) \cong \omega_S$, it follows that

$$\omega_{\mathscr{Y}_{5,1}} \cong \pi_{5,1}^* \mathscr{O}_{\mathbb{P}(\mathrm{W}_5)}(-3) \cong \rho_{5,1}^* \pi^* \mathscr{O}_{\mathbb{P}(\mathrm{W}_5)}(-3). \tag{2.46}$$

On the other hand, π is a double covering branched over a quartic, hence one has $\omega_{\mathscr{Y}} \cong \pi^* \mathscr{O}_{\mathbb{P}(\mathrm{W}_5)}(-3)$. Thus $\omega_{\mathscr{Y}_{5,1}} \cong \rho_{5,1}^* \omega_{\mathscr{Y}}$, i.e., the map $\rho_{5,1}$ is crepant. Since $\mathscr{Y}_{5,1}$ is smooth it follows that the map $\rho_{5,1}$ is an isomorphism over the smooth locus of $\mathscr{Y}$, hence the exceptional locus of $\rho_{5,1}$ is contained in

$$\rho_{5,1}^{-1}(\mathrm{Sing}(\mathscr{Y})) = \pi_{5,1}^{-1}(\mathrm{Sing}(X_{\frac{1}{4}})) = \pi_{5,1}^{-1}(\mathrm{CR}),$$

i.e., in the preimage of the Cremona–Richmond configuration of 15 lines. But by Lemma 2.35 the fibers of $\pi_{5,1}$ are at most one-dimensional, hence $\dim(\pi_{5,1}^{-1}(\mathrm{CR})) \leqslant 2$. This proves that $\rho_{5,1}$ is small.

Next, let us show that

$$\pi_{5,1}^{-1}(\mathrm{CR}) = \Big(\bigcup_{\varphi} R_\varphi\Big) \cup \Big(\bigcup_{L} R_L\Big). \tag{2.47}$$

By Lemmas 2.36 and 2.41 the surfaces R_L and R_φ are contracted onto the union of ten lines L' and five lines Γ_φ in $\mathbb{P}(\mathrm{W}_5)$, which are pairwise distinct by Lemmas 2.36 and 2.43. Therefore

$$\mathrm{CR} = \Big(\bigcup_{\varphi} \Gamma_\varphi\Big) \cup \Big(2\bigcup_{L} L'\Big).$$

It remains to show that for any $w \in \Gamma_\varphi$ or $w \in L'$ the fiber $S_w = \pi_{5,1}^{-1}(w)$ is contained either in R_φ or in R_L. If $w \in \Gamma_\varphi$, this is proved in Lemma 2.41. Now take $w \in L'$. By Lemma 2.36 we have $L \subset S_w$, hence by Lemma 2.35 the curve S_w is either the line L (hence $S_w \subset R_L$) or a conic (hence $S_w \subset R_\varphi$ for appropriate φ). This proves (2.47).

The vector space W_5 by definition (2.28) comes with a natural $\mathfrak{S}_5$-action, and, moreover, the map $\pi_{5,1} \colon \mathscr{Y}_{5,1} \to \mathbb{P}(\mathrm{W}_5)$ is $\mathfrak{S}_5$-equivariant. It follows that its branch divisor $B = X_{1/4}$ is invariant under this action. This gives an embedding $\mathfrak{S}_5 \hookrightarrow \mathrm{Aut}(X_{1/4}) \cong \mathfrak{S}_6 \subset \mathrm{Aut}(\mathscr{Y})$, such that for every element $g \in \mathfrak{S}_5$ the conjugation of the diagram (2.45) by g gives a diagram of the same form. Therefore, one has

$$g \circ \rho_{5,1} \circ g^{-1} = \sigma^{k(g)} \circ \rho_{5,1},$$

where $k \colon \mathfrak{S}_5 \to \mathbb{Z}/2\mathbb{Z}$ is a group homomorphism. If it is trivial, then $\rho_{5,1}$ is equivariant with respect to the natural action, and if k is the homomorphism of parity, then $\rho_{5,1}$ is equivariant with respect to the twisted action (as we mentioned above, we will show in Section 2.4 that k is indeed the homomorphism of parity).

To show that the embedding $\mathfrak{S}_5 \hookrightarrow \mathfrak{S}_6$ is nonstandard we use the same argument as in the proof of Proposition 2.21. The restriction of the five-dimensional representation (1.1) to the image of a standard embedding $\mathfrak{S}_5 \hookrightarrow \mathfrak{S}_6$ decomposes as a direct sum of two irreducible representations (see Lemma 5.12), while the $\mathfrak{S}_5$-representation W_5 is irreducible by (2.28) and [Shepherd-Barron 1989, Proposition 2]. □

Similarly to the case of $\rho_{4,2}$, the morphism $\rho_{5,1}$ is not uniquely defined even when the corresponding nonstandard subgroup $\mathfrak{S}_5$ is fixed. Moreover, there is a commutative diagram

$$\begin{tikzcd} & \mathscr{Y}_{5,1} \arrow[ld, "p"'] \arrow[rd, "\sigma\circ\rho_{5,1}"] \arrow[rr, dashed, "\rho_{5,1}^{-1}\circ\sigma\circ\rho_{5,1}"] & & \mathscr{Y}_{5,1} \arrow[ld, "\rho_{5,1}"'] \arrow[rd, "p"] & \\ S & & \mathscr{Y} & & S \end{tikzcd} \tag{2.48}$$

Here $\rho_{5,1}^{-1}\circ\sigma\circ\rho_{5,1}$ is a small birational map. In fact, we know that $\operatorname{rk}\operatorname{Pic}(S)^{\mathfrak{S}_5}=1$; see for instance [Cheltsov and Shramov 2016a, Lemma 6.2.2(i)]; this means that

$$\operatorname{rk}\operatorname{Cl}(\mathscr{Y}_{5,1})^{\mathfrak{S}_5}=\operatorname{rk}\operatorname{Pic}(\mathscr{Y}_{5,1})^{\mathfrak{S}_5}=2,$$

and therefore $\operatorname{rk}\operatorname{Pic}(\mathscr{Y})^{\mathfrak{S}_5}=1$. The latter implies that $\rho_{5,1}$ and $\sigma\circ\rho_{5,1}$ are the only $\mathfrak{S}_5$-equivariant small resolutions of singularities of $\mathscr{Y}$ and that $\rho_{5,1}^{-1}\circ\sigma\circ\rho_{5,1}$ is an $\mathfrak{S}_5$-flop. Consequently, the diagram (2.48) is an $\mathfrak{S}_5$-Sarkisov link between two copies of the Mori fiber space $\mathscr{Y}_{5,1}\to S$ centered at $\mathscr{Y}$ (see Section 5.1).

Remark 2.49. Recall the notation of Remark 2.25. Denote

$$\mathscr{H}_I^{5,1}:=\rho_{5,1}^{-1}(\mathscr{H}_I),$$

so that one has $\pi_{5,1}^{-1}(H_I)=\pi_{5,1}^{-1}(H_{\bar I})=\mathscr{H}_I^{5,1}\cup\mathscr{H}_{\bar I}^{5,1}$. One can check that ten out of twenty divisors $\mathscr{H}_I^{5,1}\subset\mathscr{Y}_{5,1}$ are the preimages of lines on S via the map p, and the other ten are relative hyperplane sections for p (this decomposition is the orbit decomposition for the action of $\mathfrak{S}_5$). We denote by $\mathscr{H}_+^{5,1}$ the sum of the divisors of the first type, and by $\mathscr{H}_-^{5,1}$ the sum of the divisors of the second type. The divisor $\mathscr{H}_+^{5,1}$ is the p-pullback of an ample divisor on S, hence it is $\rho_{5,1}$-ample. Consequently, $-\mathscr{H}_-^{5,1}$ is $\rho_{5,1}$-ample, hence the small birational morphism $\rho_{5,1}$ is the blow up of the Weil divisor $\rho_{5,1}(\mathscr{H}_-^{5,1})$ on $\mathscr{Y}$. See Remark 2.57 below for an explicit description of this blow up.

2.3. *Compatibility of resolutions.* In this section we relate the resolutions $\mathscr{Y}_{4,2}$ and $\mathscr{Y}_{5,1}$ of the Coble fourfold. Recall that the first of them is associated with a nonstandard embedding $\mathfrak{S}_{4,2}\hookrightarrow\mathfrak{S}_6$, and the second is associated with a nonstandard embedding $\mathfrak{S}_5\hookrightarrow\mathfrak{S}_6$. Note that each (standard or nonstandard) subgroup $\mathfrak{S}_4\subset\mathfrak{S}_6$ can be extended to a subgroup $\mathfrak{S}_{4,2}\subset\mathfrak{S}_6$ and such extension is unique. Indeed, the second factor $\mathfrak{S}_2$ in $\mathfrak{S}_{4,2}\cong\mathfrak{S}_4\times\mathfrak{S}_2$ is just the centralizer of $\mathfrak{S}_4$ in $\mathfrak{S}_6$. Recall also that for each $\mathfrak{S}_4\subset\mathfrak{S}_5=\operatorname{Aut}(S)$ there is a unique $\mathfrak{S}_4$-equivariant contraction $\varphi\colon S\to\mathbb{P}^2$ of the quintic del Pezzo surface S onto the plane.

Proposition 2.50. *Let $\mathfrak{S}_5\hookrightarrow\mathfrak{S}_6$ be a nonstandard embedding. Choose a subgroup $\mathfrak{S}_4\subset\mathfrak{S}_5$ and let $\mathfrak{S}_{4,2}\subset\mathfrak{S}_6$ be its unique extension. Let $\rho_{4,2}\colon\mathscr{Y}_{4,2}\to\mathscr{Y}$ be the $\mathfrak{S}_{4,2}$-equivariant resolution of singularities constructed in Proposition 2.21 and let $\varphi\colon S\to\mathbb{P}^2$ be the unique $\mathfrak{S}_4$-equivariant contraction of the quintic del Pezzo surface. Then there is a unique $\mathfrak{S}_5$-equivariant resolution $\rho_{5,1}\colon\mathscr{Y}_{5,1}\to\mathscr{Y}$ as in Proposition 2.44 and a unique $\mathfrak{S}_4$-equivariant small birational map $\theta_1\colon\mathscr{Y}_{5,1}\dashrightarrow\mathscr{Y}_{4,2}$ such that the diagram* (1.10) *is commutative.*

Of course, if $\rho_{5,1}$ is fixed, there is only one θ_1 such that the inner triangle in the diagram (1.10) commutes, namely, $\theta_1 = \rho_{4,2}^{-1} \circ \rho_{5,1}$. But it is a priori not clear why the outer square commutes. So, to prove Proposition 2.50 we move in the opposite direction: we first construct θ_1 such that the outer square commutes, and after that check that the inner triangle commutes for this θ_1 and for an appropriate choice of $\rho_{5,1}$.

We start with some notation and a lemma. Let $\varphi\colon S \to \mathbb{P}^2$ be the $\mathfrak{S}_4$-equivariant contraction, and, as before, denote by E_0, E_1, E_2, and E_3 the exceptional divisors of the blow up φ, by e_i their classes in $\operatorname{Pic}(S)$ and by ℓ the pullback of the line class of $\mathbb{P}^2$. Recall also the rank 3 bundle $\mathscr{U}_3$ on S.

Since $\mathscr{U}_3^\vee$ is globally generated and $\det(\mathscr{U}_3^\vee)|_{E_i} \cong \omega_S^{-1}|_{E_i} \cong \mathscr{O}_{E_i}(1)$, we have

$$\mathscr{U}_3|_{E_i} \cong \mathscr{O}_{E_i} \oplus \mathscr{O}_{E_i} \oplus \mathscr{O}_{E_i}(-1).$$

Therefore, we have a canonical surjective morphism $\mathscr{U}_3 \to \mathscr{O}_{E_i}(-1)$ of sheaves on S. The sum of these morphisms gives an exact $\mathfrak{S}_4$-equivariant sequence

$$0 \to \mathscr{E} \to \mathscr{U}_3 \to \bigoplus_{i=0}^{3} \mathscr{O}_{E_i}(-1) \to 0 \tag{2.51}$$

and defines a rank 3 vector bundle $\mathscr{E}$ on S equivariant with respect to $\mathfrak{S}_4$.

Lemma 2.52. *One has* $\mathscr{E} \cong \mathscr{O}_S(-\ell)^{\oplus 3}$.

Proof. Consider the composition of the embedding

$$\mathscr{O}_S(-2\ell + e_0 + e_1 + e_2 + e_3) \hookrightarrow \mathscr{U}_3$$

from (2.40) with the projection $\mathscr{U}_3 \to \mathscr{O}_{E_i}(-1)$. If it is equal to zero, then the map $\mathscr{U}_3 \to \mathscr{O}_{E_i}(-1)$ factors through a map $\varphi^*(\mathscr{T}_{\mathbb{P}^2}(-2)) \to \mathscr{O}_{E_i}(-1)$. But the sheaf $\varphi^*(\mathscr{T}_{\mathbb{P}^2}(-2))$ restricts to E_i trivially, hence no such map exists. This contradiction shows that the composition is nontrivial. But since

$$\mathscr{O}_S(-2\ell + e_0 + e_1 + e_2 + e_3)|_{E_i} \cong \mathscr{O}_{E_i}(-1),$$

any nontrivial morphism $\mathscr{O}_S(-2\ell + e_0 + e_1 + e_2 + e_3) \to \mathscr{O}_{E_i}(-1)$ is surjective. Therefore, the sum of these morphisms $\mathscr{O}_S(-2\ell + e_0 + e_1 + e_2 + e_3) \to \bigoplus_{i=0}^3 \mathscr{O}_{E_i}(-1)$ is surjective, hence its kernel is $\mathscr{O}_S(-2\ell)$ and we have a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathscr{O}_S(-2\ell) \arrow[r] \arrow[d] & \mathscr{O}_S(-2\ell + e_0 + e_1 + e_2 + e_3) \arrow[r] \arrow[d] & \bigoplus_{i=0}^3 \mathscr{O}_{E_i}(-1) \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathscr{E} \arrow[r] & \mathscr{U}_3 \arrow[r] & \bigoplus_{i=0}^3 \mathscr{O}_{E_i}(-1) \arrow[r] & 0
\end{tikzcd}$$

Taking into account (2.40), we see that the first column extends to an exact sequence

$$0 \to \mathscr{O}_S(-2\ell) \to \mathscr{E} \to \varphi^*(\mathscr{T}_{\mathbb{P}^2}(-2)) \to 0. \tag{2.53}$$

It remains to show that it coincides with the pullback of a twist of the Euler sequence on $\mathbb{P}^2$. Since the pullback functor φ^* is fully faithful, and the Euler sequence is the unique nonsplit extension of $\mathscr{T}_{\mathbb{P}^2}$ by $\mathscr{O}_{\mathbb{P}^2}$, it is enough to show that (2.53) is nonsplit.

Assume on the contrary that there is a splitting $\varphi^*(\mathscr{T}_{\mathbb{P}^2}(-2)) \to \mathscr{E}$. Composing it with the embedding $\mathscr{E} \hookrightarrow \mathscr{U}_3$, we obtain a splitting $\varphi^*(\mathscr{T}_{\mathbb{P}^2}(-2)) \to \mathscr{U}_3$ of (2.40). It induces an embedding

$$S \times_{\mathbb{P}^2} \mathrm{Fl}(1,2;3) \cong \mathbb{P}_S(\varphi^*(\mathscr{T}_{\mathbb{P}^2}(-2))) \hookrightarrow \mathbb{P}_S(\mathscr{U}_3) = \mathscr{Y}_{5,1},$$

such that its composition with $\pi_{5,1}$ coincides with the projection

$$S \times_{\mathbb{P}^2} \mathrm{Fl}(1,2;3) \to \mathrm{Fl}(1,2;3) \to (\mathbb{P}^2)^\vee.$$

But this contradicts the fact that $\rho_{5,1}$ is a small contraction. □

Proof of Proposition 2.50. Let us construct the map θ_1. Let V_1 be a three-dimensional vector space such that the target plane of φ is $\mathbb{P}(V_1)$. We can choose an isomorphism

$$\alpha_1 \colon \mathbb{P}(V_1) \xrightarrow{\sim} \mathbb{P}(\mathrm{W}_3)$$

such that the points of $\mathbb{P}(V_1)$ to which the divisors E_i are contracted by φ go to the points P_i of $\mathbb{P}(\mathrm{W}_3)$ defined by (2.1). Note that such an isomorphism is unique and $\mathfrak{S}_4$-equivariant.

Next, let V_2 be the three-dimensional vector space such that $\mathscr{E} \cong V_2 \otimes \mathscr{O}_S(-\ell)$. Note that the space $V_2 \cong H^0(S, \mathscr{E}(\ell))$ has a natural structure of an $\mathfrak{S}_4$-representation, and the above isomorphism $\mathscr{E} \cong V_2 \otimes \mathscr{O}_S(-\ell)$ is $\mathfrak{S}_4$-equivariant. Under this identification the first map in (2.51) becomes an $\mathfrak{S}_4$-equivariant embedding of sheaves

$$V_2 \otimes \mathscr{O}_S(-\ell) \xrightarrow{\xi} \mathscr{U}_3, \tag{2.54}$$

which is an isomorphism away from the union of E_i. Its dual map extends to an exact $\mathfrak{S}_4$-equivariant sequence

$$0 \to \mathscr{U}_3^\vee \xrightarrow{\xi^\vee} V_2^\vee \otimes \mathscr{O}_S(\ell) \to \bigoplus_{i=0}^{3} \mathscr{O}_{E_i} \to 0. \tag{2.55}$$

The second map defines four linear functions on $V_2^\vee$, i.e., four points on $\mathbb{P}(V_2)$. We can choose an isomorphism

$$\alpha_2 \colon \mathbb{P}(V_2) \xrightarrow{\sim} \mathbb{P}(\mathrm{W}_3)$$

such that these points go to the points P_i of $\mathbb{P}(\mathrm{W}_3)$ defined by (2.1). Again, such an isomorphism is unique and $\mathfrak{S}_4$-equivariant.

Now we put all the above constructions together. The morphism ξ defined by (2.54) induces a birational map

$$S \times \mathbb{P}(V_2) \cong \mathbb{P}_S(V_2 \otimes \mathscr{O}_S(-\ell)) \overset{\xi}{\dashrightarrow} \mathbb{P}_S(\mathscr{U}_3) = \mathscr{Y}_{5,1}.$$

We define θ_1 as the composition

$$\mathscr{Y}_{5,1} \xrightarrow{\xi^{-1}} S \times \mathbb{P}(V_2) \xrightarrow{\varphi \times \mathrm{id}} \mathbb{P}(V_1) \times \mathbb{P}(V_2) \xrightarrow{\alpha_1 \times \alpha_2} \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \xrightarrow{\beta^{-1}} \mathscr{Y}_{4,2},$$

where the last map is the inverse of the blow up (2.3). Clearly, θ_1 is birational and $\mathfrak{S}_4$-equivariant, since all the maps used in its definition are. Finally, its composition with p_1 equals $\varphi \circ p$ by construction, hence the outer square in (1.10) commutes.

Next, let us show an equality of the maps

$$\pi_{4,2} \circ \theta_1 = \pi_{5,1} \tag{2.56}$$

from $\mathscr{Y}_{5,1}$ to $\mathbb{P}(\mathrm{W}_5)$. For this, consider the diagram

$$\begin{tikzcd}
\mathrm{W}_5^\vee \otimes \mathscr{O}_S \arrow[r] \arrow[d, equal] & \mathrm{W}_3^\vee \otimes \mathrm{W}_3^\vee \otimes \mathscr{O}_S \arrow[r, "(\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3)"] \arrow[d, "\cong"', "(\alpha_2^\vee \otimes \alpha_1^\vee)"] & \bigoplus_{i=0}^3 \mathscr{O}_S \arrow[d, equal] \\
\mathrm{W}_5^\vee \otimes \mathscr{O}_S \arrow[r, "H^0(S,\xi^\vee)"] \arrow[d] & V_2^\vee \otimes V_1^\vee \otimes \mathscr{O}_S \arrow[r] \arrow[d] & \bigoplus_{i=0}^3 \mathscr{O}_S \arrow[d] \\
\mathscr{U}_3^\vee \arrow[r, "\xi^\vee"] & V_2^\vee \otimes \mathscr{O}_S(\ell) \arrow[r] & \bigoplus_{i=0}^3 \mathscr{O}_{E_i}
\end{tikzcd}$$

Here the bottom line is (2.55), the middle line is obtained from it by passing to global sections and tensoring with $\mathscr{O}_S$, and the maps between these lines are induced by evaluation of sections (hence the lower squares commute). The top line is obtained by identification (2.2), the upper-right square commutes by definition of α_1 and α_2. Therefore, there is a unique identification of the spaces $\mathrm{W}_5^\vee$ in this diagram (note that the one in the top line is defined by (2.2), while the other is defined by (2.28)) such that the upper-left square commutes. From now on we use implicitly the induced identification of the spaces W_5.

As a result of this commutativity two morphisms $\mathrm{W}_5^\vee \otimes \mathscr{O}_S \to V_2^\vee \otimes \mathscr{O}_S(\ell)$ in the diagram coincide. One of them induces the rational map

$$S \times \mathbb{P}(V_2) \xrightarrow{\xi} \mathscr{Y}_{5,1} \xrightarrow{\pi_{5,1}} \mathbb{P}(\mathrm{W}_5),$$

and the other induces the rational map

$$S \times \mathbb{P}(V_2) \xrightarrow{\varphi \times \mathrm{id}} \mathbb{P}(V_1) \times \mathbb{P}(V_2) \xrightarrow{\alpha_1 \times \alpha_2} \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3) \xrightarrow{\overline{\pi}_{4,2}} \mathbb{P}(\mathrm{W}_5);$$

the map φ appears here because all the global sections of $\mathscr{O}_S(\ell)$ are pullbacks via φ. So, we have an equality of rational maps

$$\overline{\pi}_{4,2} \circ (\alpha_1 \times \alpha_2) \circ (\varphi \times \mathrm{id}) = \pi_{5,1} \circ \xi$$

from $S \times \mathbb{P}(V_2)$ to $\mathbb{P}(\mathrm{W}_5)$. Composing it with the map ξ^{-1} on the right and using (2.4) and the definition of θ_1, we deduce the required equality (2.56).

