Algebra & Number Theory

Volume 8
2014
No. 2

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2014 is US $225/year for the electronic version, and $400/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers 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 Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.257

Large self-injective rings and the generating hypothesis

Leigh Shepperson and Neil Strickland

We construct a number of different examples of non-Noetherian graded rings that are injective as modules over themselves (or have some related but weaker properties). We discuss how these are related to the theory of triangulated categories, and to Freyd's generating hypothesis in stable homotopy theory.

1. Introduction

In this paper we study graded commutative rings R that are large in various senses (in particular, not Noetherian) and self-injective (meaning that R is injective as an R-module). We use graded rings because they are relevant for our applications, but ungraded rings are covered as well because they can be regarded as graded rings concentrated in degree zero. The graded setting is assumed everywhere, so "element" means "homogeneous element" and "ideal" means "homogeneous ideal" and so on. Our rings will be commutative in the graded sense, so that $ba = (-1)^{|a||b|}ab$.

It is not hard to prove that any Noetherian self-injective ring is Artinian. In particular, if R is a finitely generated algebra over a field K that is self-injective then we must have $\dim_K(R) < \infty$ and it turns out that $R \simeq \mathrm{Hom}(R, K)$ as R-modules. Examples of this situation include $R = K[x_1, \dots, x_n]/(r_1, \dots, r_n)$ for any regular sequence $r_1, \dots, r_n$, or the cohomology ring $R = H^*(M; K)$ for any closed orientable manifold M. These are the most familiar examples of self-injective rings, and they are all very small. We will be looking for examples that are much larger.

Our motivation comes from a question in stable homotopy theory, which we briefly recall. In stable homotopy theory we study a certain triangulated category $\mathcal{F}$, the Spanier–Whitehead category of finite spectra. The objects can be taken to be pairs $X = (n, A)$, where $n \in \mathbb{Z}$ and A is a finite simplicial complex. The morphism set $\mathrm{Hom}_{\mathcal{F}}((n, A), (m, B))$ is the set of homotopy classes of maps from $(\mathbb{R}^{N+n} \times A) \cup \{\infty\}$ to $(\mathbb{R}^{N+m} \times B) \cup \{\infty\}$, which is essentially independent of N when N is sufficiently large. More details are given in [Ravenel 1992], for example. For any $X, Y \in \mathcal{F}$ the set $\mathrm{Hom}_{\mathcal{F}}(X, Y)$ is a finitely generated abelian

MSC2010: primary 13C11; secondary 18E30, 55P42.
Keywords: self-injective, generating hypothesis.

group. It turns out that most methods for studying $\mathrm{Hom}_{\mathscr{F}}(X, Y)$ treat the p-primary parts separately for different primes p. We will thus fix a prime p and define $[X, Y] = \mathbb{Z}_p \otimes \mathrm{Hom}_{\mathscr{F}}(X, Y)$, where $\mathbb{Z}_p$ is the ring of p-adic integers. These are the morphism sets in a new triangulated category which we call $\mathscr{F}_p$. This has a canonical tensor structure, with the tensor product of X and Y written as $X \wedge Y$. The unit for this structure is called S, so $S \wedge X \simeq X$. As part of the triangulated structure we have a suspension functor $\Sigma\colon \mathscr{F}_p \to \mathscr{F}_p$, and we write S^n for $\Sigma^n S$. We put $R_n = [S^n, S]$. These sets form a graded commutative ring, whose structure is extremely intricate. A great deal of partial information is known, but it seems clear that there will never be a usable complete description. Some highlights are as follows.

- $R_n = 0$ for $n < 0$, and $R_0 = \mathbb{Z}_p$, and R_n is a finite abelian p-group for $n > 0$.
- Both the ranks and the exponents of the groups R_n can be arbitrarily large.
- All elements in R_n with $n > 0$ are nilpotent. Thus, the reduced quotient is $R/\sqrt{0} = \mathbb{Z}_p$.
- Various results are available describing most or all of the structure of R_n for $n < f(p)$, where $f(x)$ is a polynomial of degree at most three. The simplest of these says that $R_n = 0$ for $0 < n < 2p - 3$, and $R_{2p-3} = \mathbb{Z}/p$.

Now consider an arbitrary object $X \in \mathscr{F}_p$. We define $\pi_n(X) = [S^n, X]$ for all $n \in \mathbb{Z}$. This defines a graded abelian group $\pi_*(X)$, which has a natural structure as an R-module.

Conjecture 1.1 (Freyd's generating hypothesis). *The functor* $\pi_*\colon \mathscr{F}_p \to \mathrm{Mod}_R$ *is faithful.*

This is actually a technical modification of Freyd's conjecture [1966], because Freyd did not tensor with the p-adics. This causes various troubles in the development of the theory, which Freyd avoided in ad hoc ways. Much later, Hovey [2007] redeveloped the theory in the p-adic setting, which involves only minor modifications to Freyd's arguments but works much more smoothly.

Nearly half a century after Freyd made his conjecture, there is still no hint of a proof or a counterexample. However, there has been a certain amount of indirect progress; for example, various authors have settled the analogous questions in other triangulated categories where computations are easier [Carlson et al. 2009; Hovey et al. 2007; Benson et al. 2007; Lockridge 2007].

On the other hand, it is known that the generating hypothesis would have some very strong and surprising consequences, as we now explain.

Definition 1.2. (a) A graded ring R is *coherent* if every finitely generated ideal is finitely presented.

(b) A graded ring R is *totally incoherent* if the only finitely presented ideals are 0 and R.

Theorem 1.3 [Freyd 1966; Hovey 2007]. *Suppose that the generating hypothesis is true.*

(a) *The functor* $\pi_*\colon \mathcal{F}_p \to \mathrm{Mod}_R$ *is automatically full as well as being faithful, so it is an embedding of categories.*

(b) *For every object* $X \in \mathcal{F}_p$, *the image* $\pi_*(X)$ *is an injective* R*-module. In particular* (*by taking* $X = S$) *the ring* R *is self-injective.*

(c) *The ring* R *is totally incoherent.*

Note in particular that (a) gives a full subcategory of Mod_R that has a natural triangulation. This is very unusual; in almost all known triangulated categories, the morphisms are equivalence classes of homomorphisms under some nontrivial equivalence relation, and this equivalence structure is tightly connected to the definition of the triangulation.

Our aim in this paper is to shed light on the generating hypothesis by finding examples of self-injective rings that share some of the known or conjectured properties of the stable homotopy ring R.

Our main results are as follows. Firstly, one cannot disprove self-injectivity by looking only in a finite range of degrees:

Theorem 1.4. *Let* R *be a graded-commutative ring such that*

(a) $R_k = 0$ *for* $k < 0$,

(b) $R_0 = \mathbb{Z}/2$,

(c) R_k *is finite for all* $k \geq 0$.