From (2.56) we further deduce an equality

$$\pi \circ (\rho_{4,2} \circ \theta_1) = \pi_{4,2} \circ \theta_1 = \pi_{5,1}.$$

Hence, the composition $\rho_{4,2} \circ \theta_1$ provides one of the two possible factorizations $\rho_{5,1}$ of the morphism $\pi_{5,1}$. This shows that for one of the two choices of $\rho_{5,1}$, the inner triangle in (1.10) is commutative. $\square$

It is worth noting that if we want to replace the map p_1 in the diagram (1.10) by another projection p_2 and preserve its commutativity, we will have to replace the subgroup $\mathfrak{S}_5$ containing $\mathfrak{S}_4$ by the unique other such subgroup (more precisely, we will have to replace the embedding $\mathfrak{S}_5 \hookrightarrow \mathfrak{S}_6$ with the one obtained from it by a conjugation with the factor $\mathfrak{S}_2$ in $\mathfrak{S}_{4,2}$).

Remark 2.57. Recall the notation of Remarks 2.25 and 2.49, and assume that we are in the situation of Proposition 2.50: the resolution $\rho_{4,2}$ is defined by (2.23) and the resolution $\rho_{5,1}$ is such that the diagram (1.10) commutes. Then we have

$$\mathscr{H}_+^{5,1} = \mathscr{H}_{123}^{5,1} + \mathscr{H}_{156}^{5,1} + \mathscr{H}_{246}^{5,1} + \mathscr{H}_{345}^{5,1} + \mathscr{H}_{124}^{5,1} + \mathscr{H}_{136}^{5,1} + \mathscr{H}_{235}^{5,1} + \mathscr{H}_{145}^{5,1} + \mathscr{H}_{256}^{5,1} + \mathscr{H}_{346}^{5,1},$$
$$\mathscr{H}_-^{5,1} = \mathscr{H}_{456}^{5,1} + \mathscr{H}_{234}^{5,1} + \mathscr{H}_{135}^{5,1} + \mathscr{H}_{126}^{5,1} + \mathscr{H}_{356}^{5,1} + \mathscr{H}_{245}^{5,1} + \mathscr{H}_{146}^{5,1} + \mathscr{H}_{236}^{5,1} + \mathscr{H}_{134}^{5,1} + \mathscr{H}_{125}^{5,1}.$$

Consequently, $\rho_{5,1}$ is the blow up of the Weil divisor

$$\mathscr{H}_{456} + \mathscr{H}_{234} + \mathscr{H}_{135} + \mathscr{H}_{126} + \mathscr{H}_{356} + \mathscr{H}_{245} + \mathscr{H}_{146} + \mathscr{H}_{236} + \mathscr{H}_{134} + \mathscr{H}_{125}$$

on $\mathscr{Y}$.

2.4. *Proof of Theorem 1.9.* In Proposition 2.21 we constructed the morphism $\rho_{4,2}$ for some nonstandard subgroup $\mathfrak{S}_{4,2} \subset \mathfrak{S}_6$, and checked that it is $\mathfrak{S}_{4,2}$-equivariant for the twisted action and small. To construct $\rho_{4,2}$ for any other nonstandard embedding, we may use a conjugation by an appropriate element of $\mathfrak{S}_6$. This proves assertion (i).

Similarly to the above, in Proposition 2.44 we constructed a morphism $\rho_{5,1}$ for some nonstandard embedding $\mathfrak{S}_5 \hookrightarrow \mathfrak{S}_6$ (and the same trick as above then gives $\rho_{5,1}$ for any other nonstandard $\mathfrak{S}_5 \subset \mathfrak{S}_6$) and checked that it is small. Moreover, the compatibility isomorphism θ_1 was constructed in Proposition 2.50; by the way it proves assertion (iv).

Furthermore, we checked that the morphism $\rho_{5,1}$ is $\mathfrak{S}_5$-equivariant with respect either to the natural or to the twisted action of $\mathfrak{S}_5$ on $\mathscr{Y}$. To show that the action is twisted, we use Proposition 2.50. Choose a subgroup $\mathfrak{S}_4 \subset \mathfrak{S}_5$, a transposition $g \in \mathfrak{S}_4$, and consider the commutative diagram (1.10). Since θ_1 is $\mathfrak{S}_4$-equivariant and $g \circ \rho_{4,2} \circ g^{-1} = \sigma \circ \rho_{4,2}$ (as $\rho_{4,2}$ is equivariant with respect to the twisted action), we have

$$g \circ \rho_{5,1} \circ g^{-1} = g \circ \rho_{4,2} \circ \theta_1 \circ g^{-1} = g \circ \rho_{4,2} \circ g^{-1} \circ \theta_1 = \sigma \circ \rho_{4,2} \circ \theta_1 = \sigma \circ \rho_{5,1},$$

hence $\rho_{5,1}$ is equivariant with respect to the twisted action as well. This completes the proof of assertion (ii).

Finally, recall that we checked in Propositions 2.21 and 2.44 that $\rho_{4,2}$ and $\rho_{5,1}$ are isomorphisms over the complement of the Cremona–Richmond configuration $\mathrm{CR} = \mathrm{Sing}(X_{1/4}) \subset \mathbb{P}^4$. This gives the proof of assertion (iii) and completes the proof of Theorem 1.9. $\square$

3. Conic bundle structures on $\mathfrak{S}_6$-invariant quartics

Recall the pencil $\{X_t\}$ of $\mathfrak{S}_6$-invariant quartics defined by the (1.2) inside the hyperplane $\mathbb{P}^4 \subset \mathbb{P}^5$ given by (1.1). In this section we discuss the conic bundle structures on the quartics X_t induced by the resolutions of the Coble fourfold.

3.1. ***$\mathfrak{S}_6$-invariant quartics revisited.*** We start by collecting some facts about automorphism groups of X_t, their singularities and class groups.

Let CR be the Cremona–Richmond configuration of 15 lines with 15 intersection points; see the Appendix. The intersection points of the lines of CR form the orbit

$$\Upsilon_{15} = \{g \cdot (2:2:-1:-1:-1:-1) \mid g \in \mathfrak{S}_6\}.$$

Besides this, we consider also the orbits

$$\begin{aligned}
\Sigma_6 &= \{g \cdot (5:-1:-1:-1:-1:-1) \mid g \in \mathfrak{S}_6\},\\
\Sigma_{10} &= \{g \cdot (1:1:1:-1:-1:-1) \quad\mid g \in \mathfrak{S}_6\},\\
\Sigma_{15} &= \{g \cdot (1:-1:0:0:0:0) \qquad\mid g \in \mathfrak{S}_6\},\\
\Sigma_{30} &= \{g \cdot (1:1:\omega:\omega:\omega^2:\omega^2) \quad\mid g \in \mathfrak{S}_6\},
\end{aligned}$$

where ω is a primitive cubic root of unity and the lower index on the left-hand side stands for cardinality of the orbit. We note that

$$\Upsilon_{15} \subset \mathrm{CR}, \quad \Sigma_{30} \subset \mathrm{CR}, \quad (\Sigma_6 \cup \Sigma_{10} \cup \Sigma_{15}) \cap \mathrm{CR} = \varnothing.$$

Remark 3.1. The quartic X_∞ defined by (1.2) with $t = \infty$ is the quadric Q_∞ given by the equation

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 = 0 \tag{3.2}$$

taken with multiplicity 2. Note that

$$Q_\infty \cap \Upsilon_{15} = \varnothing, \quad Q_\infty \cap \mathrm{CR} = \Sigma_{30},$$

and the intersection is transversal.

The singularities of the quartics X_t have been described by van der Geer [1982] in terms of these orbits. Recall the discriminant set $\mathfrak{D}$ defined by (1.3).

Theorem 3.3 [van der Geer 1982, Theorem 4.1]. *One has*

t	$t \notin \mathfrak{D} \cup \{\infty\}$	$t = \frac{1}{4}$	$t = \frac{1}{2}$	$t = \frac{1}{6}$	$t = \frac{7}{10}$
$\mathrm{Sing}(X_t)$	Σ_{30}	CR	$\Sigma_{30} \cup \Sigma_{15}$	$\Sigma_{30} \cup \Sigma_{10}$	$\Sigma_{30} \cup \Sigma_6$

In particular, X_t is normal if $t \neq \infty$.

Moreover, all singular points of the quartics X_t are nodes provided that $t \neq \frac{1}{4}, \infty$.

One can describe automorphism groups of the quartics X_t.

Lemma 3.4. *The following assertions hold*:

(i) *One has* $\mathrm{Aut}(X_{1/2}) \cong \mathrm{PSp}_4(\mathbf{F}_3)$, *where* $\mathbf{F}_3$ *is the field of three elements.*

(ii) *One has* $\mathrm{Aut}(X_t) \cong \mathfrak{S}_6$ *provided that* $t \notin \left\{\frac{1}{2}, \infty\right\}$.

(iii) *If* X *is a normal quartic hypersurface with a faithful action of the group* $\mathfrak{S}_6$, *then* X *is isomorphic to one of the quartics* X_t.

Proof. Assertion (i) is well known; see e.g., [Coble 1906].

Take any $t \neq \infty$. Since the quartic X_t is normal by Theorem 3.3, its hyperplane section is the anticanonical class, hence the group $\mathrm{Aut}(X_t)$ is naturally embedded into $\mathrm{PGL}_5(\Bbbk)$. Moreover, one has $\mathfrak{S}_6 \subset \mathrm{Aut}(X_t)$ by the definition of X_t. It follows from the classification of finite subgroups of $\mathrm{PGL}_5(\Bbbk)$ that either $\mathrm{Aut}(X_t) \cong \mathfrak{S}_6$ or $\mathrm{Aut}(X_t) \cong \mathrm{PSp}_4(\mathbf{F}_3)$; see [Feit 1971, Section 8.5]. But the group $\mathrm{PSp}_4(\mathbf{F}_3)$ has a unique invariant quartic hypersurface in $\mathbb{P}^4$, which is the Burkhardt quartic $X_{1/2}$. This proves assertion (ii).

Finally, assume that X is a normal quartic hypersurface invariant under some faithful action of the group $\mathfrak{S}_6$ on $\mathbb{P}^4$. Using the classification of projective representations of the group $\mathfrak{S}_6$ we deduce that this action comes from an irreducible five-dimensional representation of $\mathfrak{S}_6$; in fact, it is enough to look at the classification of projective representations of the smaller group $\mathfrak{A}_6$, which can be found for instance in [Conway et al. 1985, page 5]. The latter $\mathfrak{S}_6$-representation is unique up to an outer automorphism and a sign twist (cf. Lemma 5.12). This implies assertion (iii). □

Corollary 3.5. *We have* $\mathrm{Aut}(\mathscr{Y}) \cong \mathfrak{S}_6 \times \boldsymbol{\mu}_2$.

Proof. The group on the right-hand side acts on $\mathscr{Y}$ by (1.6) and (1.7), and the action is clearly faithful. It remains to show that any automorphism of $\mathscr{Y}$ belongs to this group. For this we note that the morphism $\pi : \mathscr{Y} \to \mathbb{P}^4$ is defined by the ample generator of $\mathrm{Pic}(\mathscr{Y})$. Indeed, $\mathrm{rk}\,\mathrm{Pic}(\mathscr{Y}) = 1$ by Lefschetz hyperplane section theorem (see [Dolgachev 1982, Theorem 4.2.2]), because $\mathscr{Y}$ is a hypersurface in the weighted projective space $\mathbb{P}(2, 1^5)$. The pullback of the hyperplane in $\mathbb{P}^4$ via π is not divisible in $\mathrm{Pic}(\mathscr{Y})$ by degree reasons, and thus generates $\mathrm{Pic}(\mathscr{Y})$. Hence π is equivariant with respect to any automorphism of $\mathscr{Y}$. This induces a homomorphism $\mathrm{Aut}(\mathscr{Y}) \to \mathrm{PGL}_5(\Bbbk)$ whose kernel is generated by the Galois involution σ. The image of the homomorphism is the subgroup of $\mathrm{PGL}_5(\Bbbk)$ that fixes the branch divisor $X_{1/4}$ of π. Moreover, the latter subgroup acts faithfully on $X_{1/4}$, hence is contained in $\mathrm{Aut}(X_{1/4}) \cong \mathfrak{S}_6$. □

For further reference we state here a description of the class groups of X_t.

Lemma 3.6. *The following table lists the ranks of the class groups of the quartics* X_t:

t	$t \notin \mathfrak{D} \cup \{\infty\}$	$t = \frac{1}{4}$	$t = \frac{1}{2}$	$t = \frac{1}{6}$	$t = \frac{7}{10}$
$\mathrm{rk}\,\mathrm{Cl}(X_t)$	6	1	16	11	7

Proof. First, assume $t \notin \mathfrak{D} \cup \{\infty\}$. Let $\tilde{X}_t$ be the blow up of X_t at its singular points. Then $\tilde{X}_t$ is smooth by Theorem 3.3. Now the assertion follows from [Cynk 2001, Theorem 2] and [Beauville 2013, Lemma 2].

The cases $t = \frac{1}{2}$, $t = \frac{1}{6}$, and $t = \frac{7}{10}$, are discussed in [Kaloghiros 2011, Theorem 1.1(iii); Cheltsov and Shramov 2016b, Sections 5–6].

Finally, consider the case $t = \frac{1}{4}$. As it was already mentioned, the Igusa quartic $X_{1/4}$ is projectively dual to the Segre cubic threefold $Z \subset \mathbb{P}^4$. In fact, projective duality gives an $\mathfrak{S}_6$-equivariant birational map $Z \dashrightarrow X_{1/4}$ that blows up 10 ordinary double points of Z and blows down the proper transforms of 15 planes on Z; see e.g., the proof of [Prokhorov 2010, Lemma 3.10]. In particular, one has

$$\operatorname{rk} \operatorname{Cl}(X_{1/4}) = \operatorname{rk} \operatorname{Cl}(Z) + 10 - 15,$$

and since the class group of the Segre cubic Z has rank 6 (see e.g., [Prokhorov 2013, Theorem 7.1]), we obtain $\operatorname{rk} \operatorname{Cl}(X_{1/4}) = 1$. □

In Theorem 5.1 we will describe the action of the group $\mathfrak{S}_6$ on $\operatorname{Cl}(X_t) \otimes \mathbb{Q}$.

3.2. ***Wiman–Edge pencil.*** Consider the projective plane $\mathbb{P}^2$ with homogeneous coordinates w_1, w_2, and w_3 and the following two polynomials of degree six

$$\begin{aligned} \overline{\Phi}_0(w_1, w_2, w_3) &= (w_2^2 - w_3^2)(w_3^2 - w_1^2)(w_1^2 - w_2^2), \\ \overline{\Phi}_\infty(w_1, w_2, w_3) &= w_1^6 + w_2^6 + w_3^6 + (w_1^2 + w_2^2 + w_3^2)(w_1^4 + w_2^4 + w_3^4) - 12 w_1^2 w_2^2 w_3^2. \end{aligned} \tag{3.7}$$

It is easy to see that the sextic curves on $\mathbb{P}^2$ defined by these polynomials are singular at the following four points

$$(1:1:1), \quad (1:-1:-1), \quad (-1:1:-1), \quad (-1:-1:1), \tag{3.8}$$

hence they induce a pair of global sections

$$\Phi_0, \Phi_\infty \in H^0(S, \omega_S^{-2})$$

of the double anticanonical line bundle on the blow up S of $\mathbb{P}^2$ at the points (3.8), i.e., on the quintic del Pezzo surface. By [Edge 1981] the section Φ_∞ is invariant with respect to the action of $\operatorname{Aut}(S) \cong \mathfrak{S}_5$, while the Φ_0 is acted on by $\mathfrak{S}_5$ via the sign character. Therefore, there is an $\mathfrak{S}_5$-invariant pencil of $\mathfrak{A}_5$-invariant curves $\Delta_s \subset S$ given by the equation

$$\Phi_0 + s\Phi_\infty = 0, \quad s \in \Bbbk \cup \{\infty\}. \tag{3.9}$$

As we already mentioned, the curves Δ_s are double anticanonical divisors on S. We refer to the pencil (3.9) as the *Wiman–Edge pencil*. It was studied in various contexts in [Wiman 1896b; Edge 1981; Inoue and Kato 2005; Cheltsov and Shramov 2016a, Section 6.2; Dolgachev et al. 2018; Zamora 2018].

Theorem 3.10 [Edge 1981; Cheltsov and Shramov 2016a, Theorem 6.2.9]. *The Wiman–Edge pencil contains exactly five singular curves*: Δ_0, $\Delta_{\pm 1/\sqrt{125}}$, *and* $\Delta_{\pm 1/\sqrt{-3}}$. *They can be described as follows*:

- Δ_0 *is the union of* 10 *lines on* S; *it has* 15 *singular points.*
- $\Delta_{\pm 1/\sqrt{-3}}$ *are unions of* 5 *smooth conics*; *each of these curves has* 10 *singular points.*
- $\Delta_{\pm 1/\sqrt{125}}$ *are irreducible rational curves*; *each of these curves has* 6 *singular points.*

Every singular point of any of these curves is a node. The group $\mathfrak{A}_5$ *acts transitively on the set of singular points and on the set of irreducible components of each of these curves.*

Remark 3.11. The curves Δ_0 and Δ_∞ in the Wiman–Edge pencil are not just $\mathfrak{A}_5$-invariant, but also $\mathfrak{S}_5$-invariant. The first of them, as we already mentioned, is the union of 10 lines. The other one is a smooth curve of genus 6 known as the Wiman's sextic curve; see [Wiman 1896b; Edge 1981]; it should not be confused with a smooth plane sextic curve studied by Wiman [1896a]. By construction, Δ_∞ admits a faithful action of the group $\mathfrak{S}_5$, and one can show that its full automorphism group is also $\mathfrak{S}_5$.

3.3. ***Preimages of $\mathfrak{S}_6$-invariant quartics in the Coble fourfold.*** Recall that the Coble fourfold $\mathscr{Y}$ is defined as a complete intersection in the weighted projective space $\mathbb{P}(2, 1^6)$ of the hyperplane (1.1) with the hypersurface (1.5). It comes with a double covering $\pi : \mathscr{Y} \to \mathbb{P}^4$ over the projective space in which the pencil $\{X_t\}$ of $\mathfrak{S}_6$-invariant quartics sits, and with the Galois involution $\sigma : \mathscr{Y} \to \mathscr{Y}$ of the double covering.

As in Section 1, we define a pencil of hypersurfaces $\mathscr{X}_\tau \subset \mathscr{Y}$ by (1.12). By definition each of the varieties $\mathscr{X}_\tau$ is $\mathfrak{S}_6$-invariant with respect to the natural $\mathfrak{S}_6$-action. Moreover, $\mathscr{X}_0$ and $\mathscr{X}_\infty$ are invariant under the whole group $\mathrm{Aut}(\mathscr{Y}) = \mathfrak{S}_6 \times \boldsymbol{\mu}_2$.

Lemma 3.12. *For every* $\tau \neq \infty$ *we have*

$$\pi^{-1}(X_{(\tau^2+1)/4}) = \mathscr{X}_\tau \cup \mathscr{X}_{-\tau},$$

and the involution σ *induces an* $\mathfrak{S}_6$*-equivariant isomorphism* $\sigma \colon \mathscr{X}_\tau \to \mathscr{X}_{-\tau}$ *for the natural action of* $\mathfrak{S}_6$. *The map* $\pi : \mathscr{X}_\tau \to X_{(\tau^2+1)/4}$ *is an isomorphism for all* $\tau \neq \infty$, *and the map* $\pi : \mathscr{X}_\infty \to X_\infty$ *factors through the double covering over the quadric* $Q_\infty = (X_\infty)_{\mathrm{red}}$ *defined by* (3.2) *that is branched over* $X_{1/4} \cap Q_\infty$. *The map* π *is* $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$*-equivariant for* $\tau = 0, \infty$ *and* $\mathfrak{S}_6$*-equivariant otherwise.*

Proof. The hypersurface $\pi^{-1}(X_{(\tau^2+1)/4}) \subset \mathscr{Y}$ is defined by the equation

$$(x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_5^4 + x_6^4) - \tfrac{1}{4}(\tau^2+1)(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2)^2 = 0,$$

which in view of the equation (1.5) of $\mathscr{Y}$ can be rewritten as

$$\begin{aligned} 0 &= x_0^2 - \tfrac{\tau^2}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2)^2 \\ &= \big(x_0 + \tfrac{\tau}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2)^2\big)\big(x_0 - \tfrac{\tau}{2}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2)^2\big). \end{aligned}$$

Hence $\pi^{-1}(X_{(\tau^2+1)/4})$ is the union of $\mathscr{X}_\tau$ and $\mathscr{X}_{-\tau}$. The Galois involution σ acts by $x_0 \mapsto -x_0$, hence defines an isomorphism between $\mathscr{X}_\tau$ and $\mathscr{X}_{-\tau}$. To check that the map $\pi : \mathscr{X}_\tau \to X_{(\tau^2+1)/4}$ is an isomorphism, just use (1.12) to express x_0 in terms of other x_i; plugging it into the equation of the Coble fourfold $\mathscr{Y}$, we deduce the equation of the quartic X_t. For $\tau = \infty$ this of course does not work, but the equations of $\mathscr{X}_\infty$ just give

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2 = x_0^2 - (x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_5^4 + x_6^4) = 0$$

which defines a double covering of Q_∞ whose branch locus is $X_0 \cap Q_\infty = X_{1/4} \cap Q_\infty$.

The equivariance of the maps σ and π is obvious. □

Remark 3.13. The singular locus of $\mathscr{X}_\infty$ consists of the unique $\mathfrak{S}_6$-orbit of length 30 that is projected by π to the $\mathfrak{S}_6$-orbit Σ_{30}; see e.g., [Przyjalkowski and Shramov 2016, Section 6].

Now we say a couple of words about the Weil divisor class groups of the threefolds $\mathscr{X}_\tau$. Consider the set

$$\hat{\mathfrak{D}} := \left\{0, \pm 1, \pm\frac{1}{\sqrt{-3}}, \pm\frac{3}{\sqrt{5}}\right\}, \tag{3.14}$$

that is, the preimage of the discriminant set $\mathfrak{D}$ defined in (1.3) under the map (1.11).

Lemma 3.15. *The following table lists the ranks of the class groups of the threefolds* $\mathscr{X}_\tau$:

τ	$\tau \notin \hat{\mathfrak{D}}$	$\tau = 0$	$\tau = \pm 1$	$\tau = \pm\frac{1}{\sqrt{-3}}$	$\tau = \pm\frac{3}{\sqrt{5}}$
$\operatorname{rk}\operatorname{Cl}(\mathscr{X}_\tau)$	6	1	16	11	7

Proof. If we assume that $\tau \neq \infty$, then the assertion follows from Lemma 3.6 in view of Lemma 3.12. For $\tau = \infty$ we argue similarly to the proof of Lemma 3.6 (see the proof of [Przyjalkowski and Shramov 2016, Proposition 6.3]). Let $\tilde{\mathscr{X}}_\infty$ be the blow up of $\mathscr{X}_\infty$ along its singular locus, i.e., the preimage of the $\mathfrak{S}_6$-orbit Σ_{30}; see Remark 3.13. Then $\tilde{\mathscr{X}}_\infty$ is smooth, and one proceeds as in [Cynk 2001, Theorem 2], using the computation of [Beauville 2013, Lemma 2]. □

3.4. *Pencil of Verra threefolds.* We consider the pullbacks $\mathscr{X}_\tau^{5,1}$ and $\mathscr{X}_\tau^{4,2}$ of the threefolds $\mathscr{X}_\tau$ to the resolutions $\mathscr{Y}_{5,1}$ and $\mathscr{Y}_{4,2}$ of singularities of the Coble fourfold, so that $\mathscr{X}_\tau^{5,1} \subset \mathscr{Y}_{5,1}$ and $\mathscr{X}_\tau^{4,2} \subset \mathscr{Y}_{4,2}$ are defined by (1.13). In the next section we will study the first of them, but now let us consider the second one. We assume that the map $\rho_{4,2}$ is defined by (2.23).

To simplify the situation, we consider the images of the threefolds $\mathscr{X}_\tau^{4,2}$ with respect to the contraction $\beta\colon \mathscr{Y}_{4,2} \to \mathbb{P}^2 \times \mathbb{P}^2 = \mathbb{P}(\mathrm{W}_3) \times \mathbb{P}(\mathrm{W}_3)$, see Section 2.1. Define

$$\bar{\mathscr{X}}_\tau^{4,2} = \beta(\mathscr{X}_\tau^{4,2}) \subset \mathbb{P}^2 \times \mathbb{P}^2.$$

As in Section 2.1 we use $(u_1 : u_2 : u_3)$ and $(v_1 : v_2 : v_3)$ for coordinates on the factors of $\mathbb{P}^2 \times \mathbb{P}^2$, and let $\boldsymbol{P}_i = (P_i, P_i)$ with P_i defined by (2.1).

Below we consider divisors of bidegree (2, 2) in $\mathbb{P}^2 \times \mathbb{P}^2$ (and call them *Verra threefolds*) as conic bundles over the first factor. We write their equations as symmetric 3×3-matrices with coefficients being quadratic polynomials in u_1, u_2, u_3. So, if $q(u) = (q_{ij}(u))$ is such a matrix, the corresponding equation is $q(u)(v) := \sum q_{ij}(u) v_i v_j = 0$.

Proposition 3.16. *The subvariety $\overline{\mathscr{X}}^{4,2}_\tau \subset \mathbb{P}^2 \times \mathbb{P}^2$ is a Verra threefold given by the equation*

$$q_0(u)(v) + \tau q_\infty(u)(v) = 0, \tag{3.17}$$

where

$$q_0(u) = \frac{1}{2}\begin{pmatrix} 0 & u_3(u_2-u_1) & u_2(u_1-u_3) \\ u_3(u_2-u_1) & 0 & u_1(u_3-u_2) \\ u_2(u_1-u_3) & u_1(u_3-u_2) & 0 \end{pmatrix}, \quad \textit{and} \tag{3.18}$$

$$q_\infty(u) = \frac{1}{6}\begin{pmatrix} 4(u_2^2-u_2u_3+u_3^2) & u_3(u_1+u_2)-2u_1u_2-2u_3^2 & u_2(u_1+u_3)-2u_1u_3-2u_2^2 \\ u_3(u_1+u_2)-2u_1u_2-2u_3^2 & 4(u_1^2-u_1u_3+u_3^2) & u_1(u_2+u_3)-2u_2u_3-2u_1^2 \\ u_2(u_1+u_3)-2u_1u_3-2u_2^2 & u_1(u_2+u_3)-2u_2u_3-2u_1^2 & 4(u_1^2-u_1u_2+u_2^2) \end{pmatrix}. \tag{3.19}$$

Proof. By (1.12), the variety $\overline{\mathscr{X}}^{4,2}_0$ is given by the equation $x_0 = 0$. Writing the formula for x_0 from (2.23) in the matrix form, we get (3.18). Similarly, $\overline{\mathscr{X}}^{4,2}_\infty$ is given by the equation

$$\tfrac{1}{2}(x_1^2+x_2^2+x_3^2+x_4^2+x_5^2+x_6^2) = 0.$$

Substituting expressions for x_i from (2.23) and rewriting everything in the matrix form, we get (3.19). Therefore, (3.17) is the same as (1.12). $\square$

Remark 3.20. Of course, one can cancel the common factor $\frac{1}{2}$ in (3.18) and (3.19). However, we prefer to keep it so that $q_0(u)(v)$ and $q_\infty(u)(v)$ are the same as the two summands in (1.12).

Since the maps $\beta\colon \mathscr{X}^{4,2}_\tau \to \overline{\mathscr{X}}^{4,2}_\tau$ and $\pi_{4,2} = \pi\circ\rho_{4,2}\colon \mathscr{X}^{4,2}_\tau \to X_{(\tau^2+1)/4}$ are birational for all $\tau \neq \infty$, the projection $p_1\colon \overline{\mathscr{X}}^{4,2}_\tau \to \mathbb{P}^2$ provides every (reduced) $\mathfrak{S}_6$-invariant quartic with a birational structure of a conic bundle. Similarly, the map $p_1\colon \overline{\mathscr{X}}^{4,2}_\infty \to \mathbb{P}^2$ provides a birational structure of a conic bundle on the threefold $\mathscr{X}_\infty$. The explicit formulas of Proposition 3.16 allow to compute their discriminant loci.

Lemma 3.21. *The discriminant curve of the conic bundle $p_1\colon \overline{\mathscr{X}}^{4,2}_\tau \to \mathbb{P}^2$ is the curve $\overline{\Delta}_\tau \subset \mathbb{P}^2$ defined by the equation*

$$(5\tau^2+3)\overline{\Phi}_0 + (\tau^3-\tau)\overline{\Phi}_\infty = 0, \tag{3.22}$$

where $\overline{\Phi}_0$ and $\overline{\Phi}_\infty$ are the sextic polynomials (3.7), *and the coordinates $(w_1 : w_2 : w_3)$ are related to $(u_1 : u_2 : u_3)$ by the formula*

$$u_1 = w_2 + w_3, \quad u_2 = w_1 + w_3, \quad u_3 = w_1 + w_2.$$

Proof. A straightforward computation shows that

$$12\det(q_0(u) + \tau q_\infty(u)) = (5\tau^2+3)\overline{\Phi}_0 + (\tau^3-\tau)\overline{\Phi}_\infty. \qquad \square$$

The drawback of this conic bundle model is the lack of flatness. Indeed, it is easy to see that over each of the points P_i (see (2.1)) the matrix $q_0(u)$ is identically zero, so the fiber of $\overline{\mathscr{X}}^{4,2}_0$ over P_i is the whole $\mathbb{P}^2$. In the next subsection we check that using the resolution $\mathscr{Y}_{5,1}$ of the Coble fourfold, we obtain flat conic bundles.

3.5. ***Pencil of conic bundles over the quintic del Pezzo surface.*** Recall that $\mathscr{X}_\tau^{5,1} \subset \mathscr{Y}_{5,1}$ is defined in (1.13) as the preimage of the threefold $\mathscr{X}_\tau \subset \mathscr{Y}$ under the resolution $\rho_{5,1}\colon \mathscr{Y}_{5,1} \to \mathscr{Y}$. For its investigation it will be very convenient to use explicit formulas of Section 3.4. So, to benefit from those we assume that we are in the situation of Proposition 2.50, i.e., a subgroup $\mathfrak{S}_4 \subset \mathfrak{S}_5$ and a nonstandard embedding $\mathfrak{S}_5 \hookrightarrow \mathfrak{S}_6$ are chosen, the choice of $\rho_{4,2}$ is fixed as in (2.23), the map $\theta_1\colon \mathscr{Y}_{5,1} \dashrightarrow \mathscr{Y}_{4,2}$ is a birational isomorphism for which the outer square of diagram (1.10) commutes, and $\rho_{5,1} = \rho_{4,2} \circ \theta_1$.

Remark 3.23. As we already discussed, for $\tau \neq 0, \infty$ the subvariety $\mathscr{X}_\tau$ is invariant with respect to the *natural* action of $\mathfrak{S}_6$, while the map $\rho_{5,1}\colon \mathscr{Y}_{5,1} \to \mathscr{Y}$ is equivariant with respect to the *twisted* action of $\mathfrak{S}_5 \subset \mathfrak{S}_6$. As a result, the subvariety $\mathscr{X}_\tau^{5,1} \subset \mathscr{Y}^{5,1}$ is only invariant under the action of the subgroup $\mathfrak{A}_6 \cap \mathfrak{S}_5 = \mathfrak{A}_5$, on which the two actions agree. Similarly, the projection $\rho_{5,1}\colon \mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau$ is only $\mathfrak{A}_5$-equivariant. On the other hand, for $\tau = 0$ or $\tau = \infty$, the subvariety $\mathscr{X}_\tau^{5,1} \subset \mathscr{Y}$ is $\mathfrak{S}_5$-invariant and the map $\rho_{5,1}$ is $\mathfrak{S}_5$-equivariant.

Lemma 3.24. *The map $p\colon \mathscr{X}_\tau^{5,1} \to S$ is a flat conic bundle with the discriminant curve $\Delta_{s(\tau)} \subset S$ defined by* (3.9), *where*

$$s(\tau) = \frac{\tau^3 - \tau}{5\tau^2 + 3}. \tag{3.25}$$

The map p is $\mathfrak{A}_5$-equivariant for $\tau \neq 0, \infty$ and $\mathfrak{S}_5$-equivariant for $\tau = 0, \infty$.

Proof. Equivariance of the maps $p\colon \mathscr{X}_\tau^{5,1} \to S$ follows from invariance of $\mathscr{X}_\tau^{5,1}$ discussed in Remark 3.23 and $\mathfrak{S}_5$-equivariance of the $\mathbb{P}^2$-bundle $p\colon \mathscr{Y}_{5,1} \to S$. The restriction of (1.10) gives a commutative diagram

$$\begin{tikzcd}
\mathscr{X}_\tau^{5,1} \arrow[rr, dashed, "\theta_1"] \arrow[dr, "\rho_{5,1}"'] \arrow[dd, "p"'] & & \mathscr{X}_\tau^{4,2} \arrow[dl, "\rho_{4,2}"] \arrow[d, "\beta"] \\
& \mathscr{X}_\tau & \overline{\mathscr{X}}_\tau^{4,2} \arrow[d, "p_1"] \\
S \arrow[rr, "\varphi"] & & \mathbb{P}^2
\end{tikzcd} \tag{3.26}$$

The divisor $\mathscr{X}_\infty^{5,1} \subset \mathscr{Y}$ is the preimage of the quadric threefold $Q_\infty \subset \mathbb{P}(\mathrm{W}_5)$ with respect to the morphism $\pi_{5,1}\colon \mathscr{Y}_{5,1} \to \mathbb{P}(\mathrm{W}_5)$, hence it is the zero locus of a section of the line bundle $\mathscr{O}_{\mathbb{P}_S(\mathscr{U}_3)}(2)$. Since $\mathscr{X}_\tau^{5,1}$ form a pencil, all of them are the zero loci of sections of the same line bundle, hence correspond to symmetric morphisms $\mathscr{U}_3 \to \mathscr{U}_3^\vee$ on S (in particular, $p\colon \mathscr{X}_\tau^{5,1} \to S$ is a conic bundle). Therefore, the discriminant curve of $\mathscr{X}_\tau^{5,1}$ is the zero locus of a morphism

$$\omega_S \cong \det(\mathscr{U}_3) \to \det(\mathscr{U}_3^\vee) \cong \omega_S^{-1},$$

i.e., a double anticanonical divisor.

On the other hand, the above diagram shows that the discriminant locus of $\mathscr{X}_\tau^{5,1}$ contains the proper transform of the discriminant curve $\overline{\Delta}_\tau$ of $\overline{\mathscr{X}}_\tau^{4,2}$ whose equation is (3.22). If $\tau^3 - \tau \neq 0$ it is a sextic curve passing with multiplicity 2 through each of the points P_i, hence its proper transform to S is a curve on S with equation

$$(5\tau^2+3)\Phi_0+(\tau^3-\tau)\Phi_\infty=0, \tag{3.27}$$

i.e., the curve $\Delta_{s(\tau)}$. In the case when $\tau^3-\tau=0$, the curve $\bar{\Delta}_\tau$ is the union of six lines on $\mathbb{P}^2$, and its proper transform on S is the union of six lines on S. But the conic bundle $p\colon \mathscr{X}_\tau^{5,1} \to S$ is $\mathfrak{A}_5$-equivariant. Thus its discriminant curve is $\mathfrak{A}_5$-invariant, and hence it should also contain the other four lines of S. We conclude this case by noting that the sum of the ten lines Δ_0 on S is a double anticanonical divisor, and it is indeed given by the equation (3.27) with $\tau^3-\tau=0$.

It remains to show that the conic bundle is flat. For this we note that a nonflat point of a conic bundle is a point of multiplicity at least 3 on its discriminant curve. But by Theorem 3.10 all singular points of these curves are nodes. □

Before going further, we discuss some properties of the map $s\colon \mathbb{P}^1 \to \mathbb{P}^1$ defined by (3.25).

Lemma 3.28. *The map $s\colon \mathbb{P}^1 \to \mathbb{P}^1$ is a triple covering with simple ramification at four points $\tau=\pm\sqrt{-3}$ and $\tau=\pm 1/\sqrt{5}$.*

Proof. A direct computation. □

In the next table we list some special values of τ together with the values of the functions $s(\tau)$ and $t(\tau)=(\tau^2+1)/4$ at these points.

τ	0	± 1	$\pm\frac{1}{\sqrt{-3}}$	$\boxed{\pm\sqrt{-3}}$	$\mp\frac{3}{\sqrt{5}}$	$\boxed{\pm\frac{1}{\sqrt{5}}}$	∞	$\pm\sqrt{-\frac{3}{5}}$
$s(\tau)$	0		$\mp\frac{1}{\sqrt{-3}}$		$\mp\frac{1}{5\sqrt{5}}$		∞	
$t(\tau)$	$\boxed{\frac{1}{4}}$	$\boxed{\frac{1}{2}}$	$\boxed{\frac{1}{6}}$	$-\frac{1}{2}$	$\boxed{\frac{7}{10}}$	$\frac{3}{10}$	$\boxed{\infty}$	$\frac{1}{10}$

The second row contains the values of the parameter s that correspond to singular members of the Wiman–Edge pencil (see Theorem 3.10) and infinity. The first row contains their preimages; boxed cells mark ramification points of the map $s(\tau)$; see Lemma 3.28. The third row contains the values of the map $t(\tau)$ at these points; boxed cells mark the points of the discriminant set $\mathfrak{D}$ and infinity.

Since the degree of the map s is 3, the same singular curves in the Wiman–Edge pencil may appear as the discriminant loci of the preimages $\mathscr{X}_\tau^{5,1}$ of different quartics X_t. For instance, the Igusa and the Burkhardt quartics both correspond to the union Δ_0 of the ten lines on S. Note also that the quartics $X_{1/6}$ and $X_{7/10}$ share their discriminant curves with nonspecial quartics $X_{-1/2}$ and $X_{3/10}$ respectively. As we will see in Proposition 3.30, these two are characterized by the fact that the corresponding curves in the Wiman–Edge pencil are singular, while the total spaces of the threefolds $\mathscr{X}_\tau^{5,1}$ are smooth. In Section 4 we will see that this subtle difference has a drastic effect on rationality properties.

To proceed we will need the following general result. Its proof can be found in [Beauville 1977, Proposition 1.2] or [Sarkisov 1982, Proposition 1.8], except for the fact that the singularity of $\mathscr{X}_P$ is a node, but this can also be extracted from the arguments in either of these two papers.

Lemma 3.29 [Beauville 1977, Proposition 1.2; Sarkisov 1982, Proposition 1.8]. *Let $p\colon \mathscr{X} \to S$ be a flat conic bundle over a smooth surface S. Assume that its discriminant locus $\Delta \subset S$ has a node at a point $P \in S$. Then $\mathscr{X}$ has a singular point over P if and only if the fiber $\mathscr{X}_P = p^{-1}(P)$ is a conic of corank* 1 (*that is, a union of two distinct lines*), *and in this case the singularity of $\mathscr{X}$ over P is a node at the* (*unique*) *singular point of $\mathscr{X}_P$.*

The next assertion describes the singular loci of the threefolds $\mathscr{X}_\tau^{5,1}$. Recall the morphism $\pi_{5,1}$ defined in (2.34) and the discriminant set $\hat{\mathfrak{D}}$ from (3.14).

Proposition 3.30. *The threefold $\mathscr{X}_\tau^{5,1}$ is smooth for all $\tau \notin \hat{\mathfrak{D}}$ (including $\tau = \infty$). For $\tau \in \hat{\mathfrak{D}}$ the singular locus of $\mathscr{X}_\tau^{5,1}$ is mapped by $\pi_{5,1}$ isomorphically to a subset of $\mathbb{P}^4$ as follows*:

τ	0	± 1	$\pm\frac{1}{\sqrt{-3}}$	$\pm\frac{3}{\sqrt{5}}$
$\pi_{5,1}(\mathrm{Sing}(\mathscr{X}_\tau^{5,1}))$	Υ_{15}	Σ_{15}	Σ_{10}	Σ_6

For $\tau \in \hat{\mathfrak{D}}$ the singularities of $\mathscr{X}_\tau^{5,1}$ form a single $\mathfrak{A}_5$-orbit, every singular point Q of $\mathscr{X}_\tau^{5,1}$ is a node, and the fiber $p^{-1}(p(Q))$ of the conic bundle $p\colon \mathscr{X}_\tau^{5,1} \to S$ passing through Q is the union of two distinct lines intersecting at Q.

Proof. To start with, let us show that for $\tau \neq 0$ the threefold $\mathscr{X}_\tau^{5,1}$ is smooth along the exceptional locus of the morphism $\rho_{5,1}$, which by Proposition 2.44 is the reducible surface

$$\Big(\bigcup_L R_L\Big) \cup \Big(\bigcup_\varphi R_\varphi\Big) = \pi_{5,1}^{-1}(\mathrm{CR}) \subset \mathscr{Y}_{5,1}. \tag{3.31}$$

Recall that each of its irreducible components is a smooth surface in $\mathscr{Y}_{5,1}$ (see Lemmas 2.36 and 2.41). Note that a Cartier divisor in a smooth fourfold is smooth along its intersection with a smooth surface provided that their scheme intersection is a smooth curve. So, it is enough to check that the intersections $\mathscr{X}_\tau^{5,1} \cap R_L$ and $\mathscr{X}_\tau^{5,1} \cap R_\varphi$ are smooth curves for all $\tau \neq 0$. But the divisors $\mathscr{X}_\tau^{5,1}$ form a pencil, and $\mathscr{X}_0^{5,1}$ (which by definition is equal to the ramification divisor of $\pi_{5,1}$) contains all these surfaces. Therefore,

$$\mathscr{X}_\tau^{5,1} \cap R_L = \mathscr{X}_\infty^{5,1} \cap R_L \quad \text{and} \quad \mathscr{X}_\tau^{5,1} \cap R_\varphi = \mathscr{X}_\infty^{5,1} \cap R_\varphi.$$

So, it is enough to show that $\mathscr{X}_\infty^{5,1} \cap R_\varphi$ and $\mathscr{X}_\infty^{5,1} \cap R_L$ are smooth curves. But $\mathscr{X}_\infty^{5,1} = \pi_{5,1}^{-1}(Q_\infty)$, while R_φ and R_L are the preimages of the 15 lines of the Cremona–Richmond configuration CR. The quadric Q_∞ intersects all these lines transversally at two points away from the intersection points of the lines by Remark 3.1, and taking into account Lemma 2.35 and Proposition 2.44 we conclude that $\mathscr{X}_\infty^{5,1} \cap R_L$ is the union of two disjoint lines, and $\mathscr{X}_\infty^{5,1} \cap R_\varphi$ is the union of two disjoint smooth conics.

Since the map $\rho_{5,1}\colon \mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau$ is an isomorphism over $\mathbb{P}^4 \setminus \mathrm{CR}$ (because so is the map $\mathscr{Y}_{5,1} \to \mathscr{Y}$), it follows that for all $\tau \neq 0$ we have

$$\mathrm{Sing}(\mathscr{X}_\tau^{5,1}) = \mathrm{Sing}(\mathscr{X}_\tau) \setminus \mathrm{CR},$$

and in view of Lemma 3.12, Theorem 3.3, Remark 3.13, and Lemma 3.29, we obtain the required description of singularities of $\mathscr{X}_\tau^{5,1}$ for $\tau \neq 0$.

Next, consider the case $\tau = 0$. The map $p\colon \mathscr{X}_0^{5,1} \to S$ is a flat conic bundle with the discriminant locus being the curve Δ_0, i.e., the union of the 10 lines on S. It follows that $\mathscr{X}_0^{5,1}$ is smooth over the complement of the 15 intersection points of the lines on S. Since all these points are nodes of Δ_0, Lemma 3.29 shows that the threefold $\mathscr{X}_0^{5,1}$ has a singularity over such a point P if and only if the conic $(\mathscr{X}_0^{5,1})_P = p^{-1}(P)$ is the union of two distinct lines (and then the singular point is a node located at the intersection point of these lines). Since the 15 intersection points of the lines on S form a single $\mathfrak{A}_5$-orbit (see Theorem 3.10), it is enough to check everything over one of them.