Suppose given $N > 0$. *Then there is an injective map* $\phi\colon R \to R'$ *of graded rings such that*

(1) R' *also has properties* (a)–(c),

(2) $\phi\colon R_k \to R'_k$ *is an isomorphism for* $k < N$,

(3) R' *is self-injective.*

This result was a great surprise to the authors at least, although the proof is not too hard. We will restate and prove it as Theorem 6.6. We conjecture that the theorem remains true if we allow R_0 to be $\mathbb{Z}_p$, but we have not proved this.

Most of our remaining results relate to specific examples. We have aimed to give a wide spread of examples, rather than formulating each example with maximum possible generality. We will write $\mathbb{F}$ for $\mathbb{Z}/2$.

One of the simplest examples of a finite-dimensional self-injective ring is the exterior algebra

$$\mathbb{F}[x_0, \dots, x_n]/(x_0^2, \dots, x_n^2).$$

Our first infinite-dimensional example is just an obvious generalisation of this.

Proposition 1.5. *Let E be the exterior algebra over $\mathbb{F}$ with a generator $x_i \in E_{2^i}$ for all $i \in \mathbb{N}$. Then E is self-injective and coherent. The reduced quotient is $E/\sqrt{0} = \mathbb{F}$.*

Self-injectivity is proved by combining Corollary 3.7 and Proposition 4.6, as will be explained in Example 4.7. The same ingredients cover many other examples, but we will not give the relevant definitions in this introduction. Coherence is proved in Proposition 5.4, and the reduced quotient is clear. We have chosen the degrees of the generators for compatibility with our other examples, but in fact the statement would remain valid if we merely assumed that $|x_i| \to \infty$ as $i \to \infty$.

Our next example arose by applying Theorem 1.4 to the ring $\mathbb{F}[x, y]/xy$ and studying the result in low dimensions. The result is very complicated and irregular, but after studying various recurring patterns and key features we were led to the definition below.

Theorem 1.6. *Consider the ring*

$$C = \mathbb{F}[y_0, y_1, \dots]/(y_i^3 + y_i y_{i+1} \mid i \geq 0),$$

with the grading given by $|y_i| = 2^i$. Then C is self-injective and coherent. The reduced quotient is

$$C/\sqrt{0} = \mathbb{F}[x_0, x_1, \dots]/(x_i x_j \mid i \neq j) = \mathbb{F} \oplus \bigoplus_{n>0} x_n \mathbb{F}[x_n],$$

where $x_n = \sum_{i=0}^{n} y_{n-i}^{2^i}$.

This will be proved as Propositions 7.18, 7.25 and 7.26. The statement can be generalised by adjusting the degrees and the relations slightly, but this just leads to additional bookkeeping without much extra insight, so we have omitted it. It is probably also possible to generalise in more conceptual ways, but that would be a substantial project, so we leave it for future work.

For the next example, we give an axiomatic statement and then explain a special case that is relevant in chromatic homotopy theory.

Definition 1.7. For any prime p, we recall that

$$\mathbb{Z}[1/p]/\mathbb{Z} \simeq \mathbb{Q}/\mathbb{Z}_{(p)} \simeq \mathbb{Q}_p/\mathbb{Z}_p \simeq \lim_{n\to\infty} \mathbb{Z}/p^n.$$

For any module M over $\mathbb{Z}_p$, we write $M^\vee = \mathrm{Hom}_{\mathbb{Z}_p}(M, \mathbb{Q}_p/\mathbb{Z}_p)$, and call this the *Pontrjagin dual* of M. One can check that ${\mathbb{Z}_p}^\vee \simeq \mathbb{Q}_p/\mathbb{Z}_p$ and $(\mathbb{Q}_p/\mathbb{Z}_p)^\vee \simeq \mathbb{Z}_p$ and $(\mathbb{Z}/p^n)^\vee \simeq \mathbb{Z}/p^n$. Now consider a graded $\mathbb{Z}_p$-algebra R with a specified isomorphism $\zeta\colon R_d \to \mathbb{Q}_p/\mathbb{Z}_p$ for some d. This gives maps $\zeta^\#\colon R_{d-k} \to R_k^\vee$ by $\zeta^\#(a)(b) = \zeta(ab)$. We say that R is *Pontrjagin self-dual* if all these maps are isomorphisms.

Proposition 1.8. *If R is Pontrjagin self-dual, then it is self-injective.*

This will be proved as Proposition 8.2.

Now fix a prime p, and assume that $p > 2$ for simplicity. Recall that $\mathscr{F}$ denotes the Spanier–Whitehead category of finite spectra. One can construct another triangulated category $\mathscr{F}'$, called the *Bousfield localisation of $\mathscr{F}$ with respect to p-local K-theory*. Roughly speaking this is the closest possible approximation to $\mathscr{F}$ that can be analysed using topological K-theory, and it is computationally much more tractable than $\mathscr{F}$ itself. Ravenel's paper [1984] is a good introduction to both the conceptual framework and specific calculations, with references to original sources. Devinatz [1990] has shown that the most obvious analogue of the generating hypothesis for $\mathscr{F}'$ is false (his Remark 1.7), but that a related statement is true (his Theorem 1). The analogue of the stable homotopy ring for $\mathscr{F}'$ is the ring J described below.

Definition 1.9. Let p be an odd prime, and define a graded ring J as follows. We put $J_0 = \mathbb{Z}_{(p)}$ and $J_{-2} = \mathbb{Q}_p/\mathbb{Z}_p$; for notational convenience we use the symbol η for the identity map $\mathbb{Z}_{(p)} \to J_0$, and ζ for the identity map $J_{-2} \to \mathbb{Q}_p/\mathbb{Z}_p$. Next, for each nonzero integer k there is a generator $\alpha_k \in J_{2(p-1)k-1}$ generating a cyclic group of order $p^{v_p(k)+1}$, where $v_p(k)$ is the p-adic valuation of k. For the product structure, we have:

- $\eta(a)\eta(b) = \eta(ab)$ and $\eta(a)\zeta^{-1}(b) = \zeta^{-1}(ab)$ and $\eta(a)\alpha_k = a\,\alpha_k$.
- $\zeta^{-1}(a)\zeta^{-1}(b) = 0$ and $\zeta^{-1}(a)\alpha_k = 0$ for all k.
- If $k > 0$ we have
$$\alpha_k\alpha_{-k} = -\alpha_{-k}\alpha_k = \zeta^{-1}(p^{-1-v_p(k)} + \mathbb{Z}_{(p)}).$$
- $\alpha_j\alpha_k = 0$ whenever $j + k \neq 0$.

Theorem 1.10. *The ring $\hat{J} = \mathbb{Z}_p \otimes J$ is Pontrjagin self-dual and therefore self-injective. It is also totally incoherent, and the reduced quotient is $\hat{J}/\sqrt{0} = \mathbb{Z}_p$.*

Self-duality is proved as Lemma 8.3, and incoherence as Proposition 8.7. The reduced quotient is clear.