Take the intersection point $P \in S$ such that $\varphi(P) = (0:1:1)$. We know from diagram (3.26) that the conic $(\mathscr{X}_0^{5,1})_P$ is isomorphic to the conic $(\overline{\mathscr{X}}_0^{4,2})_{\varphi(P)}$, hence by Proposition 3.16 it is given by the matrix

$$q_0(0:1:1) = \tfrac{1}{2}\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{3.32}$$

Its rank equals 2, hence $(\overline{\mathscr{X}}_0^{4,2})_{\varphi(P)}$, and thus also $(\mathscr{X}_0^{5,1})_P$, is a union of two lines. Moreover, the intersection point of the irreducible components of $(\overline{\mathscr{X}}_0^{4,2})_{\varphi(P)}$ is the point $(0:1:1)$, and using (2.23) we compute that

$$\bar{\pi}_{4,2}((0:1:1),(0:1:1)) = (2:-1:-1:2:-1:-1) \in \Upsilon_{15}.$$

By $\mathfrak{A}_5$-equivariance of the map $\pi_{5,1}$ and transitivity of $\mathfrak{A}_5$-action on Υ_{15} (see Corollary A.4) we conclude that $\pi_{5,1}(\mathrm{Sing}(\mathscr{X}_0^{5,1})) = \Upsilon_{15}$. □

Corollary 3.33. *For all $\tau \neq 0, \infty$ the morphism $\pi_{5,1}\colon \mathscr{X}_\tau^{5,1} \to X_{(\tau^2+1)/4}$ is birational and small. Also, the morphism $\rho_{5,1}\colon \mathscr{X}_\infty^{5,1} \to \mathscr{X}_\infty$ is birational and small.*

Proof. Indeed, as we have seen in the proof of Proposition 3.30, for $\tau \neq 0$ the nontrivial fibers of $\mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau$ are 30 rational curves, one over each of the 30 intersection points of $\Sigma_{30} = \mathrm{CR} \cap Q_\infty$. Since the map $\pi\colon \mathscr{X}_\tau \to X_{(\tau^2+1)/4}$ is an isomorphism for $\tau \neq \infty$ by Lemma 3.12, the assertion follows. □

Remark 3.34. For $\tau = 0$ the surface (3.31) is equal to the exceptional locus of $\pi_{5,1}\colon \mathscr{X}_0^{5,1} \to X_{1/4}$, hence this morphism is not small, but is still birational.

3.6. *Proofs of Theorems 1.14 and 1.15.* For $\tau \neq 0$ the map $\rho_{5,1}\colon \mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau$ is small and birational by Corollary 3.33. The same argument works for $\rho_{4,2}\colon \mathscr{X}_\tau^{4,2} \to \mathscr{X}_\tau$ without changes. Finally, smoothness of $\mathscr{X}_\tau^{5,1}$ for nonspecial τ is proved in Proposition 3.30. The maps $\rho_{5,1}$ and $\pi \circ \rho_{5,1}$ have required equivariance by Remark 3.23 and Lemma 3.12. The same arguments prove equivariance of the maps $\rho_{4,2}$ and $\pi \circ \rho_{4,2}$. This completes the proof of Theorem 1.14.

Now let us prove Theorem 1.15. By Proposition 3.30 the total spaces of the conic bundles $p\colon \mathscr{X}_\tau^{5,1} \to S$ are smooth for $\tau \notin \hat{\mathfrak{D}}$, so since $\operatorname{rk}\operatorname{Pic}(S) = 5$, to show that p is a standard conic bundle for $\tau \notin \hat{\mathfrak{D}}$ it is enough to check that $\operatorname{rk}\operatorname{Pic}(\mathscr{X}_\tau^{5,1}) = 6$ for these τ. But since the map $\rho_{5,1}\colon \mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau$ is small, we have

$$\operatorname{Pic}(\mathscr{X}_\tau^{5,1}) \cong \operatorname{Cl}(\mathscr{X}_\tau).$$

Thus the assertion of the theorem follows from Lemma 3.15. □

Remark 3.35. Assume the notation of Remark 2.25, and suppose that $t \notin \left\{\frac{1}{4}, \frac{1}{2}, \infty\right\}$. One can check that the restrictions of each hyperplane $H_{ijk} \subset \mathbb{P}^4$ to X_t splits as the union of two smooth quadric surfaces in $H_{ijk} \cong \mathbb{P}^3$. For $t = \frac{1}{4}$ these two quadric surfaces collide into a smooth quadric with a nonreduced structure, and for $t = \frac{1}{2}$ they degenerate into unions of pairs of planes. Considering the preimages of these surfaces on $\mathscr{X}_\tau$, where as usual $t = (\tau^2+1)/4$, and using Remarks 2.25 and 2.49, one can describe the small resolutions $\rho_{4,2}$ and $\rho_{5,1}$ of singularities of $\mathscr{X}_\tau$ as blow ups of certain Weil divisors on $\mathscr{X}_\tau$.

4. Rationality

In this section we provide some applications of the results obtained earlier. Namely, we check that all quartics X_t are unirational, give a new and uniform proof of irrationality of $\mathfrak{S}_6$-invariant quartics X_t for $t \notin \mathfrak{D} \cup \{\infty\}$ (and also of the threefold $\mathscr{X}_\infty$), and rationality of X_t for $t \in \mathfrak{D}$.

4.1. *Unirationality of $\mathfrak{S}_6$-invariant quartics.* We start with a short proof of unirationality of the $\mathfrak{S}_6$-invariant quartics X_t and the threefold $\mathscr{X}_\infty$. The next fact is well known.

Lemma 4.1. *Let V be an irreducible Verra threefold, i.e., an irreducible hypersurface of bidegree $(2,2)$ in $\mathbb{P}^2 \times \mathbb{P}^2$. Then V is unirational.*

Proof. Let $p_i\colon V \to \mathbb{P}^2$, $i = 1, 2$, be the natural projections. Both p_i are (possibly nonflat) conic bundles. Let $L \subset \mathbb{P}^2$ be a general line, and put $T = p_2^{-1}(L)$. Since V is irreducible and L is general, the surface T is irreducible by Bertini's theorem. Also, the map p_2 provides the surface T with a conic bundle structure over $L \cong \mathbb{P}^1$, hence T is rational. Note also that $T = V \cap (\mathbb{P}^2 \times L)$ is a divisor of bidegree $(2,2)$ in $\mathbb{P}^2 \times \mathbb{P}^1$, hence the projection $p_1\colon T \to \mathbb{P}^2$ is dominant (actually, T is a rational 2-section of p_1). Since $p_1\colon V \to \mathbb{P}^2$ is a conic bundle, the standard base change argument implies unirationality of V. □

Combining Lemma 4.1 with Proposition 3.16, we obtain

Corollary 4.2. *The quartics X_t, $t \neq \infty$, and the threefold $\mathscr{X}_\infty$, are unirational.*

Remark 4.3. One can use the same approach to prove rationality of the Burkhardt quartic $X_{1/2}$ (this is a classical fact going back to [Todd 1936]; see also Theorem 4.6 below). For this consider the corresponding Verra threefold $\overline{\mathscr{X}}_1^{4,2} \subset \mathbb{P}^2 \times \mathbb{P}^2$ and let $T = p_2^{-1}(\overline{P_1P_2}) \subset \overline{\mathscr{X}}_1^{4,2}$ be the preimage of the line passing through two of the points (2.1), that is, the line $v_3 = 0$. As before, T is a divisor of bidegree (2, 2) in $\mathbb{P}^2 \times \mathbb{P}^1$. Using (3.18) and (3.19) we can rewrite explicitly its equation $q_0(u)(v_1, v_2, 0) + q_\infty(u)(v_1, v_2, 0) = 0$ as

$$(q_0(u) + q_\infty(u))(v_1, v_2, 0) = \tfrac{2}{3}(u_1v_2 + \omega u_2v_1 + \omega^2 u_3v_1 + \omega u_3v_2)(u_1v_2 + \omega^2 u_2v_1 + \omega u_3v_1 + \omega^2 u_3v_2),$$

where ω is a primitive cubic root of unity. Thus we see that $T = T_1 \cup T_2$, where T_i is a divisor of bidegree (1, 1). In particular, each T_i provides a rational section of the conic bundle $p_1\colon \overline{\mathscr{X}}_1^{4,2} \to \mathbb{P}^2$ and rationality of $\overline{\mathscr{X}}_1^{4,2}$ follows. Since the threefold $\overline{\mathscr{X}}_1^{4,2}$ is birational to the quartic $X_{1/2}$, the rationality of the latter follows as well.

4.2. ***Irrationality of nonspecial $\mathfrak{S}_6$-invariant quartics.*** Beauville [2013] proved that the quartic X_t is irrational provided that $t \notin \mathfrak{D} \cup \{\infty\}$ by using the $\mathfrak{S}_6$-action on the intermediate Jacobian of a suitable resolution of singularities of X_t. By [Beauville 2013], the intermediate Jacobian J_t of the blow up of the 30 singular points of X_t is five-dimensional, and the action of $\mathfrak{S}_6$ on J_t is faithful; on the other hand, if it is a product of Jacobians of curves, it cannot have a faithful $\mathfrak{S}_6$-action. Irrationality of the threefold $\mathscr{X}_\infty$ was proved using the same approach in [Przyjalkowski and Shramov 2016, Proposition 6.3]. With the help of the conic bundle structure on these varieties constructed in Theorem 1.15, we can give another proof of their irrationality.

Theorem 4.4. *If $t \notin \mathfrak{D} \cup \{\infty\}$, then X_t is irrational. Also, the variety $\mathscr{X}_\infty$ is irrational.*

Proof. By Theorem 1.14 it is enough to show that the threefold $\mathscr{X}_\tau^{5,1}$ is irrational for $\tau \notin \hat{\mathfrak{D}}$. By Theorem 1.15 the map $p\colon \mathscr{X}_\tau^{5,1} \to S$ is a standard conic bundle with the nodal discriminant curve Δ_s contained in the linear system $|-2K_S|$. Here $s = s(\tau)$ is given by the formula (1.16). The conic bundle p induces a double cover $\hat{\Delta}_s \to \Delta_s$ that by Lemma 3.29 is branched only over the nodes of the curve Δ_s. Applying [Beauville 1977, Proposition 2.8], we see that the intermediate Jacobian of the threefold $\mathscr{X}_\tau^{5,1}$ is isomorphic as a principally polarized abelian variety to the Prym variety $\mathrm{Prym}(\hat{\Delta}_s, \Delta_s)$. Now [Shokurov 1983, Main Theorem] implies that $\mathrm{Prym}(\hat{\Delta}_s, \Delta_s)$ is not a product of Jacobians of curves, hence $\mathscr{X}_\tau^{5,1}$ is irrational. □

Remark 4.5. The intermediate Jacobian of $\mathscr{X}_\tau^{5,1}$ can be described fairly explicitly. For instance, it was observed by Dimitri Markushevich that it is isogenous to the fifth power of an elliptic curve (whose j-invariant depends on τ). Here is a sketch of his argument. Let $\hat{\mathscr{X}}_\tau \to \mathscr{X}_\tau$ be a minimal $\mathfrak{S}_6$-equivariant resolution of singularities, so that $\mathrm{Jac}(\hat{\mathscr{X}}_\tau) \cong \mathrm{Jac}(\mathscr{X}_\tau^{5,1})$. The action of the group $\mathfrak{S}_6$ on $\mathrm{Jac}(\hat{\mathscr{X}}_\tau)$ can be lifted to an action of the semidirect product $\mathfrak{S}_6 \ltimes H^3(\hat{\mathscr{X}}_\tau, \mathbb{Z})$ on $H^3(\hat{\mathscr{X}}_\tau, \mathbb{C})$. Now [Bernstein and Schwarzman 2006, Theorem 3.1] proves that there is an $\mathfrak{S}_6$-equivariant isomorphism

$$H^3(\hat{\mathscr{X}}_\tau, \mathbb{Q}) \cong \mathbb{Q}(\mathfrak{S}_6) \oplus \lambda\mathbb{Q}(\mathfrak{S}_6),$$

where $\mathbb{Q}(\mathfrak{S}_6)$ is the root lattice associated with the group $\mathfrak{S}_6$ considered as a Weyl group of Dynkin type A_5, and $\lambda = \lambda(\tau)$ is a complex number with positive imaginary part. Therefore, $\mathrm{Jac}(\hat{\mathscr{X}}_\tau)$ is isogenous to $E(\lambda)^5$, where $E(\lambda) = \mathbb{C}/(\mathbb{Z} \oplus \lambda\mathbb{Z})$.

Note by the way, that there is another popular family of threefolds with five-dimensional intermediate Jacobians, namely, smooth cubic threefolds. However, it was pointed out by Beauville that the quartics X_t are not birational to smooth cubics. Indeed, if a quartic X_t is birational to a smooth cubic threefold Y, then the intermediate Jacobian $J(Y)$ is isomorphic to J_t, and thus there is a faithful $\mathfrak{S}_6$-action on $J(Y)$ (note that J_t must coincide with its Griffiths component in this case). Torelli's theorem for smooth cubic threefolds (see [Beauville 1982, Proposition 6]) implies that there is a faithful $\mathfrak{S}_6$-action on Y itself, which is impossible, because the only cubic threefold with a faithful $\mathfrak{S}_6$-action is the Segre cubic that has ten singular points.

It would be interesting to find out if the quartics X_t with $t \notin \mathfrak{D}$ are stably rational or not.

4.3. *Rationality of special $\mathfrak{S}_6$-invariant quartics.* The result of Theorem 4.4 is sharp: the threefolds $X_{1/2}$, $X_{1/4}$, $X_{1/6}$, and $X_{7/10}$ are rational. In fact, rationality of the Burkhardt quartic $X_{1/2}$ was proved by Todd [1936] (see also Remark 4.3), rationality of the Igusa quartic $X_{1/4}$ follows from rationality of its projectively dual variety (which is the Segre cubic), and rationality of the quartics $X_{1/6}$ and $X_{7/10}$ is also known; see [Todd 1933; 1935; Cheltsov and Shramov 2016b]. However, using our results one can give a uniform proof of rationality of all these threefolds; this proof does not use explicit rationality constructions.

Theorem 4.6. *The quartics $X_{1/2}$, $X_{1/4}$, $X_{1/6}$, and $X_{7/10}$ are rational.*

Proof. Suppose that $\tau \in \hat{\mathfrak{D}}$, so that $t \in \mathfrak{D}$ and $s \in \{0, \pm 1/\sqrt{125}, \pm 1/\sqrt{-3}\}$, where as usual $t = (\tau^2+1)/4$ and $s = s(\tau)$; see (3.25). By Theorem 1.14 it is enough to show that $\mathscr{X}_\tau^{5,1}$ is rational.

Consider the conic bundle $p\colon \mathscr{X}_\tau^{5,1} \to S$. The singular locus of its discriminant Δ_s is a finite set of nodes; see Theorem 3.10. Actually, by Lemma 3.24 the set $\operatorname{Sing}(\Delta_s)$ consists of 15 points when $t = \frac{1}{4}$ or $t = \frac{1}{2}$, of 10 points when $t = \frac{1}{6}$, and of 6 points when $t = \frac{7}{10}$. We also know from Proposition 3.30 that all singularities of $\mathscr{X}_\tau^{5,1}$ are nodes, and for every singular point Q of $\mathscr{X}_\tau^{5,1}$ the fiber $p^{-1}(p(Q))$ is the union of two lines, with Q being their intersection point.

The conic bundle p is not standard because the threefold $\mathscr{X}_\tau^{5,1}$ is singular, so we start by transforming it to a standard one. Let $\nu\colon \tilde{S} \to S$ be the blow up of the quintic del Pezzo surface S at $\operatorname{Sing}(\Delta_s)$, and consider the base change $p'\colon \mathscr{X}_\tau^{5,1} \times_S \tilde{S} \to \tilde{S}$ of the conic bundle p. Its discriminant curve is the preimage on $\tilde{S}$ of the discriminant curve of p. In particular, it contains all exceptional curves of the blow up ν as irreducible components of multiplicity 2, and the corank of the fibers of p' over the points of each of these curves equals 1. Modifying the conic bundle along these lines as in [Sarkisov 1982, Lemma 1.14] (see also [Debarre and Kuznetsov 2020, Section 2.5]), we can get rid of the corresponding components of the discriminant. In other words, we obtain a small birational map

$$\mathscr{X}_\tau^{5,1} \times_S \tilde{S} \dashrightarrow \tilde{\mathscr{X}}_\tau^{5,1} \tag{4.7}$$

over $\tilde{S}$, such that the threefold $\tilde{\mathscr{X}}_\tau^{5,1}$ comes with a flat conic bundle $\tilde{p}\colon \tilde{\mathscr{X}}_\tau^{5,1} \to \tilde{S}$ whose discriminant curve is the proper transform $\tilde{\Delta}_s \subset \tilde{S}$ of Δ_s with respect to ν. In particular, the curve $\tilde{\Delta}_s$ is smooth (hence also $\tilde{\mathscr{X}}_\tau^{5,1}$ is smooth), and by Theorem 3.10 has ten connected components when $t = \frac{1}{4}$ or $t = \frac{1}{2}$, five components when $t=\frac{1}{6}$, and just one component when $t=\frac{7}{10}$. Moreover, every connected component of $\tilde{\Delta}_s$ is rational.

Since $\tilde{\Delta}_s$ is smooth, the conic bundle $\tilde{p}$ has only simple degenerations. In particular, it induces an étale double covering over $\tilde{\Delta}_s$. Since every connected component $\tilde{\Delta}_s^{(i)} \subset \tilde{\Delta}_s$ is smooth and rational, the double covering is trivial, hence the preimage $\tilde{p}^{-1}(\tilde{\Delta}_s^{(i)})$ consists of two irreducible components

$$\tilde{p}^{-1}(\tilde{\Delta}_s^{(i)}) = \Theta_i' \cup \Theta_i'',$$

each being a $\mathbb{P}^1$-bundle over $\tilde{\Delta}_s^{(i)}$. Choosing for each i one of them and contracting all chosen components simultaneously over $\tilde{S}$ (see [Sarkisov 1982, 1.17]), we obtain a commutative diagram

$$\begin{tikzcd}
\tilde{\mathscr{X}}^{5,1}_\tau \arrow[rr] \arrow[dr, "\tilde{p}"'] & & \bar{\mathscr{X}}^{5,1}_\tau \arrow[dl, "\bar{p}"] \\
& \tilde{S} &
\end{tikzcd}$$

Here the horizontal arrow is a birational morphism, and $\bar{p}$ is an everywhere nondegenerate conic bundle. Since $\tilde{S}$ is a rational surface, its Brauer group is trivial, hence this $\mathbb{P}^1$-bundle is a projectivization of a vector bundle, hence birational to $\tilde{S} \times \mathbb{P}^1$, hence rational. This means that $\mathscr{X}^{5,1}_\tau$ is also rational. □

Remark 4.8. The birational transformation $\mathscr{X}^{5,1}_\tau \dashrightarrow \mathscr{X}^{5,1}_\tau \times_S \tilde{S} \dashrightarrow \tilde{\mathscr{X}}^{5,1}_\tau$ can be described very explicitly; see Construction I in the proof of [Cheltsov et al. 2019b, Theorem 4.2]. It is a composition of the blow ups of all singular points $Q \in \mathscr{X}^{5,1}_\tau$ followed by the Atiyah flops in the union of proper transforms of the two irreducible components of the conic $p^{-1}(p(Q))$; see Proposition 3.30.

The construction that we used in the proof of Theorem 4.6 has the following consequence, which we will need in Section 5. Recall the notation of (1.17).

Corollary 4.9. *For $\tau \in \hat{\mathfrak{D}}$ the relative divisor class group $\mathrm{Cl}(\mathscr{X}^{5,1}_\tau/S) \otimes \mathbb{Q}$ has the following structure as a representation of the group $\mathfrak{A}_5$:*

τ	0	± 1	$\pm\frac{1}{\sqrt{-3}}$	$\pm\frac{3}{\sqrt{5}}$
$\mathrm{Cl}(\mathscr{X}^{5,1}_\tau/S) \otimes \mathbb{Q}$	$\mathbf{1} \oplus \mathrm{Ind}^{\mathfrak{A}_5}_{\mathfrak{A}_{3,2}}(\mathbf{1})$	$\mathbf{1} \oplus \mathrm{Ind}^{\mathfrak{A}_5}_{\mathfrak{A}_{3,2}}(-\mathbf{1})$	$\mathbf{1} \oplus \mathrm{Ind}^{\mathfrak{A}_5}_{\mathfrak{A}_4}(\mathbf{1})$	$\mathbf{1} \oplus \mathbf{1}$

Here $\mathrm{Ind}^{\mathfrak{A}_5}_G$ stands for the induction functor from the subgroup $G = \mathfrak{A}_4$ or $G = \mathfrak{A}_{3,2} \cong \mathfrak{S}_3$ in $\mathfrak{A}_5$, while $\mathbf{1}$ stands for the trivial representation, and $-\mathbf{1}$ stands for the sign representation of $\mathfrak{S}_3$. The first summand $\mathbf{1}$ in each cell is generated by the canonical class of $\mathscr{X}^{5,1}_\tau$.

Proof. The canonical class $K_{\mathscr{X}^{5,1}_\tau}$ is invariant with respect to the group action, hence generates a trivial subrepresentation in $\mathrm{Cl}(\mathscr{X}^{5,1}_\tau/S) \otimes \mathbb{Q}$. Consider the quotient

$$\mathrm{Cl}_0(\mathscr{X}^{5,1}_\tau/S) \otimes \mathbb{Q} := (\mathrm{Cl}(\mathscr{X}^{5,1}_\tau/S) \otimes \mathbb{Q})/\mathbb{Q}K_{\mathscr{X}^{5,1}_\tau}.$$

To describe it we use the notation introduced in the proof of Theorem 4.6. First, we have

$$\mathrm{Cl}_0(\mathscr{X}_\tau^{5,1}/S) \cong \mathrm{Cl}_0((\mathscr{X}_\tau^{5,1} \times_S \tilde{S})/\tilde{S}).$$

Furthermore, since (4.7) is a small birational map, we have

$$\mathrm{Cl}_0((\mathscr{X}_\tau^{5,1} \times_S \tilde{S})/\tilde{S}) \cong \mathrm{Cl}_0(\tilde{\mathscr{X}}_\tau^{5,1}/\tilde{S}).$$

Finally, it is clear that $\mathrm{Cl}_0(\tilde{\mathscr{X}}_\tau^{5,1}/\tilde{S}) \otimes \mathbb{Q}$ is contained in an $\mathfrak{A}_5$-equivariant exact sequence

$$0 \to \bigoplus \mathbb{Q}[\tilde{\Delta}_{s(\tau)}^{(i)}] \to \bigoplus (\mathbb{Q}[\Theta_i'] \oplus \mathbb{Q}[\Theta_i'']) \to \mathrm{Cl}_0(\tilde{\mathscr{X}}_\tau^{5,1}/\tilde{S}) \otimes \mathbb{Q} \to 0,$$

where we sum up over the set of irreducible components of $\tilde{\Delta}_{s(\tau)}$, and the first map takes the class $[\tilde{\Delta}_{s(\tau)}^{(i)}] \in \mathrm{Pic}(\tilde{S})$ to $[\Theta_i'] + [\Theta_i''] \in \mathrm{Pic}(\tilde{\mathscr{X}}_\tau^{5,1})$. It follows that $\mathrm{Cl}_0(\tilde{\mathscr{X}}_\tau^{5,1}/\tilde{S}) \otimes \mathbb{Q}$ has the basis $[\Theta_i'] - [\Theta_i'']$, and the group $\mathfrak{A}_5$ permutes the basis vectors, possibly changing their signs.

Recall that the group $\mathfrak{A}_5$ acts transitively on the set of irreducible components of $\tilde{\Delta}_{s(\tau)}$ by Theorem 3.10. Let $G \subset \mathfrak{A}_5$ be the stabilizer of some irreducible component of $\tilde{\Delta}_{s(\tau)}$, say, of $\tilde{\Delta}_{s(\tau)}^{(0)}$. The action of G on the set $\{\Theta_0', \Theta_0''\}$ defines a homomorphism $\upsilon\colon G \to \{\pm 1\}$, i.e., a one-dimensional representation of G, and we conclude that

$$\mathrm{Cl}_0(\tilde{\mathscr{X}}_\tau^{5,1}/\tilde{S}) \otimes \mathbb{Q} \cong \mathrm{Ind}_G^{\mathfrak{A}_5}(\upsilon).$$

So, it remains to identify the possible stabilizers G for various τ, and the homomorphisms υ.

When $\tau = \pm 3/\sqrt{5}$, the curve $\tilde{\Delta}_{s(\tau)}$ is irreducible, hence G is the whole group $\mathfrak{A}_5$, and since it has no nontrivial one-dimensional representations, we conclude that

$$\mathrm{Cl}_0(\tilde{\mathscr{X}}_{\pm 3/\sqrt{5}}^{5,1}/\tilde{S}) \otimes \mathbb{Q} \cong \mathrm{Ind}_{\mathfrak{A}_5}^{\mathfrak{A}_5}(\mathbf{1}) \cong \mathbf{1}.$$

When $\tau = \pm 1/\sqrt{-3}$, the curve $\tilde{\Delta}_{s(\tau)}$ has five components, G is the subgroup $\mathfrak{A}_4$ of $\mathfrak{A}_5$, and since again it has no nontrivial one-dimensional representations, we conclude that

$$\mathrm{Cl}_0(\tilde{\mathscr{X}}_{\pm 1/\sqrt{-3}}^{5,1}/\tilde{S}) \otimes \mathbb{Q} \cong \mathrm{Ind}_{\mathfrak{A}_4}^{\mathfrak{A}_5}(\mathbf{1}).$$

When $\tau = 0$ or $\tau = \pm 1$, the curve $\tilde{\Delta}_{s(\tau)}$ has ten components (corresponding to the lines on S) and G is the subgroup $\mathfrak{A}_{3,2} \cong \mathfrak{S}_3$ of $\mathfrak{A}_5$. It remains to show that it fixes the components Θ_0' and Θ_0'' when $\tau = 0$, and swaps them when $\tau = \pm 1$.

The stabilizer $\mathfrak{A}_{3,2}$ of a line $L \subset S$ permutes three points of its intersection with other lines on S. Each of these points, in its turn, is stabilized by a transposition in $\mathfrak{A}_{3,2} \cong \mathfrak{S}_3$. So, it is enough to check how these transpositions act on Θ_0' and Θ_0''.

Consider the point $P = (0:1:1)$ as in the proof of Proposition 3.30. Then it is easy to see that the subgroup of $\mathfrak{A}_5$ that preserves both lines passing through P is generated by the automorphism

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

of order two of the plane, while the fiber $p^{-1}(P)$ is given by (3.32) in the case $\tau = 0$, and by

$$q_0(0:1:1) + q_\infty(0:1:1) = \frac{1}{3}\begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}$$

in the case $\tau = 1$. Now verifying that g fixes the components of the conic $p^{-1}(P)$ if $\tau = 0$ and swaps the components if $\tau = 1$ is straightforward.

The computation in the case $\tau = -1$ is similar to that in the case $\tau = 1$. □

The part of the above argument that identifies the relative class group of a conic bundle in terms of the induced representation is completely general and can be proved for any conic bundle with only simple degenerations, and for an arbitrary group acting on it.

5. Representation structure of the class groups

The main result of this section is the description of the $\mathfrak{S}_6$-action on the class groups of the Coble fourfold and of the quartics X_t, and its applications to the equivariant birational geometry of these varieties. We will be mostly interested in the quartics X_t with $t \neq \frac{1}{4}, \infty$, because the quartic $X_{1/4}$ has nonisolated singularities, and at the same time its class group is not very intriguing by Lemma 3.6 (see Remark 5.24 below), while the quartic X_∞ is nonreduced; however, we will also perform the same computations for the threefold $\mathscr{X}_\infty$.

5.1. ***The result and its applications.*** We start by stating our main result and its consequences. We will use the following notation for representations of the symmetric groups. For each partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)$ of an integer n (i.e., a nonincreasing sequence of positive integers summing up to n) we denote by

$$\mathrm{R}(\lambda) = \mathrm{R}(\lambda_1, \lambda_2, \dots, \lambda_r)$$

the irreducible $\mathbb{Q}$-representation of the group $\mathfrak{S}_n$ as described in [Fulton and Harris 1991, Section 4.1]. For instance, $\mathrm{R}(n)$ is the trivial representation, while $\mathrm{R}(1^n)$ is the sign representation. Note that the standard permutation representation is the direct sum $\mathrm{R}(n) \oplus \mathrm{R}(n-1, 1)$.

We denote by $\mathrm{R}(\lambda) \boxtimes 1$ and $\mathrm{R}(\lambda) \boxtimes (-1)$ the representations of the group $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$, which are isomorphic to $\mathrm{R}(\lambda)$ when restricted to $\mathfrak{S}_6$ and on which the nontrivial element of $\boldsymbol{\mu}_2$ acts by 1 or -1, respectively.

Theorem 5.1. *The group* $\mathrm{Cl}(\mathscr{Y})$ *is torsion free and there are the following isomorphisms of* $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$*-representations*:

$$\mathrm{Cl}(\mathscr{Y}) \otimes \mathbb{Q} \cong \mathrm{Cl}(\mathscr{X}_\infty) \otimes \mathbb{Q} \cong (\mathrm{R}(6) \boxtimes 1) \oplus (\mathrm{R}(3,3) \boxtimes (-1)).$$

In particular, for the natural action of $\mathfrak{S}_6$ *there are isomorphisms of* $\mathfrak{S}_6$*-representations*

$$\mathrm{Cl}(\mathscr{Y}) \otimes \mathbb{Q} \cong \mathrm{Cl}(\mathscr{X}_\infty) \otimes \mathbb{Q} \cong \mathrm{R}(6) \oplus \mathrm{R}(3,3),$$

while for the twisted action of $\mathfrak{S}_6$ *there are isomorphisms of* $\mathfrak{S}_6$*-representations*

$$\mathrm{Cl}(\mathscr{Y})\otimes\mathbb{Q}\cong\mathrm{Cl}(\mathscr{X}_\infty)\otimes\mathbb{Q}\cong\mathrm{R}(6)\oplus\mathrm{R}(2,2,2).$$

Finally, there are the following isomorphisms of $\mathfrak{S}_6$*-representations*:

$$\begin{aligned}\mathrm{Cl}(X_t)\otimes\mathbb{Q}&\cong\mathrm{R}(6)\oplus\mathrm{R}(3,3) &&\text{for } t\notin\mathfrak{D}\cup\{\infty\}\\ \mathrm{Cl}(X_{1/2})\otimes\mathbb{Q}&\cong\mathrm{R}(6)\oplus\mathrm{R}(3,3)\oplus\mathrm{R}(3,1^3),\\ \mathrm{Cl}(X_{1/6})\otimes\mathbb{Q}&\cong\mathrm{R}(6)\oplus\mathrm{R}(3,3)\oplus\mathrm{R}(2,2,2),\\ \mathrm{Cl}(X_{7/10})\otimes\mathbb{Q}&\cong\mathrm{R}(6)\oplus\mathrm{R}(3,3)\oplus\mathrm{R}(1^6).\end{aligned}$$

The proof of Theorem 5.1 takes the next subsection, and now we discuss its applications to equivariant birational geometry.

Recall that an n-dimensional variety X with an action of a group G is G*-rational* if there exists a G-equivariant birational map between X and $\mathbb{P}^n$ for some action of G on $\mathbb{P}^n$. Also recall that a G-equivariant morphism $\phi\colon X\to S$ of normal varieties acted on by a finite group G is called a G*-Mori fiber space*, if X has terminal singularities, one has $\mathrm{rk}\,\mathrm{Pic}(X)^G=\mathrm{rk}\,\mathrm{Cl}(X)^G$, the fibers of ϕ are connected and of positive dimension, the anticanonical divisor $-K_X$ is ϕ-ample, and the relative G-invariant Picard rank $\mathrm{rk}\,\mathrm{Pic}(X/S)^G$ equals 1.

The first application of Theorem 5.1 is due to the following expectation, which is proved in several particular cases, see [Mella 2004; Shramov 2008; Cheltsov et al. 2019a, Proof of Theorem 1.1].

Conjecture 5.2. *Let X be either a nodal quartic threefold*, *or a nodal double covering of a smooth three-dimensional quadric branched over its intersection with a quartic. Let G be a finite subgroup in* $\mathrm{Aut}(X)$ *such that*

$$\mathrm{rk}\,\mathrm{Cl}(X)^G=1.$$

If there is a G-equivariant birational map $X\dashrightarrow X'$, *where $X'\to S'$ is a G-Mori fiber space*, *then $X\cong X'$. In particular*, *X is not G-rational.*

Of course, this applies to each of the $\mathfrak{S}_6$-invariant quartics X_t with $t\neq\frac{1}{4},\infty$, and to the threefold $\mathscr{X}_\infty$ as well. For each subgroup $G\subset\mathfrak{S}_6$ the rank of the invariant class group $\mathrm{Cl}(X_t)^G$ can be easily computed from the result of Theorem 5.1 by restricting the representation and computing the multiplicity of the trivial summand. We used GAP [2017] to perform this computation; see http://www.mi-ras.ru/~akuznet/GAP-code/rk-code.txt for the source code. To state our result in a precise form we first introduce our notation for the (conjugacy classes of) subgroups of $\mathfrak{S}_6$ that will be used until the end of Section 5.1. We will also use notation (1.17).

Notation 5.3. Given a subgroup $G\subset\mathfrak{S}_6$ we denote by $\overline{G}\subset\mathfrak{S}_6$ the image of G under an outer automorphism of $\mathfrak{S}_6$ (it is well-defined up to conjugation). Furthermore, if $G_1\subset\mathfrak{S}_{n_1},\dots,G_r\subset\mathfrak{S}_{n_r}$ are subgroups and $n_1+\dots+n_r\leqslant 6$, then by $G_1\times\dots\times G_r$ we denote the corresponding subgroup in

$$\mathfrak{S}_{n_1}\times\dots\times\mathfrak{S}_{n_r}\cong\mathfrak{S}_{n_1,\dots,n_r}\subset\mathfrak{S}_{n_1+\dots+n_r}\subset\mathfrak{S}_6.$$

Next, we use the notation $\boldsymbol{\mu}_d[c_1, \dots, c_r]$ for a cyclic subgroup of order d generated by a permutation of cycle type $[c_1, \dots, c_r]$. We abbreviate $\boldsymbol{\mu}_5[5]$ to just $\boldsymbol{\mu}_5$.

By V_4 we denote the Klein four-group, i.e., the unique subgroup of order 4 in $\mathfrak{A}_4 \subset \mathfrak{S}_4$. By $\mathrm{V}_{4,2}$ we denote a subgroup of $\mathfrak{S}_{4,2} \subset \mathfrak{S}_6$ whose projection to the first factor $\mathfrak{S}_4$ gives an isomorphism with V_4, while the projection to the second factor $\mathfrak{S}_2$ is surjective.

By D_{2n} we denote the dihedral group of order $2n$. It is naturally embedded into the group $\mathfrak{S}_n$, so for $n \leqslant 6$ it is a subgroup of $\mathfrak{S}_6$; note that $\mathrm{D}_{12} = \overline{\mathfrak{S}}_{3,2}$.

There are four conjugacy classes of subgroups isomorphic to D_8 in $\mathfrak{S}_6$. They can be described as follows. The first class contains subgroups of (the standard) $\mathfrak{S}_4$ in $\mathfrak{S}_6$; according to the above conventions, we will refer to subgroups from this conjugacy class simply as D_8. There are three nontrivial homomorphisms

$$\upsilon_\circ \colon \mathrm{D}_8 \to \boldsymbol{\mu}_2, \quad \upsilon_+ \colon \mathrm{D}_8 \to \boldsymbol{\mu}_2, \quad \upsilon_\times \colon \mathrm{D}_8 \to \boldsymbol{\mu}_2,$$

determined by their kernels

$$\mathrm{Ker}(\upsilon_\circ) = \boldsymbol{\mu}_4[4], \quad \mathrm{Ker}(\upsilon_+) = \mathrm{V}_4, \quad \mathrm{Ker}(\upsilon_\times) = \mathfrak{S}_{2,2}.$$

Thinking of these as of subgroups of symmetries of a square, the first is generated by rotations, the second by reflections with respect to the lines passing through the middle points of its opposite sides, and the third by reflections with respect to the diagonals; this is the mnemonics for the notation $\circ$, $+$, and $\times$. We denote by D_8°, D_8^+, and $\mathrm{D}_8^\times$ the images of the map

$$\mathrm{D}_8 \xrightarrow{(\mathrm{id}, \upsilon)} \mathfrak{S}_4 \times \boldsymbol{\mu}_2 \cong \mathfrak{S}_{4,2} \subset \mathfrak{S}_6$$

for $\upsilon = \upsilon_\circ$, υ_+, and $\upsilon_\times$, respectively. Note that $\mathrm{D}_8^\circ = \overline{\mathrm{D}}_8$.

The intersection $\mathfrak{S}_5 \cap \overline{\mathfrak{S}}_5$ of a standard and a nonstandard subgroups $\mathfrak{S}_5$ is a subgroup of order 20 isomorphic to $\boldsymbol{\mu}_5 \rtimes \boldsymbol{\mu}_4$, and such groups form a unique conjugacy class of subgroups of order 20 in $\mathfrak{S}_6$. Also, the subgroups $\boldsymbol{\mu}_4 \times \boldsymbol{\mu}_2$, $\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3$, D_{10}, $\mathrm{D}_8 \times \mathfrak{S}_2$, $(\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_2$, $(\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_4$, and $\mathfrak{S}_{3,3} \rtimes \boldsymbol{\mu}_2$ of $\mathfrak{S}_6$ are unique up to conjugation.

Finally, recall the definitions (1.6) of the natural and (1.8) of the twisted actions of $\mathfrak{S}_6$ on the Coble fourfold $\mathscr{Y}$ and on the threefold $\mathscr{X}_\infty \subset \mathscr{Y}$. Theorem 5.1 implies:

Corollary 5.4. Figure 1 *contains a complete list* (*ordered by cardinality*) *of subgroups* $G \subset \mathfrak{S}_6$ *such that* $\mathrm{rk}\,\mathrm{Cl}(X)^G = 1$, *where* X *is either* X_t, *or* $\mathscr{X}_\infty$, *or* $\mathscr{Y}$. *If* X *is either* $\mathscr{X}_\infty$ *or* $\mathscr{Y}$, *and* G *is any subgroup of* $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$ *that contains the second factor*, *then one also has* $\mathrm{rk}\,\mathrm{Cl}(X)^G = 1$.

In particular, Conjecture 5.2 suggests that the varieties listed in Corollary 5.4 are not G-rational with respect to the corresponding groups.

Another interesting case of G-equivariant behavior arises when $\mathrm{rk}\,\mathrm{Cl}(X)^G = 2$. The following result is well known to experts.

X, *action of* $\mathfrak{S}_6$	G
X_t, $t \notin \mathfrak{D} \cup \{\infty\}$; $\mathscr{X}_\infty$, *natural action*; $\mathscr{Y}$, *natural action*	$\mathfrak{S}_6, \mathfrak{A}_6, \mathfrak{S}_5, \overline{\mathfrak{S}}_5, \mathfrak{S}_{3,3} \rtimes \boldsymbol{\mu}_2, \mathfrak{A}_5, \mathfrak{S}_{4,2}, \overline{\mathfrak{S}}_{4,2}, (\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_4,$ $\overline{\mathfrak{S}}_{3,3}, \mathfrak{S}_4, \overline{\mathfrak{S}}_4, \mathfrak{A}_{4,2}, \mathfrak{A}_4 \times \mathfrak{S}_2, \boldsymbol{\mu}_5 \rtimes \boldsymbol{\mu}_4, \overline{\mathfrak{S}_3 \times \boldsymbol{\mu}_3}, \mathrm{D}_8 \times \mathfrak{S}_2, \mathfrak{A}_4,$ $\overline{\mathfrak{S}}_{3,2}, \boldsymbol{\mu}_4 \times \boldsymbol{\mu}_2, \mathrm{V}_4 \times \boldsymbol{\mu}_2, \overline{\mathrm{D}}_8, \mathrm{D}_8^\times, \overline{\mathfrak{S}}_3, \mathrm{V}_{4,2}$
$\mathscr{X}_\infty$, *twisted action*; $\mathscr{Y}$, *twisted action*	$\mathfrak{S}_6, \mathfrak{A}_6, \mathfrak{S}_5, \mathfrak{S}_{3,3} \rtimes \boldsymbol{\mu}_2, \mathfrak{A}_5, \mathfrak{S}_{4,2}, (\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_4, \mathfrak{S}_{3,3}, \mathfrak{S}_4,$ $\mathfrak{A}_{4,2}, \mathfrak{A}_4 \times \mathfrak{S}_2, \mathfrak{S}_3 \times \boldsymbol{\mu}_3, \mathfrak{S}_{3,2}, \mathfrak{A}_4, \mathfrak{S}_3, \boldsymbol{\mu}_6[3, 2]$
$X_{1/2}$	$\mathfrak{S}_6, \mathfrak{A}_6, \mathfrak{S}_5, \overline{\mathfrak{S}}_5, \mathfrak{S}_{3,3} \rtimes \boldsymbol{\mu}_2, \mathfrak{A}_5, \mathfrak{S}_{4,2}, \overline{\mathfrak{S}}_{4,2}, (\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_4, \overline{\mathfrak{S}}_{3,3}, \mathfrak{S}_4,$ $\mathfrak{A}_{4,2}, \boldsymbol{\mu}_5 \rtimes \boldsymbol{\mu}_4, \mathrm{D}_8 \times \mathfrak{S}_2$
$X_{1/6}$	$\mathfrak{S}_6, \mathfrak{A}_6, \mathfrak{S}_5, \mathfrak{S}_{3,3} \rtimes \boldsymbol{\mu}_2, \mathfrak{A}_5, \mathfrak{S}_{4,2}, (\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_4, \mathfrak{S}_4, \mathfrak{A}_{4,2}, \mathfrak{A}_4 \times \mathfrak{S}_2, \mathfrak{A}_4$
$X_{7/10}$	$\mathfrak{S}_6, \mathfrak{S}_5, \overline{\mathfrak{S}}_5, \mathfrak{S}_{3,3} \rtimes \boldsymbol{\mu}_2, \mathfrak{S}_{4,2}, \overline{\mathfrak{S}}_{4,2}, \overline{\mathfrak{S}}_{3,3}, \mathfrak{S}_4, \overline{\mathfrak{S}}_4, \mathfrak{A}_4 \times \mathfrak{S}_2, \boldsymbol{\mu}_5 \rtimes \boldsymbol{\mu}_4,$ $\overline{\mathfrak{S}_3 \times \boldsymbol{\mu}_3}, \mathrm{D}_8 \times \mathfrak{S}_2, \overline{\mathfrak{S}}_{3,2}, \boldsymbol{\mu}_4 \times \boldsymbol{\mu}_2, \overline{\mathrm{D}}_8, \mathrm{D}_8^\times, \mathrm{V}_4 \times \boldsymbol{\mu}_2, \overline{\mathfrak{S}}_3, \mathrm{V}_{4,2}$

Figure 1. Subgroups $G \subset \mathfrak{S}_6$ such that $\operatorname{rk} \operatorname{Cl}(X)^G = 1$, where X is either X_t, or $\mathscr{X}_\infty$, or $\mathscr{Y}$.