Remark 1.11. Tensoring with $\mathbb{Z}_p$ here just has the effect of replacing $\mathbb{Z}_{(p)}$ in degree zero with $\mathbb{Z}_p$. Note that this is not the same as the p-completion of J, because $(\mathbb{Q}_p/\mathbb{Z}_p)_p = 0$. Moreover, a derived version of p-completion would replace $\mathbb{Q}_p/\mathbb{Z}_p$ by a copy of $\mathbb{Z}_p$ shifted by one degree, which is different again. The ring J itself is not self-injective. However, this does not account for Devinatz's example showing the failure of the generating hypothesis in $\mathcal{F}'$; that has a deeper topological origin.

We now note that the ring $\mathbb{F}[x]/x^N$ is another easy example of a finite-dimensional self-injective ring. Our next example arose by trying to generalise this. An obvious possibility is to consider the ring $\bigcup_{n>0} \mathbb{F}[x^{1/n}]$ modulo the ideal generated by x. Any element of this ring can be expressed as $\sum_q a(q)x^q$, for some function $a\colon \mathbb{Q} \cap [0, 1) \to \mathbb{F}$ with finite support. However, this ring needs to be adjusted to make it self-injective. Firstly, it turns out to be better not to kill x itself, but just the powers x^q with $q > 1$. Next, self-injectivity forces certain modules to be isomorphic to their double duals and thus to have strong completeness properties. To handle this, we must allow some infinite sums, or equivalently weaken the condition that a has finite support. It is also convenient (but not strictly necessary) to include powers x^q where q is irrational. This leads us to the following definition.

Definition 1.12. Let K be a field. For any map $a\colon [0, 1] \to K$ we put

$$\operatorname{supp}(a) = \{q \in [0, 1] \mid a(q) \neq 0\}.$$

We say that a is an *infinite root series* if every nonempty subset of $\operatorname{supp}(a)$ has a smallest element (so $\operatorname{supp}(a)$ is well-ordered). We let P denote the set of infinite root series, and call this the *infinite root algebra.*

Theorem 1.13. *The formula*

$$(ab)(q) = \sum_{0 \leq r \leq q} a(r)\, b(q - r)$$

gives a well-defined ring structure on P. With this structure, P is self-injective and totally incoherent. The reduced quotient is $P/\sqrt{0} = K$.

This will be proved in Propositions 9.20 and 9.21, and Corollary 9.13.

We will also discuss two rings that are not self-injective, but have a related property that we now explain.

Definition 1.14. Let R be a graded commutative ring, and let J be an ideal in R. We put $\operatorname{ann}_R(J) = \{a \in R \mid aJ = 0\}$. It is tautological that the ideal $\operatorname{ann}^2_R(J) = \operatorname{ann}_R(\operatorname{ann}_R(J))$ contains J. We say that R satisfies the *double annihilator condition* if $\operatorname{ann}^2_R(J) = J$ for all finitely generated ideals J.

Proposition 1.15. *If R is self-injective then it satisfies the double annihilator condition. Conversely, if R is Noetherian and satisfies the double annihilator condition, then it is self-injective.*

This is proved in Remark 2.4 and Theorem 4.1.

Definition 1.16. For any integer n we let $B(n)$ be the set of exponents i such that 2^i occurs in the binary expansion of n, so $B(n)$ is the unique finite subset of $\mathbb{N}$ such that $n = \sum_{i\in B(n)} 2^i$.

The *Rado graph* has vertex set $\mathbb{N}$, with an edge from i to j if ($i \in B(j)$ or $j \in B(i)$). The *Rado ideal* in the exterior algebra E has a generator $x_i x_j$ for each pair (i, j) such that there is no edge from i to j in the Rado graph. The *Rado algebra* Q is the quotient of E by the Rado ideal.

Remark 1.17. See [Rado 1964; Cameron 2001] for discussion of the Rado graph. Although the definition looks very specialised, the appearance is deceptive. Roughly speaking, any countable random graph is isomorphic to the Rado graph with probability one. The proof of this uses a kind of injectivity property of the Rado graph, which is what suggested it to us as being potentially relevant for the present project.

Theorem 1.18. *The Rado algebra is totally incoherent* (*and in particular*, *not Noetherian*). *It satisfies the double annihilator condition*, *but is not self-injective. The reduced quotient is* $Q/\sqrt{0} = \mathbb{F}$.

This will be proved as Propositions 10.5, 10.6 and 10.8 (apart from the fact that $Q/\sqrt{0} = \mathbb{F}$, which is clear).

One major difference between the Rado algebra and the stable homotopy ring is that the former has Krull dimension zero (because all elements in the maximal ideal square to zero) whereas the latter is $\mathbb{Z}_2$ in degree 0 and so has Krull dimension one. Our final example aims to do something similar to the Rado construction but without making all the generators nilpotent. To do this we must work in base ω rather than base 2; this involves some theory of ordinals, which we briefly recall (the book [Johnstone 1987] is an admirably concise reference). There is an exponentiation operation for ordinals (different from the usual one for cardinals). There is a countable ordinal called ϵ_0 such that $\epsilon_0 = \omega^{\epsilon_0}$, and no ordinal $\alpha < \epsilon_0$ satisfies $\alpha = \omega^\alpha$. Any ordinal $\alpha < \epsilon_0$ has a unique Cantor normal form

$$\alpha = \omega^{\beta_1} n_1 + \cdots + \omega^{\beta_r} n_r,$$

where the n_i are positive integers and $\alpha > \beta_1 > \cdots > \beta_r$.

Definition 1.19. We write $\mu_0(\alpha, \beta)$ for the coefficient of ω^β in the Cantor normal form of α. We then put

$$\mu(\alpha, \beta) = \max(\mu_0(\alpha, \beta), \mu_0(\beta, \alpha)),$$

and

$$A = \mathbb{F}[x_\alpha \mid \alpha < \epsilon_0]/(x_\alpha x_\beta^{1+\mu(\alpha,\beta)} \mid \alpha, \beta < \epsilon_0, \alpha \neq \beta).$$

We call A the ϵ_0-*algebra*.

Given any function $\delta\colon \epsilon_0 \to \mathbb{N}$, we can give A a grading such that $|x_\alpha| = \delta(\alpha)$. In Section 11 we will describe a particular function δ with the property that $\delta(\alpha) > 0$ for all α, and all the sets $\delta^{-1}\{n\}$ are finite. This will ensure that the homogeneous pieces A_d are finite for all d.

Theorem 1.20. *If J is any ideal in A that is generated by a finite set of monomials, then $J = \operatorname{ann}_A^2(J)$. However, there are nonmonomial ideals J with $J \neq \operatorname{ann}_A^2(J)$, so A does not satisfy the double annihilator condition, and is not self-injective. Moreover, A is totally incoherent, and the reduced quotient is*

$$A/\sqrt{0} = \mathbb{F}[x_\alpha \mid \alpha < \epsilon_0]/(x_\alpha x_\beta \mid \alpha \neq \beta).$$

This will be proved as Propositions 11.17, 11.21 and 11.22, and Corollary 11.19.

2. General theory of self-injective rings

Let R be a graded commutative ring, and Mod_R the category of graded R-modules. Suppose that R is self-injective. For $M \in \mathrm{Mod}_R$ we put $DM = \mathrm{Hom}_R(M, R)$ (regarded as a graded R-module in the usual way). This construction defines a functor $D\colon\ \mathrm{Mod}_R \to \mathrm{Mod}_R^{\mathrm{op}}$, which is exact because R is self-injective. It follows that D^2 gives an exact covariant functor from Mod_R to itself. There is a natural map $\kappa\colon M \to D^2M$ given by $\kappa(m)(u) = u(m)$. Properties of D^2 are studied under different technical hypotheses in [Bruns and Herzog 1993, Theorem 3.2.13], for example.

Definition 2.1. We let $\mathcal{U} = \mathcal{U}_R$ denote the full subcategory of Mod_R consisting of the modules M for which $\kappa\colon M \to D^2M$ is an isomorphism.

Proposition 2.2. *The category $\mathcal{U}$ is closed under finite direct sums, suspensions and desuspensions, kernels, cokernels, images and extensions. It also contains R itself.*

Proof. This is clear from the exactness of the functor D^2 and the five lemma. □

Corollary 2.3. *If $J \leq R$ is a finitely generated ideal, then J and R/J lie in $\mathcal{U}$.*

Proof. They are the image and cokernel of some map $\bigoplus_{i=1}^n \Sigma^{d_i} R \to R$. □

Remark 2.4. If J is an ideal in R then

$$D(R/J) \simeq \{a \in R \mid aJ = 0\} = \operatorname{ann}_R(J).$$

By dualising the sequence $J \to R \to R/J$, we see that $D(J) = R/\operatorname{ann}_R(J)$. It follows that $D^2(J) = \operatorname{ann}_R(\operatorname{ann}_R(J)) = \operatorname{ann}_R^2(J)$. Thus, we have $J \in \mathcal{U}$ if and only if $J = \operatorname{ann}_R^2(J)$. In particular, if J is finitely generated then $J = \operatorname{ann}_R^2(J)$.

Lemma 2.5. *For any $a \in R_d$ there is an isomorphism $D(Ra) \simeq \Sigma^{-d} Ra$.*

Proof. Given $u \in D(Ra)_e$ we put $\alpha(u) = u(a) \in R_{d+e}$. This defines a map $\alpha\colon D(Ra) \to \Sigma^{-d}R$, which is clearly injective. Note that if $b \in \operatorname{ann}_R(a)$ then $\alpha(a)b = \alpha(ab) = \alpha(0) = 0$. This proves that $\alpha(a) \in \operatorname{ann}^2_R(Ra)_{d+e} = (Ra)_{d+e}$. In the opposite direction, if $c \in (Ra)_{d+e}$ then we have $c = ma$ for some $m \in R_e$, and the rule $\mu_m(x) = mx$ defines an element $\mu_m \in D(Ra)_e$ with $\alpha(\mu_m) = c$. This proves that the image of α is $\Sigma^{-d}Ra$, as required. □

Proposition 2.6. *If R is self-injective and $a \in R$ then $R/\operatorname{ann}(a)$ is also self-injective.*

Proof. Put $Q = R/\operatorname{ann}(a)$, and let $i\colon Q \to R$ be induced by $x \mapsto xa$, so i is injective, with image Ra. For $M \in \operatorname{Mod}_Q$ we write

$$D_Q(M) = \operatorname{Hom}_Q(M, Q) = \operatorname{Hom}_R(M, Q) \quad \text{and} \quad D_R(M) = \operatorname{Hom}_R(M, R).$$

We are given that D_R is exact, and we must show that D_Q is exact. The map $i\colon Q \to R$ gives a natural monomorphism $i\colon D_Q(M) \to D_R(M)$, and it will suffice to show that this is also an epimorphism. For any $\phi\colon M \to R$ we see that $\operatorname{ann}(a).\phi(M) = \phi(\operatorname{ann}(a)M) = \phi(0) = 0$, so $\phi(M) \leq \operatorname{ann}^2_R(a) = Ra$, and $i\colon Q \to Ra$ is an isomorphism, so $\phi = i(\psi)$ for some $\psi \in D_Q(M)$, as required. □

Proposition 2.7. *If R is self-injective and I and J are ideals in R, then*

$$\operatorname{ann}_R(I+J) = \operatorname{ann}_R(I) \cap \operatorname{ann}_R(J) \quad \textit{and} \quad \operatorname{ann}_R(I \cap J) = \operatorname{ann}_R(I) + \operatorname{ann}_R(J).$$

Proof. There is a short exact sequence

$$R/(I \cap J) \xrightarrow{\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]} R/I \oplus R/J \xrightarrow{[1\ -1]} R/(I+J).$$

By applying the exact functor D, we get a short exact sequence

$$\operatorname{ann}_R(I \cap J) \xleftarrow{[1\ 1]} \operatorname{ann}_R(I) \oplus \operatorname{ann}_R(J) \xleftarrow{\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]} \operatorname{ann}_R(I+J).$$

The claim follows. □

Corollary 2.8. *If R is local and self-injective and I and J are nontrivial ideals, then $I \cap J$ is also nontrivial.*

Proof. Let $\mathfrak{m}$ be the maximal ideal. As I and J are nontrivial we have $\operatorname{ann}(I) < R$ and $\operatorname{ann}(J) < R$, so $\operatorname{ann}(I) \leq \mathfrak{m}$ and $\operatorname{ann}(J) \leq \mathfrak{m}$, so

$$\operatorname{ann}(I \cap J) = \operatorname{ann}(I) + \operatorname{ann}(J) \leq \mathfrak{m} < R,$$

so $I \cap J$ is nontrivial. □

3. Criteria for self-injectivity

We first record a graded version of the standard Baer criterion for injectivity.

Definition 3.1. Let R be a graded ring, and let I be a graded R-module. We say that I satisfies the *Baer condition* if for every graded ideal $J \leq R$, every integer d and every R-module homomorphism $\phi\colon \Sigma^d J \to I$, there exists $m \in I_d$ such that $\phi(a) = am$ for all $a \in I$. We say that I satisfies the *finite Baer condition* if the same condition holds for all finitely generated graded ideals J.

Proposition 3.2. *In the above context, the module I is injective if and only if it satisfies the Baer condition.*

Proof. This was originally done in the ungraded context in [Baer 1940], as an application of Zorn's lemma. The proof is also given in many textbooks, such as [Lam 1999, page 63]. It can be modified in an obvious way to keep track of gradings, which gives our statement above. □

Proposition 3.3. *Suppose that I_d is finite for all d, and that I satisfies the finite Baer condition. Then I also satisfies the full Baer condition and so is injective.*

Proof. Consider a graded ideal $J \leq R$ and a homomorphism $\phi\colon \Sigma^d J \to I$. For each finitely generated ideal $K \subseteq J$ we put

$$M(K) = \{m \in I_d \mid \phi(a) = am \text{ for all } a \in K\}.$$

The finite Baer condition means that this is a nonempty subset of the finite set I_d. Choose K such that $|M(K)|$ is as small as possible, and choose $m \in M(K)$. For $a \in J$ it is clear that $M(K + Ra) \subseteq M(K)$, so by the minimality property we must have $M(K + Ra) = M(K)$, so $m \in M(K + Ra)$, so $\phi(a) = am$. This proves the full Baer condition. □

Definition 3.4. Let R be a graded ring, and let I be an R-module. A *test pair* of length r and degree d is a pair (u, v) where $u \in R^r$ and $v \in I^r$ such that the entries u_i and v_i are homogeneous with $|v_i| = |u_i| + d$ for all i. A *block* for such a pair is a vector $b \in R^r$ such that $b.u = 0$ but $b.v \neq 0$ (where $b.x = \sum_i b_i x_i$). A *transporter* is an element $m \in I_d$ such that $v_i = mu_i$ for all i.