Proposition 5.5 (cf. [Corti 1995; Hacon and McKernan 2013]). *Let X be a terminal Fano variety (so that, in particular, the canonical class K_X is a $\mathbb{Q}$-Cartier divisor). Let G be a finite subgroup in* $\operatorname{Aut}(X)$ *such that* $\operatorname{rk} \operatorname{Cl}(X)^G = 2$ *and* $\operatorname{rk} \operatorname{Pic}(X)^G = 1$. *Then there exists a unique G-equivariant diagram*:

$$\begin{tikzcd}[column sep=small] & X'_+ \arrow[dl, "p_+"'] & & X_+ \arrow[ll, dashed, "\psi_+"'] \arrow[rr, dashed, "\iota"] \arrow[dr, "f_+"'] & & X_- \arrow[rr, dashed, "\psi_-"] \arrow[dl, "f_-"] & & X'_- \arrow[dr, "p_-"] & \\ Z_+ & & & & X & & & & Z_- \end{tikzcd} \tag{5.6}$$

Here $X_\pm$ are varieties with terminal singularities such that

$$\operatorname{rk} \operatorname{Pic}(X_\pm)^G = \operatorname{rk} \operatorname{Cl}(X_\pm)^G = 2, \quad \operatorname{rk} \operatorname{Pic}(X_\pm/X)^G = 1,$$

the maps $f_\pm$ are small birational morphisms, the map ι is a nontrivial G-flop, the maps $\psi_\pm$ are small and birational (and possibly are just isomorphisms), the varieties $X'_\pm$ have terminal singularities,

$$\operatorname{rk} \operatorname{Pic}(X'_\pm)^G = \operatorname{rk} \operatorname{Cl}(X'_\pm)^G = 2,$$

and each of the maps $p_\pm$, is either a $K_{X'_\pm}$-negative divisorial contraction onto a terminal Fano variety $Z_\pm$ with $\operatorname{rk} \operatorname{Cl}(Z_\pm)^G = 1$, *or a G-Mori fibration.*

The diagram (5.6) is a special case of a so-called *G-Sarkisov link* (that is a G-equivariant version of a usual Sarkisov link; see e.g., [Corti 1995, Definition 3.4] or [Cheltsov 2005, Theorem 1.6.14] for notation). One sometimes says that the link (5.6) is *centered at X*.

Theorem 5.1 allows us to write down a complete list of subgroups $G \subset \mathfrak{S}_6$ for which Proposition 5.5 can be used (as before, we obtained it with the help of GAP [2017]; see http://www.mi-ras.ru/~akuznet/GAP-code/rk-code.txt for the source code).

X, *action of* $\mathfrak{S}_6$	G
X_t, $t \notin \mathfrak{D} \cup \{\infty\}$; $\mathscr{X}_\infty$, *natural action*; $\mathscr{Y}$, *natural action*	$\overline{\mathfrak{A}}_5$, $\mathfrak{S}_{3,3}$, $\overline{\mathfrak{A}}_{4,2}$, $\overline{\mathfrak{A}_4 \times \mathfrak{S}_2}$, $\mathfrak{S}_3 \times \boldsymbol{\mu}_3$, $(\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_2$, $\mathfrak{S}_{3,2}$, D_{10}, $\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3$, D_8, D_8^+, $\boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$, $\mathfrak{S}_3$, $\mathfrak{A}_{3,2}$, $\boldsymbol{\mu}_6[6]$, $\boldsymbol{\mu}_6[3, 2]$, $\boldsymbol{\mu}_5$, $\boldsymbol{\mu}_4[4]$, $\boldsymbol{\mu}_4[4, 2]$, $\boldsymbol{\mu}_2[2, 2] \times \boldsymbol{\mu}_2[2]$, $\boldsymbol{\mu}_3[3]$, $\boldsymbol{\mu}_2[2, 2, 2]$
$\mathscr{X}_\infty$, *twisted action*; $\mathscr{Y}$, *twisted action*	$\overline{\mathfrak{S}}_5$, $\overline{\mathfrak{A}}_5$, $\overline{\mathfrak{S}}_{4,2}$, $\overline{\mathfrak{S}}_{3,3}$, $\overline{\mathfrak{A}}_{4,2}$, $\overline{\mathfrak{A}_4 \times \mathfrak{S}_2}$, $\boldsymbol{\mu}_5 \rtimes \boldsymbol{\mu}_4$, $\overline{\mathfrak{S}}_3 \times \boldsymbol{\mu}_3$, $(\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_2$, $\mathrm{D}_8 \times \mathfrak{S}_2$, D_{10}, $\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3$, D_8, $\mathrm{D}_8^\times$, D_8^+, $\boldsymbol{\mu}_4 \times \boldsymbol{\mu}_2$, $\boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$, $\mathfrak{A}_{3,2}$, $\boldsymbol{\mu}_5$, $\boldsymbol{\mu}_4[4, 2]$, $\boldsymbol{\mu}_2[2] \times \boldsymbol{\mu}_2[2]$, $\boldsymbol{\mu}_3[3]$
$X_{1/2}$	$\overline{\mathfrak{A}}_5$, $\mathfrak{S}_{3,3}$, $\overline{\mathfrak{S}}_4$, $\overline{\mathfrak{A}}_{4,2}$, $\mathfrak{A}_4 \times \mathfrak{S}_2$, $\overline{\mathfrak{A}_4 \times \mathfrak{S}_2}$, $\overline{\mathfrak{S}}_3 \times \boldsymbol{\mu}_3$, $(\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_2$, $\mathfrak{S}_{3,2}$, $\overline{\mathfrak{S}}_{3,2}$, $\mathfrak{A}_4$, D_{10}, D_8, $\overline{\mathrm{D}}_8$, $\mathrm{D}_8^\times$, D_8^+, $\boldsymbol{\mu}_4 \times \boldsymbol{\mu}_2$, $\boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$, $\mathrm{V}_4 \times \boldsymbol{\mu}_2$
$X_{1/6}$	$\overline{\mathfrak{S}}_5$, $\overline{\mathfrak{S}}_{4,2}$, $\mathfrak{S}_{3,3}$, $\overline{\mathfrak{S}}_{3,3}$, $\boldsymbol{\mu}_5 \rtimes \boldsymbol{\mu}_4$, $\mathfrak{S}_3 \times \boldsymbol{\mu}_3$, $\overline{\mathfrak{S}}_3 \times \boldsymbol{\mu}_3$, $\mathrm{D}_8 \times \boldsymbol{\mu}_2$, $\mathfrak{S}_{3,2}$, $\mathrm{D}_8^\times$, $\boldsymbol{\mu}_4 \times \boldsymbol{\mu}_2$, $\mathfrak{S}_3$, $\boldsymbol{\mu}_6[3, 2]$
$X_{7/10}$	$\mathfrak{A}_6$, $\mathfrak{A}_5$, $\mathfrak{S}_{3,3}$, $(\boldsymbol{\mu}_3 \times \boldsymbol{\mu}_3) \rtimes \boldsymbol{\mu}_4$, $\mathfrak{A}_{4,2}$, $\overline{\mathfrak{A}_4 \times \mathfrak{S}_2}$, $\mathfrak{S}_3 \times \boldsymbol{\mu}_3$, $\mathfrak{S}_{3,2}$, $\mathfrak{A}_4$, D_8, $\boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$, $\mathfrak{S}_3$, $\boldsymbol{\mu}_6[6]$, $\boldsymbol{\mu}_6[3, 2]$, $\boldsymbol{\mu}_4[4]$, $\boldsymbol{\mu}_2[2, 2] \times \boldsymbol{\mu}_2[2]$, $\boldsymbol{\mu}_2[2, 2, 2]$

Figure 2. Subgroups $G \subset \mathfrak{S}_6$ such that $\operatorname{rk} \mathrm{Cl}(X)^G = 2$, where X is either X_t, or $\mathscr{X}_\infty$, or $\mathscr{Y}$.

Corollary 5.7. Figure 2 *contains a complete list* (*ordered by cardinality*) *of subgroups* $G \subset \mathfrak{S}_6$ *such that* $\operatorname{rk} \mathrm{Cl}(X)^G = 2$, *where* X *is either* X_t, *or* $\mathscr{X}_\infty$, *or* $\mathscr{Y}$. *In particular, for each of these varieties there is a* G*-Sarkisov link* (5.6) *centered at* X *with respect to the corresponding groups.*

Example 5.8. If $t \notin \mathfrak{D} \cup \{\infty\}$ and $G = \overline{\mathfrak{A}}_5$, the G-Sarkisov link (5.6) is obtained by restricting the diagram (2.48):

$$\begin{tikzcd}
 & \mathscr{X}_\tau^{5,1} \arrow[dl, "p"'] \arrow[dr, "\pi\circ\rho_{5,1}"] \arrow[rr, dashed, "\iota"] & & \mathscr{X}_{-\tau}^{5,1} \arrow[dl, "\pi\circ\rho_{5,1}"'] \arrow[dr, "p"] & \\
S & & X_t & & S
\end{tikzcd}$$

Here $t = (\tau^2 + 1)/4$, ι is the restriction of the map $\rho_{5,1}^{-1} \circ \sigma \circ \rho_{5,1}$ to $\mathscr{X}_\tau^{5,1}$ (it is a composition of 30 Atiyah flops), and $\psi_\pm$ are the identity maps. The map ι can be also defined as the map induced by an action of an odd permutation in the subgroup $\overline{\mathfrak{S}}_5 \subset \mathfrak{S}_6$ containing $\overline{\mathfrak{A}}_5$.

Example 5.9. If $G = \overline{\mathfrak{S}}_5$, then the G-Sarkisov link (5.6) for $\mathscr{X}_\infty$ comes from a restriction of the commutative diagram (2.48) to $\mathscr{X}_\infty$ (recall that $\mathscr{X}_\infty$ is $\mathrm{Aut}(\mathscr{Y})$-invariant).

5.2. ***Class group computation.*** In this section we prove Theorem 5.1. We start with a description of the $\mathfrak{S}_5$-action on the Picard group of the quintic del Pezzo surface.

Lemma 5.10. *There is an isomorphism of* $\mathfrak{S}_5$*-representations*

$$\mathrm{Pic}(S) \otimes \mathbb{Q} \cong \mathrm{R}(5) \oplus \mathrm{R}(4, 1).$$

Proof. The surface S can be obtained as a blow up of $\mathbb{P}^2$ in four points, and this blow up is $\mathfrak{S}_4$-invariant. Therefore, one has

$$(\mathrm{Pic}(S) \otimes \mathbb{Q})|_{\mathfrak{S}_4} \cong \mathrm{R}(4) \oplus \mathrm{R}(4) \oplus \mathrm{R}(3, 1).$$

Here the first summand is the pullback of the line class, and the last two form the permutation representation spanned by the classes of the exceptional divisors of the blow up. Now the assertion easily follows, since

$$\mathrm{R}(5)|_{\mathfrak{S}_4} \cong \mathrm{R}(4), \quad \mathrm{R}(4, 1)|_{\mathfrak{S}_4} \cong \mathrm{R}(4) \oplus \mathrm{R}(3, 1), \tag{5.11}$$

and moreover, by Pieri's rule [Fulton and Harris 1991, Exercise 4.44] the irreducible $\mathfrak{S}_5$-representations R(5) and R(4, 1) are the only ones that restrict to $\mathfrak{S}_4$ as sums of R(4)'s and R(3, 1)'s. □

Further on we will use a similar argument to describe an $\mathfrak{S}_6$-representation from its restriction to a nonstandard subgroup $\mathfrak{S}_5$. For this the following calculation is quite useful.

Lemma 5.12. *The following table contains all irreducible representations V of $\mathfrak{S}_6$, their images $\overline{V}$ under an outer automorphism of $\mathfrak{S}_6$, and the restrictions of V and $\overline{V}$ to a standard subgroup $\mathfrak{S}_5$.*

$\dim V$	V	$\overline{V}$	$V\|_{\mathfrak{S}_5}$	$\overline{V}\|_{\mathfrak{S}_5}$
1	R(6)		R(5)	
1	R(1^6)		R(1^5)	
5	R(5, 1)	R(2^3)	R(5) ⊕ R(4, 1)	R(2^2, 1)
5	R(2, 1^4)	R(3^2)	R(2, 1^3) ⊕ R(1^5)	R(3, 2)
9	R(4, 2)		R(4, 1) ⊕ R(3, 2)	
9	R(2^2, 1^2)		R(2^2, 1) ⊕ R(2, 1^3)	
10	R(4, 1^2)	R(3, 1^3)	R(4, 1) ⊕ R(3, 1^2)	R(3, 1^2) ⊕ R(2, 1^3)
16	R(3, 2, 1)		R(3, 2) ⊕ R(3, 1^2) ⊕ R(2^2, 1)	

Proof. The restrictions to $\mathfrak{S}_5$ are computed by Pieri's rule, so we only need to explain the action of an outer automorphism. For this note that an outer automorphism acts on the conjugacy classes of $\mathfrak{S}_6$ by swapping the following cycle types

$$[2] \leftrightarrow [2, 2, 2], \quad [3] \leftrightarrow [3, 3], \quad [6] \leftrightarrow [3, 2],$$

and fixing the other types. By using the character table of $\mathfrak{S}_6$ (see for instance [James and Liebeck 1993, Example 19.17]) it is then straightforward to check that an outer automorphism swaps

$$\mathrm{R}(5, 1) \leftrightarrow \mathrm{R}(2, 2, 2), \quad \mathrm{R}(2, 1^4) \leftrightarrow \mathrm{R}(3, 3), \quad \mathrm{R}(4, 1^2) \leftrightarrow \mathrm{R}(3, 1^3),$$

and fixes the other irreducible representations. □

Now we are ready to prove the part of Theorem 5.1 concerning the Coble fourfold.

Proposition 5.13. *The group* $\mathrm{Cl}(\mathscr{Y})$ *is torsion free, and there is an isomorphism*

$$\mathrm{Cl}(\mathscr{Y})\otimes\mathbb{Q}\cong(\mathrm{R}(6)\boxtimes 1)\oplus(\mathrm{R}(3,3)\boxtimes(-1))$$

of representations of the group $\mathrm{Aut}(\mathscr{Y})\cong\mathfrak{S}_6\times\boldsymbol{\mu}_2$.

Proof. Since $\mathscr{Y}_{5,1}\to\mathscr{Y}$ is a small $\mathfrak{S}_5$-equivariant resolution, we have an $\mathfrak{S}_5$-equivariant isomorphism $\mathrm{Cl}(\mathscr{Y})\cong\mathrm{Pic}(\mathscr{Y}_{5,1})$ with respect to the twisted action of a nonstandard subgroup $\mathfrak{S}_5$. Since $\mathscr{Y}_{5,1}$ is a $\mathbb{P}^2$-bundle over the quintic del Pezzo surface S, we have an $\mathfrak{S}_5$-equivariant direct sum decomposition

$$\mathrm{Pic}(\mathscr{Y}_{5,1})=\mathbb{Z}H\oplus p^*(\mathrm{Pic}(S)).$$

Here the first summand is generated by the pullback of the hyperplane class of $\mathbb{P}^4$ under the map $\pi\circ\rho_{5,1}$, and so is $\mathfrak{S}_5$-invariant. This proves that $\mathrm{Cl}(\mathscr{Y})$ is torsion free.

Furthermore, it follows from Lemma 5.10 that there is an isomorphism of $\mathfrak{S}_5$-representations

$$(\mathrm{Cl}(\mathscr{Y})\otimes\mathbb{Q})|_{\mathfrak{S}_5}\cong\mathrm{R}(5)\oplus\mathrm{R}(5)\oplus\mathrm{R}(4,1).$$

Since the embedding of $\mathfrak{S}_5\hookrightarrow\mathfrak{S}_6$ is nonstandard, it follows from Lemma 5.12 that

$$(\mathrm{Cl}(\mathscr{Y})\otimes\mathbb{Q})|_{\mathfrak{S}_6}\cong\mathrm{R}(6)\oplus\mathrm{R}(2,2,2);$$

we emphasize the fact that this isomorphism holds for the *twisted* action of $\mathfrak{S}_6$ on $\mathscr{Y}$. The first summand $\mathrm{R}(6)$ is generated by the class H, hence lifts to $\mathrm{R}(6)\boxtimes 1$ as a representation of $\mathfrak{S}_6\times\boldsymbol{\mu}_2$. Since the quotient of $\mathscr{Y}$ by the Galois involution σ is $\mathbb{P}^4$ and its class group is of rank 1, it follows that the action of $\boldsymbol{\mu}_2$ on the second summand $\mathrm{R}(2,2,2)$ is nontrivial. Hence the *natural* action of $\mathfrak{S}_6$ on the second summand is obtained from $\mathrm{R}(2,2,2)$ by the sign twist, i.e., the corresponding representation is $\mathrm{R}(3,3)$ (recall that the sign twist modifies an irreducible representation by a transposition of its partition), and the assertion of the proposition follows. $\square$

Below we will also need to describe certain $\mathfrak{S}_5$-representations from their restrictions to $\mathfrak{A}_5$. For this the following calculation is useful. Denote by R_1, R_3', R_3'', R_4, and R_5 the irreducible representations of the group $\mathfrak{A}_5$ of dimensions 1, 3, 3, 4, and 5, respectively; see for instance [Fulton and Harris 1991, Exercise 3.5].

Lemma 5.14. *The following table contains all irreducible representation of* $\mathfrak{S}_5$ *and their restrictions to* $\mathfrak{A}_5$.

$\mathrm{R}(\lambda)$	$\mathrm{R}(5)$ $\mathrm{R}(1^5)$	$\mathrm{R}(4,1)$ $\mathrm{R}(2,1^3)$	$\mathrm{R}(3,2)$ $\mathrm{R}(2^2,1)$	$\mathrm{R}(3,1^2)$
$\mathrm{R}(\lambda)\vert_{\mathfrak{A}_5}$	R_1	R_4	R_5	$R_3'\oplus R_3''$

Proof. It is enough to know that a restriction of an $\mathfrak{S}_5$-representation $\mathrm{R}(\lambda)$ to $\mathfrak{A}_5$ contains the trivial subrepresentation R_1 if and only if $\mathrm{R}(\lambda)$ is trivial or is the sign representation, i.e., if $\lambda=(5)$ or $\lambda=(1^5)$.

This follows from Frobenius reciprocity, because

$$\mathrm{Ind}_{\mathfrak{A}_5}^{\mathfrak{S}_5}(R_1) \cong \mathrm{R}(5) \oplus \mathrm{R}(1^5).$$

With this in mind, there is only one way to represent the dimensions of $\mathrm{R}(\lambda)$ as sums of dimensions of irreducible $\mathfrak{A}_5$-representations. It remains to notice that the $\mathfrak{S}_5$-representation $\mathrm{R}(3, 1^2)$ is defined over $\mathbb{Q}$, while both three-dimensional $\mathfrak{A}_5$-representations R_3' and R_3'' are not, so the restriction of $\mathrm{R}(3, 1^2)$ to $\mathfrak{A}_5$ splits as $R_3' \oplus R_3''$. □

Now we are almost ready to attack the class groups of the quartics X_t. For each τ we have a natural composition

$$\mathrm{Cl}(\mathscr{Y}) \cong \mathrm{Cl}(\mathscr{Y} \setminus \mathrm{CR}) \cong \mathrm{Pic}(\mathscr{Y} \setminus \mathrm{CR}) \xrightarrow{\mathsf{res}} \mathrm{Pic}(\mathscr{X}_\tau \setminus \mathrm{CR}) \hookrightarrow \mathrm{Cl}(\mathscr{X}_\tau \setminus \mathrm{CR}) \cong \mathrm{Cl}(\mathscr{X}_\tau). \tag{5.15}$$

Here res denotes the restriction map. The first and the last isomorphisms take place since the Cremona–Richmond configuration $\mathrm{CR} = \mathrm{Sing}(\mathscr{Y})$ has codimension greater than 1 both in $\mathscr{Y}$ and $\mathscr{X}_\tau$, and the second isomorphism follows from smoothness of $\mathscr{Y} \setminus \mathrm{CR}$.

Lemma 5.16. *For all $\tau \neq 0$ the composition $\mathrm{Cl}(\mathscr{Y}) \to \mathrm{Cl}(\mathscr{X}_\tau)$ of the maps in (5.15) is an $\mathfrak{S}_6$-equivariant embedding with respect to the natural action of $\mathfrak{S}_6$. For $\tau = \infty$ it is an $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$-equivariant embedding. Moreover, for $\tau \notin \hat{\mathfrak{D}}$ it is an isomorphism.*

Proof. All the maps in (5.15) are equivariant with respect to the natural action of $\mathfrak{S}_6$ (or of the whole group $\mathfrak{S}_6 \times \boldsymbol{\mu}_2$ in case $\tau = \infty$), hence so is the composition, and it remains to prove injectivity. For this we forget about the $\mathfrak{S}_6$-action and consider the diagram

$$\begin{tikzcd} \mathrm{Pic}(\mathscr{Y}_{5,1}) \arrow[r, "\mathsf{res}"] \arrow[d, "(\rho_{5,1})_*"'] & \mathrm{Pic}(\mathscr{X}_\tau^{5,1}) \arrow[d, "(\rho_{5,1})_*"] \\ \mathrm{Cl}(\mathscr{Y}) \arrow[r] & \mathrm{Cl}(\mathscr{X}_\tau) \end{tikzcd} \tag{5.17}$$

which is easily seen to be commutative. The vertical arrows are isomorphisms, since the birational maps $\rho_{5,1} \colon \mathscr{Y}_{5,1} \to \mathscr{Y}$ and $\rho_{5,1} \colon \mathscr{X}_\tau^{5,1} \to \mathscr{X}_\tau$ for $\tau \neq 0$ are small by Theorems 1.9 and 1.14. So, it is enough to check that the morphism res is injective, which is obvious, since $\mathscr{Y}_{5,1}$ is a $\mathbb{P}^2$-bundle over S and $\mathscr{X}_\tau^{5,1}$ is a (flat) conic bundle inside $\mathscr{Y}_{5,1}$.

Moreover, for $\tau \notin \hat{\mathfrak{D}}$ the conic bundle is standard, hence the image of the top arrow is a sublattice of index 2 or 1, depending on whether the conic bundle has a rational section or not. Since we also know from [Beauville 2013] or Theorem 4.4 that for $\tau \notin \hat{\mathfrak{D}}$ the threefold $\mathscr{X}_\tau^{5,1}$ is not rational, we conclude that the conic bundle $p \colon \mathscr{X}_\tau^{5,1} \to S$ has no rational sections, and thus res is an isomorphism. □

Remark 5.18. Recall that by Lemma 3.12 for $\tau \neq 0, \infty$ one has an isomorphism $\mathscr{X}_\tau \cong X_t$ for $t = (\tau^2+1)/4$. Thus Proposition 5.13 and Lemma 5.16 provide a description of $\mathrm{Cl}(X_t)$ for all $t \notin \mathfrak{D} \cup \{\infty\}$.

It remains to analyze the class groups of the special quartics X_t.