Remark 3.5. We implicitly formulate the theory of graded groups in such a way that the zero elements in different degrees are distinct. Thus, the notation $|u|$ is meaningful even if $u = 0$.

Proposition 3.6. *The module I satisfies the finite Baer condition if and only if every test pair has either a block or a transporter.*

Proof. Suppose that every test pair has either a block or a transporter. Consider a finitely generated graded ideal $J \leq R$, and a homomorphism $\phi\colon \Sigma^d J \to R$.

Choose a list $u = (u_1, \dots, u_r)$ of homogeneous elements that generates J, and put $v_i = \phi(u_i) \in I$. Note that if $b \in R^r$ with $b.u = 0$ then we can apply ϕ to see that $b.v = 0$. It follows that the pair (u, v) has no block, so it must have a transporter. This means that there is an element $m \in I_d$ with $\phi(u_i) = u_i m$ for all i, and it follows easily that $\phi(a) = am$ for all $a \in J$, as required.

Conversely, suppose that I satisfies the finite Baer condition. Consider a test pair (u, v) of degree d with no block, and let J be the ideal generated by the entries u_i. Define $\phi\colon \Sigma^d J \to I$ by $\phi\big(\sum_i b_i u_i\big) = \sum_i b_i v_i$ (the absence of a block means that this is well-defined). The finite Baer condition means that there is an element $m \in I_d$ with $\phi(a) = am$ for all $a \in J$, and this m is clearly a transporter for (u, v). □

Corollary 3.7. *Let R be a graded commutative ring such that R_k is finite for all k. Suppose also that there are subrings*

$$R(0) \leq R(1) \leq R(2) \leq \cdots \leq R$$

such that each $R(n)$ is self-injective and $R = \bigcup_n R(n)$. Then R is self-injective.

Proof. Any test pair $(u, v) \in R^r \times R^r$ can be regarded as a test pair over $R(n)$ for sufficiently large n. As $R(n)$ is self-injective, there must be a block in $R(n)^r$ or a transporter in $R(n)$. It is clear from the definitions that such a block or transporter still qualifies as a block or transporter over R, so we see that R satisfies the finite Baer condition. As we have assumed that R_k is finite for all k, we can use Proposition 3.3 to see that R is injective as an R-module. □

Theorem 3.8. *Let R be a graded commutative ring such that R_k is finite for all k. The following conditions are equivalent*:

(a) *R is self-injective.*

(b) *For all finitely generated ideals $J, K \leq R$ we have $\operatorname{ann}_R^2(J) = J$ and*

$$\operatorname{ann}_R(J \cap K) = \operatorname{ann}_R(J) + \operatorname{ann}_R(K).$$

(c) *For all elements $a \in R$ and every finitely generated ideal $J \leq R$ we have $\operatorname{ann}_R^2(a) = Ra$ and*

$$\operatorname{ann}_R(J \cap Ra) = \operatorname{ann}_R(J) + \operatorname{ann}_R(a).$$

Proof. It follows from Remark 2.4 and Proposition 2.7 that (a) implies (b). If (b) holds, then (c) follows immediately. Now suppose (c) holds. As we have assumed that R_k is finite for all k, we may use the theory of blocks and transporters. We proceed by induction on the length of a test pair to show that every test pair over the ring R has either a block or a transporter. Let $(u; v)$ be a test pair of length 1 and degree d. Suppose this test pair has neither block nor transporter. Then $\operatorname{ann}_R(u) \leq \operatorname{ann}_R(v)$ and by assumption we have $Rv = \operatorname{ann}_R^2(v) \leq \operatorname{ann}_R^2(u) = Ru$, that is, $v = um$ for some $m \in R_d$. Since m is a transporter for this test pair, we have a contradiction.

Now suppose each test pair of length $\leq k$ and arbitrary degree has either a block or a transporter. A test pair of length $k + 1$ and degree d takes the form $(u, u_{k+1}; v, v_{k+1})$, where $(u; v)$ is a test pair of length k and degree d and (u_{k+1}, v_{k+1}) is a test pair of length 1 and degree d. By the inductive hypothesis, both the test pairs $(u; v)$ and (u_{k+1}, v_{k+1}) have either a block or a transporter. If $(u; v)$ has block r, then $(r, 0)$ is a block for the test pair $(u, u_{k+1}; v, v_{k+1})$. Similarly, if (u_{k+1}, v_{k+1}) has block r_{k+1}, then $(0, \ldots, 0, r_{k+1})$ is a block for the test pair $(u, u_{k+1}; v, v_{k+1})$. Otherwise, $(u; v)$ must have transporter $m \in R_d$ and (u_{k+1}, v_{k+1}) must have transporter $n \in R_d$. In this situation, suppose the test pair $(u, u_{k+1}; v, v_{k+1})$ has neither block nor transporter and let J be the ideal generated by the entries of u. The absence of a block implies that there is a well defined map $\phi : \Sigma^d(J + Ru_{k+1}) \to R$ defined by $\phi\big(\sum_{i=1}^{k+1} b_i u_i\big) = \sum_{i=1}^{k+1} b_i v_i$. Now let s be an element in the intersection $J \cap Ru_{k+1}$. Then we must have $s = \sum_{i=1}^{k} s_i u_i = s_{k+1}u_{k+1}$ for elements $s_i \in R$ for each i. Applying the map ϕ to the zero element $\big(\sum_{i=1}^{k} s_i u_i\big) - s_{k+1}u_{k+1}$ gives

$$0 = \left(\sum_{i=1}^{k} s_i v_i\right) - s_{k+1}v_{k+1} = \left(\sum_{i=1}^{k} s_i u_i m\right) - s_{k+1}u_{k+1}n = s(m - n).$$

Thus it follows that the element $m - n$ is in the annihilator ideal $\operatorname{ann}_R(J \cap Ru_{k+1})$. By assumption, we have $\operatorname{ann}_R(J \cap Ru_{k+1}) = \operatorname{ann}_R(J) + \operatorname{ann}_R(u_{k+1})$. Now let $m - n = x - y$, where $x \in \operatorname{ann}_R(J)$ and $y \in \operatorname{ann}_R(u_{k+1})$, and put $z = m - x = n - y$. Since $u_i z = u_i(m - x) = u_i m = v_i$ for each $i \leq k$ and $u_{k+1}z = u_{k+1}(n - y) = u_{k+1}n = v_{k+1}$ it follows that z is a transporter for the test pair $(u, u_{k+1}; v, v_{k+1})$. As this gives a contradiction, it follows that every test pair of length $k + 1$ and arbitrary degree must have either a block or transporter. We deduce that every test pair in the ring R must have either a block or transporter, and since R_k is finite for each k, we can use Proposition 3.6 to show that R is injective as an R-module. $\square$

4. The Noetherian case

Theorem 4.1. *Let R be a Noetherian graded commutative ring. Then the following are equivalent*:

(a) *R is self-injective.*

(b) *For every ideal $J \leq R$ we have* $\operatorname{ann}_R^2(J) = J$.

(c) *R is Artinian* (*and thus is a finite product of Artinian local rings*), *and each of the local factors has one-dimensional socle.*

Statements similar to this are certainly well-known (see, for example, [Bruns and Herzog 1993, Exercise 3.2.15]), but we do not know a reference for this precise formulation. For completeness we will give a self-contained proof after some lemmas.

Lemma 4.2. *Let R be an Artinian local graded ring, with maximal ideal $\mathfrak{m}$, and put $K = R/\mathfrak{m}$. Suppose that the socle $\operatorname{soc}(R) = \operatorname{ann}_R(\mathfrak{m})$ has dimension one over K. Then every nonzero ideal in R contains $\operatorname{soc}(R)$.*

Proof. Let I be a nonzero ideal. By the Artinian condition, we can choose an ideal J that is minimal among nonzero ideals contained in I. Recall that every Artinian ring is Noetherian (see, for example, [Matsumura 1980, Theorem 3.2]), so we can use Nakayama's lemma to see that $\mathfrak{m}J < J$ and thus (by minimality) that $\mathfrak{m}J = 0$. This means that J is a nontrivial K-subspace of $\operatorname{soc}(R)$, but $\operatorname{soc}(R)$ has dimension one, so $J = \operatorname{soc}(R)$, so $\operatorname{soc}(R) \leq I$. □

Lemma 4.3. *Suppose that R is as in Lemma 4.2. Then for all ideals $J \leq R$ we have $\operatorname{ann}_R^2(J) = J$.*

Proof. First, it is standard that we can fit together a composition series for J with a composition series for R/J to get a chain

$$0 = I_0 < I_1 < \cdots < I_r = R$$

with $I_i/I_{i-1} \simeq K$ for all i, and $J = I_t$ for some t. Now let A_j be the annihilator of I_j, so we have

$$R = A_0 \geq A_1 \geq \cdots \geq A_r = 0.$$

Now $\mathfrak{m}A_i I_{i+1} = A_i(\mathfrak{m}I_{i+1}) \leq A_i I_i = 0$, so $A_i I_{i+1} \leq \operatorname{soc}(R)$. On the other hand, we have $A_i I_i = 0$ and $A_{i+1} I_{i+1} = 0$. We therefore have a natural map

$$\xi_i \colon A_i/A_{i+1} \to \operatorname{Hom}_K(I_{i+1}/I_i, \operatorname{soc}(R))$$

given by $\xi_i(a + A_{i+1})(b + I_i) = ab$. It is clear from the definitions that this is injective, and the codomain is isomorphic to K, so A_i/A_{i+1} is either 0 or K. It is standard that any two composition series have the same length, so we must have $A_i/A_{i+1} \simeq K$ for all i, so A_i has length $r - i$. After applying the same logic to the composition series $\{A_{r-i}\}_{i=0}^{r}$ we see that the ideal $\operatorname{ann}(A_i) = \operatorname{ann}^2(I_i)$ has length i. We also know that $I_i \leq \operatorname{ann}^2(I_i)$ and that I_i also has length i; it follows that $I_i = \operatorname{ann}^2(I_i)$, as required. □

Corollary 4.4. *Suppose that R is as in Lemma 4.3. Then R is self-injective.*

Proof. Consider an ideal $I \leq R$ and an R-module map $f \colon I \to R$. Choose a composition series $0 = J_0 < J_1 < \cdots < J_r = I$. We have $J_i/J_{i-1} \simeq K$ so we can find $a_i \in J_i \setminus J_{i-1}$ such that $J_i = J_{i-1} + Ra_i$ with $\mathfrak{m}a_i \leq J_{i-1}$.

We will construct elements $x_0, \dots, x_r \in R$ such that $f(a) = ax_i$ for all $a \in J_i$. We start with $x_0 = 0$. Now suppose we have found x_{i-1}. Put $u_i = f(a_i) - x_{i-1}a_i$. Using the fact that $\mathfrak{m}a_i \leq I_{i-1}$ we find that $\mathfrak{m}u_i = 0$, so $u_i \in \mathrm{soc}(R)$. Next, we have $a_i \notin I_{i-1} = \mathrm{ann}^2(I_{i-1})$, so $\mathrm{ann}(I_{i-1})a_i \neq 0$. As every nontrivial ideal contains the socle, we see that $u_i \in \mathrm{ann}(I_{i-1})a_i$, so we can write $u_i = y_i a_i$ for some y_i with $y_i I_{i-1} = 0$. We now put $x_i = x_{i-1} + y_i$. By construction we have $f(a) = ax_i$ for $a \in I_{i-1}$ or for $a = a_i$, and it follows that this equation holds for all $a \in I_i$ as required. At the end of the induction we have an element x_r which fulfils Baer's criterion. □

Proof of Theorem 4.1. It follows from Remark 2.4 that (a) implies (b). Now suppose that (b) holds. Consider a descending chain of ideals $I_0 \geq I_1 \geq I_2 \geq \cdots$ in R. The ideals $\mathrm{ann}(I_k)$ then form an ascending chain, which must eventually stabilise because R is Noetherian. We can thus take annihilators again to see that the original chain also stabilises. This shows that R is Artinian. It follows in a standard way that there are only finitely many maximal ideals, and that R is the product of its maximal localisations. We thus have a splitting $R = \prod_{i=1}^n R_i$ say, where each factor R_i an Artinian local ring. It follows that the lattice of ideals in R is the product of the corresponding lattices for the factors R_i, and thus that each R_i satisfies condition (b). We can thus reduce to the case where R is local, with maximal ideal $\mathfrak{m}$ say. Recall that the socle is $\mathrm{soc}(R) = \{a \in R \mid a\mathfrak{m} = 0\} = \mathrm{ann}_R(\mathfrak{m})$, which is naturally a vector space over the field $K = R/\mathfrak{m}$. If $\mathrm{soc}(R)$ were zero we would have $\mathfrak{m} = \mathrm{ann}^2(\mathfrak{m}) = \mathrm{ann}(\mathrm{soc}(R)) = \mathrm{ann}(0) = R$, which is a contradiction. We can therefore choose a nonzero element $u \in \mathrm{soc}(R)$. We find that $Ku = Ru$ is a nonzero ideal in R, so $\mathrm{ann}(Ku)$ is a proper ideal containing $\mathrm{ann}(\mathrm{soc}(R)) = \mathfrak{m}$, so $\mathrm{ann}(Ku) = \mathfrak{m}$ by maximality. We can now take annihilators again to see that $Ku = \mathrm{ann}(\mathfrak{m}) = \mathrm{soc}(R)$, so $\mathrm{soc}(R)$ is one-dimensional. This proves (c).