We can think of the map (5.15) as of a map $\mathrm{Cl}(\mathscr{Y}) \to \mathrm{Cl}(X_t)$; this map is $\mathfrak{S}_6$-equivariant, where the action of $\mathfrak{S}_6$ on $\mathscr{Y}$ is natural. We denote the cokernel of this map by

$$\mathrm{ExCl}(X_t) := \mathrm{Cl}(X_t)/\,\mathrm{Cl}(\mathscr{Y}),$$

and refer to this group as the *excess class group* of X_t. To prove Theorem 5.1 we need to compute the latter group for $t = \frac{1}{2}$, $\frac{1}{6}$, and $\frac{7}{10}$ as an $\mathfrak{S}_6$-representation. For this we need a couple of observations.

Lemma 5.19. *For a standard subgroup* $\mathfrak{S}_4 \subset \mathfrak{S}_6$ *we have* $\operatorname{rk}\mathrm{Cl}(X_t)^{\mathfrak{S}_4} = 1$ *for any* $t \neq \infty$. *In particular, we have* $\operatorname{rk}\mathrm{ExCl}(X_t)^{\mathfrak{S}_4} = 0$ *for any* $t \neq \infty$.

Proof. We may assume that $\mathfrak{S}_4$ preserves the homogeneous coordinates x_5 and x_6 on $\mathbb{P}^5$. Denote $\boldsymbol{p}_i := x_1^i + \cdots + x_6^i$. Consider the quotients $\mathbb{P}^5/\mathfrak{S}_4$ and $X_t/\mathfrak{S}_4$. Then

$$\mathbb{P}^5/\mathfrak{S}_4 \cong \mathbb{P}(1, 1, 1, 2, 3, 4),$$

where the weighted homogeneous coordinates of weights 1, 1, 1, 2, 3, and 4 correspond to the $\mathfrak{S}_4$-invariants x_5, x_6, $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$, and $\boldsymbol{p}_4$, respectively. The quotient variety $X_t/\mathfrak{S}_4$ is given in $\mathbb{P}(1, 1, 1, 2, 3, 4)$ by the equations

$$\boldsymbol{p}_1 = \boldsymbol{p}_4 - t\,\boldsymbol{p}_2^2 = 0,$$

so that $X_t/\mathfrak{S}_4 \cong \mathbb{P}(1, 1, 2, 3)$. Therefore, we have $\operatorname{rk}\mathrm{Cl}(X_t)^{\mathfrak{S}_4} = \operatorname{rk}\mathrm{Cl}(X_t/\mathfrak{S}_4) = 1$; see for instance [Fulton 1984, 1.7.5]. Since also $\operatorname{rk}\mathrm{Cl}(\mathscr{Y})^{\mathfrak{S}_4} = 1$ (see Corollary 5.4), it follows that $\operatorname{rk}\mathrm{ExCl}(X_t)^{\mathfrak{S}_4} = 0$. □

Remark 5.20 [Cheltsov et al. 2019a, Remark 2.11]. An argument similar to the proof of Lemma 5.19 was (incorrectly!) used in the proof of [Cheltsov and Shramov 2014, Theorem 1.20] for the standard subgroup $\mathfrak{A}_{4,2} \cong \mathfrak{S}_4$ in $\mathfrak{S}_6$ to deduce that $\operatorname{rk}\mathrm{Cl}(X_{1/2})^{\mathfrak{A}_6} = 1$. However, the assertion is correct: it was later obtained in [Cheltsov et al. 2019a, Corollary 2.10] by a different method. Using Theorem 5.1 we can find this rank as well: indeed, one has $\operatorname{rk}\mathrm{Cl}(X_{1/2})^{\mathfrak{A}_6} = \operatorname{rk}\mathrm{Cl}(X_{1/2})^{\mathfrak{A}_{4,2}} = 1$ by Corollary 5.4.

Lemma 5.21. *For a nonstandard subgroup* $\mathfrak{S}_5 \subset \mathfrak{S}_6$ *we have* $\operatorname{rk}\mathrm{Cl}(X_{1/6})^{\mathfrak{S}_5} = 2$. *In particular, we have* $\operatorname{rk}\mathrm{ExCl}(X_{1/6})^{\mathfrak{S}_5} = 1$.

Proof. By [Cheltsov and Shramov 2016b, Section 6] the quartic $X_{1/6}$ is $\mathfrak{S}_5$-equivariantly isomorphic away from codimension 2 to the blow up $\hat{X}_{1/6}$ of ten lines in $\mathbb{P}^3$, that form a so-called double-five configuration. Therefore we have $\mathrm{Cl}(X_{1/6}) \cong \mathrm{Cl}(\hat{X}_{1/6})$ as $\mathfrak{S}_5$-representations. Furthermore, the group $\mathfrak{S}_5$ acts transitively on this configuration of lines, hence $\operatorname{rk}\mathrm{Cl}(\hat{X}_{1/6})^{\mathfrak{S}_5} = 2$. Since also $\operatorname{rk}\mathrm{Cl}(\mathscr{Y})^{\mathfrak{S}_5} = 1$ (see Corollary 5.4, and keep in mind that according to Notation 5.3 the nonstandard subgroup $\mathfrak{S}_5 \subset \mathfrak{S}_6$ is denoted by $\overline{\mathfrak{S}}_5$), it follows that $\operatorname{rk}\mathrm{ExCl}(X_{1/6})^{\mathfrak{S}_5} = 1$. □

Now we are ready to describe the excess class groups for the special quartics.

Proposition 5.22. *There are the following isomorphisms of* $\mathfrak{S}_6$*-representations*:

$$\mathrm{ExCl}(X_{1/2}) \otimes \mathbb{Q} \cong \mathrm{R}(3, 1^3), \quad \mathrm{ExCl}(X_{1/6}) \otimes \mathbb{Q} \cong \mathrm{R}(2, 2, 2), \quad \mathrm{ExCl}(X_{7/10}) \otimes \mathbb{Q} \cong \mathrm{R}(1^6).$$

Proof. We replace the quartics $X_{1/2}$, $X_{1/6}$, and $X_{7/10}$ by their partial resolutions of singularities $\mathscr{X}_1^{5,1}$, $\mathscr{X}_{1/\sqrt{-3}}^{5,1}$, and $\mathscr{X}_{3/\sqrt{5}}^{5,1}$, respectively. Similarly to the proof of Lemma 5.16, we obtain isomorphisms of $\mathfrak{A}_5$-representations

$$\mathrm{Cl}(\mathscr{X}_\tau^{5,1}/S)\otimes\mathbb{Q}\cong(\mathrm{Cl}(\mathscr{Y}^{5,1}/S)\oplus(\mathrm{Cl}(\mathscr{X}_\tau^{5,1})/\,\mathrm{Cl}(\mathscr{Y}^{5,1})))\otimes\mathbb{Q}\cong R_1\oplus(\mathrm{ExCl}(X_t)\otimes\mathbb{Q})|_{\mathfrak{A}_5}, \quad (5.23)$$

with the summand R_1 on the right generated by the canonical class. Next we use the computation of Corollary 4.9 to describe the left-hand side of (5.23). Namely, by Corollary 4.9 the left-hand side is isomorphic to $R_1\oplus\mathrm{Ind}_G^{\mathfrak{A}_5}(\upsilon)$ for a certain subgroup $G\subset\mathfrak{A}_5$ and its one-dimensional representation υ. Canceling the R_1 summands, we obtain an isomorphism

$$(\mathrm{ExCl}(X_t)\otimes\mathbb{Q})|_{\mathfrak{A}_5}\cong\mathrm{Ind}_G^{\mathfrak{A}_5}(\upsilon).$$

It only remains to use the description of the subgroup G and its representation υ also provided by Corollary 4.9.

In the case $t=\frac{1}{2}$, so that $\tau=1$, it gives

$$(\mathrm{ExCl}(X_{1/2})\otimes\mathbb{Q})|_{\mathfrak{A}_5}\cong\mathrm{Ind}_{\mathfrak{A}_{3,2}}^{\mathfrak{A}_5}(-\mathbf{1})\cong R_3'\oplus R_3''\oplus R_4.$$

Therefore, by Lemma 5.14 we deduce that $(\mathrm{ExCl}(X_{1/2})\otimes\mathbb{Q})|_{\mathfrak{S}_5}$ is isomorphic either to $\mathrm{R}(3,1^2)\oplus\mathrm{R}(4,1)$ or to $\mathrm{R}(3,1^2)\oplus\mathrm{R}(2,1^3)$, hence by Lemma 5.12 we have either $\mathrm{ExCl}(X_{1/2})\otimes\mathbb{Q}\cong\mathrm{R}(4,1^2)$ or $\mathrm{ExCl}(X_{1/2})\otimes\mathbb{Q}\cong\mathrm{R}(3,1^3)$. The first case is impossible by Lemma 5.19, because by Pieri's rule the restriction of the $\mathfrak{S}_6$-representation $\mathrm{R}(4,1^2)$ to a standard subgroup $\mathfrak{S}_4$ contains a trivial subrepresentation, hence the required result.

Similarly, in the case $t=\frac{1}{6}$, so that $\tau=1/\sqrt{-3}$, we have

$$(\mathrm{ExCl}(X_{1/6})\otimes\mathbb{Q})|_{\mathfrak{A}_5}\cong\mathrm{Ind}_{\mathfrak{A}_4}^{\mathfrak{A}_5}(\mathbf{1})\cong R_1\oplus R_4.$$

Therefore, by Lemma 5.14 we deduce that $(\mathrm{ExCl}(X_{1/6})\otimes\mathbb{Q})|_{\mathfrak{S}_5}$ is isomorphic to the sum of one of the representations $\mathrm{R}(5)$ and $\mathrm{R}(1^5)$, and one of the representations $\mathrm{R}(4,1)$ and $\mathrm{R}(2,1^3)$. On the other hand, $(\mathrm{ExCl}(X_{1/6})\otimes\mathbb{Q})|_{\mathfrak{S}_5}$ should contain $\mathrm{R}(5)$ by Lemma 5.21, so it follows that $(\mathrm{ExCl}(X_{1/6})\otimes\mathbb{Q})|_{\mathfrak{S}_5}$ is either $\mathrm{R}(5)\oplus\mathrm{R}(4,1)$, or $\mathrm{R}(5)\oplus\mathrm{R}(2,1^3)$. By Lemma 5.12 only the first of them can be obtained as a restriction of a representation of $\mathfrak{S}_6$ with respect to a nonstandard embedding of $\mathfrak{S}_5$, and the corresponding representation of $\mathfrak{S}_6$ is $\mathrm{R}(2,2,2)$. Thus, we have $\mathrm{ExCl}(X_{1/6})\otimes\mathbb{Q}\cong\mathrm{R}(2,2,2)$.

Finally, in the case $t=\frac{7}{10}$, so that $\tau=3/\sqrt{5}$, we have

$$(\mathrm{ExCl}(X_{7/10})\otimes\mathbb{Q})|_{\mathfrak{A}_5}\cong R_1,$$

hence $\mathrm{ExCl}(X_{7/10})\otimes\mathbb{Q}$ is either $\mathrm{R}(6)$ or $\mathrm{R}(1^6)$. Again, the first case is impossible by Lemma 5.19, hence the required result. □

Now we are ready to prove the main result of this section.

Proof of Theorem 5.1. The description of $\mathrm{Cl}(\mathscr{Y}) \otimes \mathbb{Q}$ is given by Proposition 5.13, and the descriptions of $\mathrm{Cl}(\mathscr{X}_\infty) \otimes \mathbb{Q}$ and $\mathrm{Cl}(X_t) \otimes \mathbb{Q}$ for $t \notin \mathfrak{D} \cup \{\infty\}$ follow from a combination of Proposition 5.13 with Lemma 5.16. The last three isomorphisms follow from Proposition 5.22 in view of the definition of the excess class group. □

Remark 5.24. To study G-equivariant birational maps of the remaining $\mathfrak{S}_6$-invariant quartic $X_{1/4}$ to G-Mori fiber spaces, one can replace $X_{1/4}$ by its projective dual, which is the Segre cubic Z. This may be simpler because Z has terminal singularities. The corresponding problem for Z was partially solved in [Avilov 2016, Theorem 1.3]. In particular, if G is a standard subgroup $\mathfrak{A}_5$ in $\mathfrak{S}_6$, then $\mathrm{rk}\,\mathrm{Cl}(Z)^G = 1$ by [Avilov 2016, Proposition 3.1], and we expect that Z, and thus also $X_{1/4}$, is not G-rational. In this case the induced action of G on Z is also given by a standard embedding $\mathfrak{A}_5 \cong G \hookrightarrow \mathrm{Aut}(Z) \cong \mathfrak{S}_6$; see e.g., [Howard et al. 2008, Section 2.2]. On the contrary, if G is a nonstandard subgroup $\mathfrak{A}_5$ in $\mathfrak{S}_6$, then Z is known to be G-rational; see [Prokhorov 2010, 3.16].

Remark 5.25. One of the geometric interpretations of the nontrivial summands of $\mathrm{Cl}(X_t) \otimes \mathbb{Q}$ that appear in Theorem 5.1 is as follows. Suppose that $t \neq \frac{1}{4}, \infty$, so that the singularities of X_t are nodes by Theorem 3.3. Let $\nu\colon \tilde{X}_t \to X_t$ be the blow up of all singular points of X_t, and let $D_1, \ldots, D_r$ be the exceptional divisors of ν. Then $\tilde{X}_t$ is smooth, and $D_i \cong \mathbb{P}^1 \times \mathbb{P}^1$. Let M_i^+ and M_i^- be the rulings from two different families on D_i. One can check that there is a natural perfect pairing between the vector subspace in $H^4(\tilde{X}_t, \mathbb{C})$ spanned by the one-cycles $M_i^+ - M_i^-$ and the space $(\mathrm{Cl}(X_t)/\operatorname{Pic}(X_t)) \otimes \mathbb{C}$. Note also that the structure of this subspace of $H^4(\tilde{X}_t, \mathbb{C})$ as an $\mathfrak{S}_6$-representation can be independently deduced from [Schoen 1985, Proposition 1.3] and [Beauville 2013, Lemma 1].

Appendix: Cremona–Richmond configuration

The *Cremona–Richmond configuration* is the configuration CR of 15 lines with 15 triple intersection points in $\mathbb{P}^4$ formed by the singular locus of the Igusa quartic. By a small abuse of terminology, we will sometimes say that the singular locus is the configuration CR itself. We refer the reader to [Cremona 1877; Richmond 1900; Dolgachev 2004, Section 9] for basic properties.

Explicitly, the configuration CR can be described as follows. Consider $\mathbb{P}^4$ as the hyperplane given by (1.1) in $\mathbb{P}^5$ with the usual $\mathfrak{S}_6$-action. For each pairs-splitting

$$\{1, \ldots, 6\} = I_1 \sqcup I_2 \sqcup I_3,$$

where $|I_1| = |I_2| = |I_3| = 2$, let $\mathrm{L}_{(I_1|I_2|I_3)}$ be the line in $\mathbb{P}^4$ given by equations

$$x_i = x_j \text{ if } \{i, j\} = I_p \text{ for some } p \in \{1, 2, 3\}.$$

This gives 15 lines in $\mathbb{P}^4$; for instance, $\mathrm{L}_{(1,2|3,4|5,6)}$ is the line given by equations

$$x_1 = x_2, \quad x_3 = x_4, \quad x_5 = x_6, \tag{A.1}$$

and the other lines are obtained from this by the $\mathfrak{S}_6$-action.

Similarly, for every two-element subset $I \subset \{1, \dots, 6\}$ let P_I be the point in $\mathbb{P}^4$ given by equations

$$x_i = x_j \text{ if either } i, j \in I \text{ or } i, j \in \bar{I},$$

where $\bar{I}$ is the complement of I in $\{1, \dots, 6\}$. This gives 15 points in $\mathbb{P}^4$; for instance,

$$\mathrm{P}_{1,2} = (2 : 2 : -1 : -1 : -1 : -1), \tag{A.2}$$

and the other points are obtained from this by the $\mathfrak{S}_6$-action (so, this is the set Υ_{15} defined in Section 3.1).

It is easy to see that P_I lies on $\mathrm{L}_{(I_1|I_2|I_3)}$ if and only if $I = I_p$ for some $p \in \{1, 2, 3\}$, i.e., if I is one of the pairs in the pairs-splitting, or, equivalently, the pairs-splitting extends the pair I. In particular, there are three lines through each of the points (corresponding to three pairs-splittings of $\bar{I}$), and there are three points on each line (corresponding to three pairs in a pairs-splitting). Moreover, the points P_I are the only intersection points of the lines $\mathrm{L}_{(I_1|I_2|I_3)}$. Because of this CR is often referred to as a (15_3)-*configuration*.

In this section we discuss some properties of CR. In particular, in Theorem A.8 we show that CR is determined uniquely up to a projective transformation of $\mathbb{P}^4$ by its combinatorial structure (under a mild nondegeneracy assumption), and that the Igusa quartic is the only quartic whose singular locus contains CR.

We start by a discussion of combinatorics of CR.

Lemma A.3. *The configuration* CR *is combinatorially self-dual*: *an outer automorphism of* $\mathfrak{S}_6$ *induces a bijection between the set of points* P_I *and the set of lines* $\mathrm{L}_{(I_1|I_2|I_3)}$ *that preserves the incidence correspondence.*

Proof. There is a natural bijection between subsets of cardinality two in the set $\{1, \dots, 6\}$, and transpositions in the group $\mathfrak{S}_6$. Similarly, there is a natural bijection between pairs-splittings of the set $\{1, \dots, 6\}$, and elements of cycle type $[2, 2, 2]$ in $\mathfrak{S}_6$. Let us denote the transposition corresponding to a subset $I \subset \{1, \dots, 6\}$ by $w(I)$, and the element of cycle type $[2, 2, 2]$ corresponding to a pairs-splitting (I_1, I_2, I_3) of $\{1, \dots, 6\}$ by $w(I_1, I_2, I_3)$. The incidence relation of lines and points of CR can be reformulated in group-theoretic terms: the line $\mathrm{L}_{(I_1|I_2|I_3)}$ is incident to the point P_I if and only if the permutations $w(I)$ and $w(I_1, I_2, I_3)$ commute (or, which is the same, the composition $w(I) \circ w(I_1, I_2, I_3)$ has cycle type $[2, 2]$).

Choose an outer automorphism α of the group $\mathfrak{S}_6$. The automorphism α interchanges transpositions with elements of cycle type $[2, 2, 2]$. Thus α defines a map from the set of points of CR to the set of lines of CR, and a map from the set of lines of CR to the set of points of CR. Moreover, this map preserves the incidence relation. $\square$

Lemma A.3 implies the following result that we used in the main part of the paper.

Corollary A.4. *Every standard subgroup* $\mathfrak{A}_5 \subset \mathfrak{S}_6$ *acts transitively on the set of lines of* CR, *and every nonstandard subgroup* $\mathfrak{A}_5 \subset \mathfrak{S}_6$ *acts transitively on the set of points of* CR.

Proof. The first assertion is evident from combinatorics, and the second assertion follows from the first one in view of the bijection of Lemma A.3. $\square$

The following description of CR is very useful. Choose a triples-splitting

$$\{1, \dots, 6\} = K_0 \sqcup K_1, \quad |K_0| = |K_1| = 3.$$

For each bijection $g\colon K_0 \xrightarrow{\sim} K_1$ let $\Gamma(g)$ be the pairs-splitting formed by all pairs $\{k_0, g(k_0)\}$, where k_0 runs through K_0 (and hence $g(k_0)$ runs through K_1). The 6 lines and 9 points

$$\{\mathrm{L}_{\Gamma(g)}\}_{g\in \mathrm{Iso}(K_0,K_1)} \quad \text{and} \quad \{\mathrm{P}_{k_0,k_1}\}_{(k_0,k_1)\in K_0\times K_1}$$

form a subconfiguration $\mathrm{CR}'_{K_0,K_1} \subset \mathrm{CR}$ of the Cremona–Richmond configuration; see Figure 3. Because of its characteristic shape we call it a *jail configuration*. Note that CR'_{K_0,K_1} is contained in the hyperplane

$$H_{K_0} := \left\{\sum_{k\in K_0} x_k = 0\right\} = \left\{\sum_{k\in K_1} x_k = 0\right\} =: H_{K_1}. \tag{A.5}$$

We call it the *jail hyperplane*.

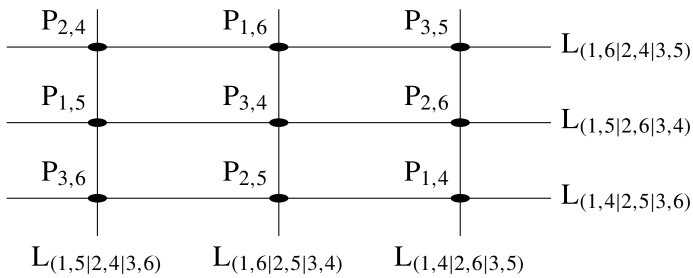


Figure 3. The jail subconfiguration $\mathrm{CR}'_{\{1,2,3\},\{4,5,6\}}$ in the Cremona–Richmond configuration CR.

The remaining 9 lines and 6 points

$$\{\mathrm{L}_{(k_0,k_1|K_0\setminus k_0|K_1\setminus k_1)}\}_{(k_0,k_1)\in K_0\times K_1} \quad \text{and} \quad \{\mathrm{P}_I\}_{I\subset K_0 \text{ or } I\subset K_1}$$

form a complete bipartite graph; see Figure 4; we call it a *bipartite configuration*.

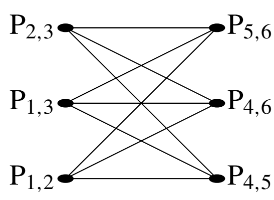


Figure 4. The bipartite subconfiguration $\mathrm{CR}''_{\{1,2,3\},\{4,5,6\}}$ in the Cremona–Richmond configuration CR.

For any decomposition

$$\mathrm{CR} = \mathrm{CR}'_{K_0,K_1} \cup \mathrm{CR}''_{K_0,K_1}$$

into a jail and bipartite subconfiguration its components interact quite weakly: every line $\mathrm{L}_{(k_0,k_1|K_0\setminus k_0|K_1\setminus k_1)}$ from the bipartite component passes through a single point P_{k_0,k_1} in the jail component. This gives a bijection between bipartite lines and jail points (compatible with the natural bijection of both sets with $K_0 \times K_1$).