Finally, we will assume (c) and prove (a). It is again easy to reduce to the case where R is local, and the local case is covered by Corollary 4.4. □

Definition 4.5. Let K be a field. A *Poincaré duality algebra* over K is a graded commutative K-algebra R equipped with a K-linear map $\theta\colon R_d \to K$ for some $d \geq 0$ such that:

- For $i < 0$ or $i > d$ we have $R_i = 0$.
- $R_0 = K$.
- For $0 \leq i \leq d$ we have $\dim_K(R_i) < \infty$, and the map $(a, b) \mapsto \theta(ab)$ defines a perfect pairing between R_i and R_{d-i}.

Proposition 4.6. *Every Poincaré duality algebra is self-injective.*

Proof. Let R be a Poincaré duality algebra of top dimension d, and put $\mathfrak{m} = \bigoplus_{i>0} R_i$. It is clear that $R/\mathfrak{m} = K$ and $\mathfrak{m}^{d+1} = 0$, and it follows that $\mathfrak{m}$ is the unique maximal ideal. As R has finite total dimension over K it is clearly Artinian. The perfect pairing condition implies that $\mathrm{soc}(R) = R_d$ and that this has dimension one. It follows by Theorem 4.1 that R is self-injective.

Alternatively, for any R-module M we can define a natural map

$$\tau\colon \mathrm{Hom}_R(M, R) \to \mathrm{Hom}_K(M_d, K)$$

by $\tau(\phi) = \theta \circ \phi_d$. Using the perfectness of the pairing we see that this is an isomorphism. As K is a field, the functor $M \mapsto \mathrm{Hom}_K(M_d, K)$ is exact, and it follows that the functor $M \mapsto \mathrm{Hom}_R(R, R)$ is also exact, or in other words that R is injective as an R-module. □

Example 4.7. Put

$$E = \mathbb{F}[x_0, x_1, x_2, \dots]/(x_i^2 \mid i \geq 0),$$

with $|x_i| = 2^i$. For any finite set $I \subset \mathbb{N}$ we put $x_I = \prod_{i\in I} x_i$, so $|x_I| = \sum_{i\in I} 2^i$ and the elements x_I form a basis for E over $\mathbb{F}$. It follows that $E_k \simeq \mathbb{F}$ for all $k \geq 0$, and $E_k = 0$ for $k < 0$. Let $E(n)$ be the subalgebra of E generated by $x_0, \dots, x_{n-1}$. This is a Poincaré duality algebra, with socle generated by the element $\prod_{i<n} x_i$, and it is clear that $E = \bigcup_n E(n)$. Corollary 3.7 therefore tells us that E is self-injective.

5. Coherence

We now briefly recall some standard ideas about finite presentation.

Definition 5.1. Let R be a graded commutative ring, and let M be a graded R-module. Then we see from [Lam 1999, Section 4D] the following are equivalent:

(a) There exists an exact sequence

$$P_1 \xrightarrow{f} P_0 \xrightarrow{g} M \to 0,$$

where P_0 and P_1 are finitely generated free modules.

(b) M is finitely generated, and for every epimorphism $g\colon P_0 \to M$ (with P_0 a finitely generated free module) the module $\ker(g)$ is also finitely generated.

If these conditions hold, we say that M is *finitely presented*.

Remark 5.2. By *finitely generated free module*, we mean one of the form $\bigoplus_{i=1}^{r} \Sigma^{d_i} R$; we do not assume that the degree shift d_i is zero.

Corollary 5.3. *If R is Noetherian, then every finitely generated ideal is finitely presented.*

Proof. Condition (b) is clearly satisfied. □

As we stated in Definition 1.2, a graded ring R is said to be *coherent* if every finitely generated ideal is finitely presented, and *totally incoherent* if the only finitely presented ideals are 0 and R. It is clear that every Noetherian ring is coherent. We mention as background that if R is coherent, then the category of finitely generated modules is closed under images, kernels, cokernels and extensions, so it is an abelian category. The following example is standard:

Proposition 5.4. *The infinite exterior algebra E (as in Example 4.7) is coherent.*

Proof. Let $E(n)$ be the subalgebra generated by $x_0, \dots, x_{n-1}$, and let $E'(n)$ be generated by the remaining variables, so $E = E(n) \otimes_{\mathbb{F}} E'(n)$. Any finitely generated ideal is the image of some E-linear map $g\colon E^r \to E$, which will have the form $g(u) = u.v$ for some vector $v \in E^r$. We must show that the module $K = \ker(g)$ is finitely generated. Choose n large enough that $v_i \in E(n)$ for all i. Now v gives a map $g'\colon E(n)^r \to E(n)$ of $E(n)$-modules, and $E(n)$ is Noetherian, so the module $K' = \ker(g')$ is finitely generated over $E(n)$. We can identify g with $g' \otimes 1$ with respect to the splitting $E = E(n) \otimes E'(n)$, and it follows that $K = K' \otimes E(n)'$, and thus that any finite generating set for K' over $E(n)$ also generates K over E. $\square$

The following result will be our main tool for proving incoherence results.

Lemma 5.5. *Let A be a local graded ring, with maximal ideal $\mathfrak{m}$, and let I be a finitely presented ideal in A. Then for each $u \in I \setminus \mathfrak{m}I$, the image of $\operatorname{ann}_A(u)$ in $\mathfrak{m}/\mathfrak{m}^2$ has finite dimension over $A/\mathfrak{m}$.*

Note here that as $u \notin \mathfrak{m}I$ we have $u \neq 0$, so $\operatorname{ann}_A(u) \leq \mathfrak{m}$ and it is meaningful to talk about the image in $\mathfrak{m}/\mathfrak{m}^2$.

Proof. As I is finitely generated over A, we see that $I/\mathfrak{m}I$ is a finite-dimensional vector space over $A/\mathfrak{m}$. We can choose a basis for this space containing the image of u, and then choose elements of I lifting these basis elements. This gives a list $v_1, \dots, v_n \in I$ with $v_1 = u$ such that the corresponding map $g\colon A^n \to I$ induces an isomorphism $\bar{g}\colon (A/\mathfrak{m})^n \to I/\mathfrak{m}I$. Now $\operatorname{cok}(g)$ is a finitely generated module with $\mathfrak{m}.\operatorname{cok}(g) = \operatorname{cok}(g)$, so $\operatorname{cok}(g) = 0$ by Nakayama's lemma, and so g is an epimorphism. As I is assumed to be finitely presented, we see that $\ker(g)$ is also finitely generated over A. Moreover, as $\bar{g}$ is an isomorphism we see that $\ker(g) \leq \mathfrak{m}^n$. It follows that the image of $\ker(g)$ in $(\mathfrak{m}/\mathfrak{m}^2)^n$ is finite-dimensional. The intersection of $\ker(g)$ with the first copy of A in A^n is just the annihilator of u, so we see that the image of $\operatorname{ann}_A(u)$ in $\mathfrak{m}/\mathfrak{m}^2$ is finite-dimensional. $\square$

Corollary 5.6. *Let A be a local graded ring, with maximal ideal $\mathfrak{m}$. Suppose that all $u \in A$ satisfy one of the following conditions: $u = 0$; the image of $\operatorname{ann}_A(u)$ in $\mathfrak{m}/\mathfrak{m}^2$ has infinite dimension; or u is invertible.*

Then A is totally incoherent.