Lemma A.6. *Let* C *be a configuration of* 15 *lines with* 15 *intersection points in* $\mathbb{P}^4$ *which is not contained in* $\mathbb{P}^3$ *and is combinatorially isomorphic to the Cremona–Richmond configuration. If* $\mathrm{C} = \mathrm{C}' \cup \mathrm{C}''$ *is a jail–bipartite decomposition then the jail component* C' *spans a hyperplane, and the bipartite component* C'' *spans* $\mathbb{P}^4$.

Proof. The jail component C' has the shape shown in Figure 3. Two vertical lines do not intersect, hence they span a hyperplane $H' \subset \mathbb{P}^4$. Three horizontal lines intersect each of them, hence they are contained in H'. The last vertical line intersects the horizontal lines, hence it is also contained in H'.

The bipartite component C'' has the shape shown in Figure 4. Assume it is contained in a hyperplane $H'' \subset \mathbb{P}^4$. Then every line of the bipartite component is contained in H''. Since every point of the jail component lies on a line of the bipartite component, it follows that the jail component is also contained in H''. Thus $\mathrm{C} \subset H''$, which contradicts the assumptions of the lemma. ☐

Remark A.7. The set $\{1, 2, 3, 4, 5, 6\}$ has 10 distinct triples-splittings, giving rise to 10 distinct jail-bipartite decompositions of the Cremona–Richmond configuration. The 10 hyperplanes supporting the jail components of CR appeared in Remark 2.25.

Theorem A.8. *Let* C *be a configuration of* 15 *lines with* 15 *intersection points in* $\mathbb{P}^4$ *which is not contained in* $\mathbb{P}^3$ *and is combinatorially isomorphic to the Cremona–Richmond configuration. Then it is projectively isomorphic to the Cremona–Richmond configuration.*

Proof. Choose a jail-bipartite decomposition $\mathrm{C} = \mathrm{C}' \cup \mathrm{C}''$. Choose five points $P_1, \dots, P_5$ in the bipartite component C'' that are not contained in a hyperplane (this is possible by Lemma A.6), and let H' be the hyperplane containing the jail component C'. Note that $P_i \notin H'$ for all i. Indeed, if $P_i \in H'$ then every line of the bipartite component passing through P_i would be contained in H' (since it also contains a point of the jail component), hence the three points of C'' that are connected to P_i by lines in C'' will be also contained in H'. Applying the same argument to one of these points, we would deduce that the whole bipartite component is contained in H', hence $\mathrm{C} \subset H'$, which contradicts our assumptions.

Assume that the points P_1, P_3, and P_5 are not connected to each other by lines in C''; that is, they are contained in one part of the bipartite component, and P_2, P_4 are contained in the other. Since the points P_i do not lie on a hyperplane, they can be taken to points

$$\begin{aligned} &P_1 = (1:0:0:0:0), \quad P_3 = (0:0:1:0:0), \quad P_5 = (0:0:0:0:1), \\ &P_2 = (0:1:0:0:0), \quad P_4 = (0:0:0:1:0), \end{aligned} \tag{A.9}$$

of $\mathbb{P}^4$ by a projective transformation. Since the hyperplane H' does not pass through the points P_i, it can be simultaneously taken to the hyperplane defined by the equation

$$x_1 - x_2 + x_3 - x_4 + x_5 = 0.$$

Now for each odd i and even j consider the line passing through P_i and P_j. By assumption it belongs to the bipartite component C''. The intersection points of these lines with H' are the following six points

$$\begin{aligned} &P_{12} = (1:1:0:0:0), \quad P_{32} = (0:1:1:0:0), \quad P_{52} = (0:1:0:0:1), \\ &P_{14} = (1:0:0:1:0), \quad P_{34} = (0:0:1:1:0), \quad P_{54} = (0:0:0:1:1). \end{aligned} \tag{A.10}$$

It follows that P_{ij} are points of the jail component C'. Consequently, the following six lines belong to the jail component C':

$$\begin{aligned} &\langle P_{12}, P_{34}\rangle = \{x_1 - x_2 = x_3 - x_4 = x_5 = 0\}, \quad \langle P_{12}, P_{54}\rangle = \{x_1 - x_2 = x_5 - x_4 = x_3 = 0\}, \\ &\langle P_{32}, P_{14}\rangle = \{x_3 - x_2 = x_1 - x_4 = x_5 = 0\}, \quad \langle P_{32}, P_{54}\rangle = \{x_3 - x_2 = x_5 - x_4 = x_1 = 0\}, \\ &\langle P_{52}, P_{14}\rangle = \{x_5 - x_2 = x_1 - x_4 = x_3 = 0\}, \quad \langle P_{52}, P_{34}\rangle = \{x_5 - x_2 = x_3 - x_4 = x_1 = 0\}, \end{aligned}$$

and their three extra intersection points

$$P_{1234} = (1:1:1:1:0), \quad P_{1245} = (1:1:0:1:1), \quad P_{2345} = (0:1:1:1:1) \tag{A.11}$$

also belong to C'. Finally, the last point P_0 of the bipartite component is the point

$$P_0 = \langle P_1, P_{2345}\rangle \cap \langle P_3, P_{1245}\rangle \cap \langle P_5, P_{1234}\rangle = (1:1:1:1:1). \tag{A.12}$$

This proves that such configuration is unique up to a projective transformation. The explicit transformation from $\mathbb{P}^4$ to $\mathbb{P}^5$ that takes the points (A.9), (A.10), (A.11), and (A.12) to the points $\mathrm{P}_{i,j}$ that were defined in (A.2) is given by the matrix

$$\begin{pmatrix} 1 & 1 & -2 & 1 & -2 \\ -2 & 1 & 1 & 1 & -2 \\ -2 & 1 & -2 & 1 & 1 \\ 1 & -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & -2 & 1 \\ 1 & 1 & 1 & -2 & 1 \end{pmatrix};$$

in particular, the point P_5 is mapped to the point $\mathrm{P}_{1,2}$ in (A.2). This completes the proof of Theorem A.8. □

Remark A.13. Let C be a configuration combinatorially isomorphic to CR. Then one can always project C isomorphically to $\mathbb{P}^3$. In particular, the assumption of Theorem A.8 requiring that the configuration is not contained in $\mathbb{P}^3$ is necessary.

Corollary A.14. *Let* C *be a configuration of* 15 *lines with* 15 *intersection points in* $\mathbb{P}^4$ *which is not contained in* $\mathbb{P}^3$ *and is combinatorially isomorphic to the Cremona–Richmond configuration. Suppose that* X *is a quartic threefold that contains* C *in its singular locus. Then it is projectively isomorphic to the Igusa quartic.*

Proof. By Theorem A.8 it is enough to show that the Igusa quartic X is the unique quartic singular along C. Suppose that X' is another quartic with this property. Since X is irreducible, the intersection $Z = X \cap X'$ is two-dimensional, and $\deg Z = 16$. Let C' be one of the jail subconfigurations of C. Then C' is contained in a unique two-dimensional smooth quadric T; this quadric is swept out by lines that meet three of the lines in C'. The lines of C' are singular both on X and X', so we conclude that T is contained in Z. It remains to notice that C contains 10 jail subconfigurations, all of them giving rise to different two-dimensional quadrics contained in Z. The degree of the union of these quadrics is 20; this is greater than $\deg Z$, which gives a contradiction. □

Acknowledgements

We are grateful to A. Beauville, S. Bloch, I. Dolgachev, G. Kapustka, D. Markushevich, Yu. Prokhorov, and E. Tevelev, for useful discussions. We also thank the referee for valuable remarks. This paper was written during the Cheltov's stay at the Max Planck Institute for Mathematics in 2017. He would like to thank the institute for the excellent working conditions. All authors were partially supported by the HSE University Basic Research Program, Russian Academic Excellence Project "5–100". Kuznetsov and Shramov were also supported by RFBR grants 15-01-02164 and 15-01-02158. Shramov was also supported by Young Russian Mathematics award.

References

[Avilov 2016] A. A. Avilov, "Automorphisms of three-dimensional singular cubic hypersurfaces and the Cremona group", *Mat. Zametki* **100**:3 (2016), 461–464. In Russian; translated in *Math. Notes* **100**:3-4 (2016), 482–485. MR Zbl

[Bauer and Verra 2010] I. Bauer and A. Verra, "The rationality of the moduli space of genus-4 curves endowed with an order-3 subgroup of their Jacobian", *Michigan Math. J.* **59**:3 (2010), 483–504. MR Zbl

[Beauville 1977] A. Beauville, "Variétés de Prym et jacobiennes intermédiaires", *Ann. Sci. École Norm. Sup.* (4) **10**:3 (1977), 309–391. MR Zbl

[Beauville 1982] A. Beauville, "Les singularités du diviseur Θ de la jacobienne intermédiaire de l'hypersurface cubique dans $\mathbb{P}^4$", pp. 190–208 in *Algebraic threefolds* (Varenna, Italy, 1981), edited by A. Conte, Lecture Notes in Math. **947**, Springer, 1982. MR Zbl

[Beauville 2013] A. Beauville, "Non-rationality of the $\mathfrak{S}_6$-symmetric quartic threefolds", *Rend. Semin. Mat. Univ. Politec. Torino* **71**:3-4 (2013), 385–388. MR Zbl

[Bernstein and Schwarzman 2006] J. Bernstein and O. Schwarzman, "Complex crystallographic Coxeter groups and affine root systems", *J. Nonlinear Math. Phys.* **13**:2 (2006), 163–182. MR Zbl

[Cheltsov 2005] I. A. Cheltsov, "Birationally rigid Fano varieties", *Uspekhi Mat. Nauk* **60**:5 (2005), 71–160. In Russian; translated in *Russ. Math. Surv.* **60**:5 (2005), 875–965. MR Zbl

[Cheltsov and Shramov 2014] I. Cheltsov and C. Shramov, "Five embeddings of one simple group", *Trans. Amer. Math. Soc.* **366**:3 (2014), 1289–1331. MR Zbl

[Cheltsov and Shramov 2016a] I. Cheltsov and C. Shramov, *Cremona groups and the icosahedron*, CRC Press, Boca Raton, FL, 2016. MR Zbl

[Cheltsov and Shramov 2016b] I. Cheltsov and C. Shramov, "Two rational nodal quartic 3-folds", *Q. J. Math.* **67**:4 (2016), 573–601. MR Zbl

[Cheltsov et al. 2019a] I. Cheltsov, V. Przyjalkowski, and C. Shramov, "Burkhardt quartic, Barth sextic, and the icosahedron", *Int. Math. Res. Not.* **2019**:12 (2019), 3683–3703. MR

[Cheltsov et al. 2019b] I. Cheltsov, V. Przyjalkowski, and C. Shramov, "Which quartic double solids are rational?", *J. Algebraic Geom.* **28**:2 (2019), 201–243. MR Zbl

[Clingher et al. 2019] A. Clingher, A. Malmendier, and T. Shaska, “Six line configurations and string dualities”, *Comm. Math. Phys.* **371**:1 (2019), 159–196. MR Zbl

[Coble 1906] A. B. Coble, “An invariant condition for certain automorphic algebraic forms”, *Amer. J. Math.* **28**:4 (1906), 333–366. MR Zbl

[Coble 1915] A. B. Coble, “Point sets and allied Cremona groups, I”, *Trans. Amer. Math. Soc.* **16**:2 (1915), 155–198. MR Zbl

[Coble 1916] A. B. Coble, “Point sets and allied Cremona groups, II”, *Trans. Amer. Math. Soc.* **17**:3 (1916), 345–385. MR Zbl

[Coble 1917] A. B. Coble, “Point sets and allied Cremona groups, III”, *Trans. Amer. Math. Soc.* **18**:3 (1917), 331–372. MR Zbl

[Conway et al. 1985] J. H. Conway, R. T. Curtis, S. P. Norton, R. A. Parker, and R. A. Wilson, *Atlas of finite groups: maximal subgroups and ordinary characters for simple groups*, Oxford Univ. Press, 1985. MR Zbl

[Corti 1995] A. Corti, “Factoring birational maps of threefolds after Sarkisov”, *J. Algebraic Geom.* **4**:2 (1995), 223–254. MR Zbl

[Cremona 1877] L. Cremona, “Teoremi stereometrici dai quali si deducono le proprietà dell’esagrammo di Pascal”, *Atti R. Accad. Lincei Mem. Cl. Sci. Fis. Mat. Natur.* (3) **1** (1877), 854–874.

[Cynk 2001] S. Cynk, “Defect of a nodal hypersurface”, *Manuscripta Math.* **104**:3 (2001), 325–331. MR Zbl

[Debarre and Kuznetsov 2018] O. Debarre and A. Kuznetsov, “Gushel–Mukai varieties: classification and birationalities”, *Algebr. Geom.* **5**:1 (2018), 15–76. MR Zbl

[Debarre and Kuznetsov 2020] O. Debarre and A. Kuznetsov, “Gushel–Mukai varieties: moduli”, *International Journal of Mathematics* (2020).

[Dolgachev 1982] I. Dolgachev, “Weighted projective varieties”, pp. 34–71 in *Group actions and vector fields* (Vancouver, 1981), edited by J. B. Carrell, Lecture Notes in Math. **956**, Springer, 1982. MR Zbl

[Dolgachev 2004] I. V. Dolgachev, “Abstract configurations in algebraic geometry”, pp. 423–462 in *The Fano Conference* (Turin, 2002), edited by A. Collino et al., Univ. Torino, 2004. MR Zbl

[Dolgachev 2012] I. V. Dolgachev, *Classical algebraic geometry: a modern view*, Cambridge Univ. Press, 2012. MR Zbl

[Dolgachev 2016] I. Dolgachev, “Corrado Segre and nodal cubic threefolds”, pp. 429–450 in *From classical to modern algebraic geometry*, edited by G. Casnati et al., Birkhäuser, Cham, 2016. MR Zbl

[Dolgachev and Ortland 1988] I. Dolgachev and D. Ortland, *Point sets in projective spaces and theta functions*, Astérisque **165**, 1988. MR Zbl

[Dolgachev et al. 2005] I. Dolgachev, B. van Geemen, and S. Kondō, “A complex ball uniformization of the moduli space of cubic surfaces via periods of $K3$ surfaces”, *J. Reine Angew. Math.* **588** (2005), 99–148. MR Zbl

[Dolgachev et al. 2018] I. Dolgachev, B. Farb, and E. Looijenga, “Geometry of the Wiman–Edge pencil, I: Algebro-geometric aspects”, *Eur. J. Math.* **4**:3 (2018), 879–930. MR Zbl

[Edge 1981] W. L. Edge, “A pencil of four-nodal plane sextics”, *Math. Proc. Cambridge Philos. Soc.* **89**:3 (1981), 413–421. MR Zbl

[Farkas and Verra 2016] G. Farkas and A. Verra, “The universal abelian variety over $\mathscr{A}_5$”, *Ann. Sci. École Norm. Sup.* (4) **49**:3 (2016), 521–542. MR Zbl

[Feit 1971] W. Feit, “The current situation in the theory of finite simple groups”, pp. 55–93 in *Actes du Congrès International des Mathématiciens, I* (Nice, France, 1970), edited by P. Montel, Gauthier-Villars, Paris, 1971. MR Zbl

[Finkelnberg 1987] H. Finkelnberg, “Small resolutions of the Segre cubic”, *Nederl. Akad. Wetensch. Indag. Math.* **49**:3 (1987), 261–277. MR Zbl

[Fulton 1984] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik (3) **2**, Springer, 1984. MR Zbl

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory: a first course*, Grad. Texts in Math. **129**, Springer, 1991. MR Zbl

[GAP 2017] The GAP Group, “GAP: groups, algorithms, and programming”, 2017, available at https://www.gap-system.org. Version 4.8.8.

[van der Geer 1982] G. van der Geer, “On the geometry of a Siegel modular threefold”, *Math. Ann.* **260**:3 (1982), 317–350. MR Zbl

[Hacking et al. 2009] P. Hacking, S. Keel, and J. Tevelev, "Stable pair, tropical, and log canonical compactifications of moduli spaces of del Pezzo surfaces", *Invent. Math.* **178**:1 (2009), 173–227. MR Zbl

[Hacon and McKernan 2013] C. D. Hacon and J. McKernan, "The Sarkisov program", *J. Algebraic Geom.* **22**:2 (2013), 389–405. MR Zbl

[Howard et al. 2008] B. Howard, J. Millson, A. Snowden, and R. Vakil, "A description of the outer automorphism of S_6, and the invariants of six points in projective space", *J. Combin. Theory Ser. A* **115**:7 (2008), 1296–1303. MR Zbl

[Hunt 1996] B. Hunt, *The geometry of some special arithmetic quotients*, Lecture Notes in Math. **1637**, Springer, 1996. MR Zbl

[Inoue and Kato 2005] N. Inoue and F. Kato, "On the geometry of Wiman's sextic", *J. Math. Kyoto Univ.* **45**:4 (2005), 743–757. MR Zbl

[James and Liebeck 1993] G. James and M. Liebeck, *Representations and characters of groups*, Cambridge Univ. Press, 1993. MR Zbl

[de Jong et al. 1990] A. J. de Jong, N. I. Shepherd-Barron, and A. Van de Ven, "On the Burkhardt quartic", *Math. Ann.* **286**:1-3 (1990), 309–328. MR Zbl

[Kaloghiros 2011] A.-S. Kaloghiros, "The defect of Fano 3-folds", *J. Algebraic Geom.* **20**:1 (2011), 127–149. MR Zbl

[Kuznetsov 2016] A. Kuznetsov, "Küchle fivefolds of type $c5$", *Math. Z.* **284**:3-4 (2016), 1245–1278. MR Zbl

[Looijenga 2009] E. Looijenga, "The period map for cubic fourfolds", *Invent. Math.* **177**:1 (2009), 213–233. MR Zbl

[Matsumoto et al. 1992] K. Matsumoto, T. Sasaki, and M. Yoshida, "The monodromy of the period map of a 4-parameter family of $K3$ surfaces and the hypergeometric function of type (3, 6)", *Int. J. Math.* **3**:1 (1992), 164. MR Zbl

[Mella 2004] M. Mella, "Birational geometry of quartic 3-folds, II: The importance of being $\mathbb{Q}$-factorial", *Math. Ann.* **330**:1 (2004), 107–126. MR Zbl

[Naruki 1982] I. Naruki, "Cross ratio variety as a moduli space of cubic surfaces", *Proc. Lond. Math. Soc.* (3) **45**:1 (1982), 1–30. MR Zbl

[Prokhorov 2010] Y. G. Prokhorov, "Fields of invariants of finite linear groups", pp. 245–273 in *Cohomological and geometric approaches to rationality problems*, edited by F. Bogomolov and Y. Tschinkel, Progr. Math. **282**, Birkhäuser, Boston, 2010. MR Zbl

[Prokhorov 2013] Y. Prokhorov, "G-Fano threefolds, I", *Adv. Geom.* **13**:3 (2013), 389–418. MR Zbl

[Przyjalkowski and Shramov 2016] V. V. Przyjalkowski and C. A. Shramov, "Double quadrics with large automorphism groups", *Tr. Mat. Inst. Steklova* **294** (2016), 167–190. In Russian; translated in *Proc. Steklov Inst. Math.* **294**:1 (2016), 154–175. MR Zbl

[Richmond 1900] H. W. Richmond, "The figure formed from six points in space of four dimensions", *Math. Ann.* **53**:1-2 (1900), 161–176. MR Zbl

[Sarkisov 1982] V. G. Sarkisov, "On conic bundle structures", *Izv. Akad. Nauk SSSR Ser. Mat.* **46**:2 (1982), 371–408. In Russian; translated in *Math. USSR-Izv.* **20**:2 (1983), 355–390. MR Zbl

[Schoen 1985] C. Schoen, "Algebraic cycles on certain desingularized nodal hypersurfaces", *Math. Ann.* **270**:1 (1985), 17–27. MR Zbl

[Shepherd-Barron 1989] N. I. Shepherd-Barron, "Invariant theory for S_5 and the rationality of M_6", *Compositio Math.* **70**:1 (1989), 13–25. MR Zbl

[Shokurov 1983] V. V. Shokurov, "Prym varieties: theory and applications", *Izv. Akad. Nauk SSSR Ser. Mat.* **47**:4 (1983), 785–855. In Russian; translated in *Math. USSR-Izv.* **23**:1 (1984), 83–147. MR Zbl

[Shramov 2008] K. A. Shramov, "On the birational rigidity and $\mathbb{Q}$-factoriality of a singular double covering of a quadric with branching over a divisor of degree 4", *Mat. Zametki* **84**:2 (2008), 300–311. In Russian; translated in *Math. Notes* **84**:1-2 (2008), 280–289. MR Zbl

[Todd 1933] J. A. Todd, "Configurations defined by six lines in space of three dimensions", *Proc. Cambridge Philos. Soc.* **29**:1 (1933), 52–68. Zbl

[Todd 1935] J. A. Todd, "A note on two special primals in four dimensions", *Q. J. Math.* **6**:1 (1935), 129–136. Zbl

[Todd 1936] J. A. Todd, "On a quartic primal with forty-five nodes, in space of four dimensions", *Q. J. Math.* **7**:1 (1936), 168–174. Zbl

[Wiman 1896a] A. Wiman, “Ueber eine einfache Gruppe von 360 ebenen Collineationen”, *Math. Ann.* **47**:4 (1896), 531–556. MR Zbl

[Wiman 1896b] A. Wiman, “Zur Theorie der endlichen Gruppen von birationalen Transformationen in der Ebene”, *Math. Ann.* **48**:1-2 (1896), 195–240. MR Zbl

[Zamora 2018] A. G. Zamora, “Some remarks on the Wiman–Edge pencil”, *Proc. Edinb. Math. Soc.* (2) **61**:2 (2018), 401–412. MR Zbl

Communicated by Gavril Farkas
Received 2019-02-02 Revised 2019-07-01 Accepted 2019-09-01

i.cheltsov@ed.ac.uk *Department of Mathematics, University of Edinburgh, United Kingdom*

National Research University Higher School of Economics, Russian Federation

akuznet@mi-ras.ru *Algebraic Geometry Section, Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, Russia*

National Research University Higher School of Economics, Russian Federation

costya.shramov@gmail.com *Algebraic Geometry Section, Steklov Mathematical Institute of Russian Academy of Sciences, Moscow, Russia*

National Research University Higher School of Economics, Russian Federation

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 14 No. 1 2020

1937-0652(2020)14:1;1-0