Proof. Let I be a finitely presented ideal. If $\mathfrak{m}I = I$ then $I = 0$ by Nakayama's lemma. Otherwise, we can choose $u \in I \setminus \mathfrak{m}I$. As $u \notin \mathfrak{m}I$ we have $u \neq 0$. By the lemma, the image of $\mathrm{ann}_A(u)$ in $\mathfrak{m}/\mathfrak{m}^2$ must have finite dimension. Thus, the first two possibilities are excluded, and u must be invertible. As $u \in I$ we conclude that $I = A$. □

Next we record a graded version of Chase's theorem for coherent rings.

Theorem 5.7. *Let R be a graded commutative ring. The following conditions are equivalent*:

(a) *R is coherent.*

(b) *For all elements $a \in R$ and for every finitely generated ideal $J \leq R$, the conductor ideal*
$$(J : a) = \{r \in R \mid ra \in J\}$$
is finitely generated.

(c) *For all elements $a \in R$, the annihilator ideal $\mathrm{ann}_R(a)$ is finitely generated, and for all finitely generated ideals $J, K \leq R$, the intersection $J \cap K$ is finitely generated.*

Proof. The ungraded version of the proof is given in many textbooks such as [Lam 1999, page 142]. It can be modified in an obvious way to keep track of gradings, which gives our statement above. □

Theorem 5.8. *Let R be a graded commutative ring such that R_k is finite for all k. The following conditions are equivalent*:

(a) *R is coherent and self-injective.*

(b) *R is coherent and for all finitely generated ideals $J \leq R$ we have $\mathrm{ann}_R^2(J) = J$.*

(c) *For every finitely generated ideal $J \leq R$, the ideal $\mathrm{ann}_R(J)$ is finitely generated and $\mathrm{ann}_R^2(J) = J$.*

(d) *R is self injective and for all finitely generated ideals $J \leq R$, the ideal $\mathrm{ann}_R(J)$ is finitely generated.*

Proof. It follows from Remark 2.4 that (a) implies (b). To show that (b) implies (c) we need to show that the ideal $\mathrm{ann}_R(J)$ is finitely generated for each finitely generated ideal $J \leq R$. If we let $(r_1, \ldots, r_n)$ be generators for the ideal J, then we can take the annihilator of J to give $\mathrm{ann}_R(J) = \bigcap_i \mathrm{ann}_R(r_i)$. Since R is assumed to be coherent, it follows from part (c) of Theorem 5.7 that $\mathrm{ann}_R(r_i)$ is finitely generated for each i and that a finite intersection of finitely generated ideals is also finitely generated. Thus $\mathrm{ann}_R(J)$ is finitely generated as claimed. Now suppose that part (c) holds. To prove that (c) implies (d), we need to show that R is injective as an R-module. For all ideals $J, K \leq R$ we have

$$\operatorname{ann}_R(\operatorname{ann}_R(J)+\operatorname{ann}_R(K)) = \operatorname{ann}_R^2(J)\cap\operatorname{ann}_R^2(K) = J\cap K.$$

By assumption, the ideal sum $\operatorname{ann}_R(J)+\operatorname{ann}_R(K)$ must be finitely generated. Thus we can take double annihilators to give

$$\operatorname{ann}_R(J)+\operatorname{ann}_R(K) = \operatorname{ann}_R(J\cap K).$$

Since R_k is finite for each k, we can use part (b) of Theorem 3.8 to complete the claim. We now conclude by showing that (d) implies (a). By assumption, the annihilator ideal $\operatorname{ann}_R(a)$ is finitely generated for all elements $a \in R$. Then for all ideals $J, K \leq R$ we know that the ideal sum $\operatorname{ann}_R(J)+\operatorname{ann}_R(K)$ is finitely generated by assumption. By taking annihilators we then have

$$\operatorname{ann}_R(\operatorname{ann}_R(J)+\operatorname{ann}_R(K)) = \operatorname{ann}_R^2(J)\cap\operatorname{ann}_R^2(K) = J\cap K,$$

where the double annihilator condition holds by Remark 2.4. However, by assumption, the annihilator of a finitely generated ideal is also finitely generated. Thus the intersection $J\cap K$ must be finitely generated. It follows from part (c) of Theorem 5.7 that the ring R is coherent as claimed. □

6. Self-injective adjustment

Definition 6.1. We write $\mathcal{R}$ for the category of commutative graded $\mathbb{F}$-algebras such that:

(a) $R_k = 0$ for all $k < 0$.

(b) $R_0 = \mathbb{F}$.

(c) R_k is finite for all $k > 0$.

Proposition 6.2. *Let R be a ring in $\mathcal{R}$, and let $\mathcal{P}$ be a finite set of test pairs in R that have no transporters. Let m be a positive integer. Then there is an extension $R' \geq R$ of graded rings such that:*

(a) *R' is also in $\mathcal{R}$.*

(b) *$R'_k = R_k$ for all $k < m$.*

(c) *Each test pair in $\mathcal{P}$ has a block in R'.*

Proof. List the elements of $\mathcal{P}$ as $(u_0, v_0), \dots, (u_{p-1}, v_{p-1})$ say. Suppose that (u_t, v_t) has length r_t, and let d_t be the maximum of the degrees of the entries $u_{t,j}$ for $0 \leq j < r_t$. Let P be the polynomial ring obtained from R by adjoining variables $b_{t,j}$ for $0 \leq t < p$ and $0 \leq j < r_t$, with $|b_{t,j}| = m + d_t - |u_{t,j}| \geq m > 0$. Put $w_t = \sum_{j=0}^{r_t-1} b_{t,j}u_{t,j} \in P$ and $R' = P/(w_0, \dots, w_{p-1})$. There is an evident ring map

$\eta \colon R \to R'$, and also a ring map $\pi : R' \to R$ given by $\pi(b_{t,j}) = 0$ for all t and j. It is clear that $\pi\eta = 1$, so η is injective, and we can use it to regard R' as an extension of R. As $|b_{t,j}| \geq m > 0$, it is easy to see that $R' \in \mathcal{R}$ and that the map $R_k \to R'_k$ is surjective (and therefore bijective) for $k < m$. By construction we have $b_t.u_t = 0$ in R'. We claim that $b_t.v_t \neq 0$ in R', or equivalently that $b_t.v_t$ cannot be written as $\sum_s c_s w_s$ in P. To see this, let c^* denote the constant term in the polynomial c_t. By examining the coefficient of $b_{t,j}$ in the equation $b_t.v_t = \sum_s c_s w_s$ we obtain $v_{t,j} = c^* u_{t,j}$ for all j, which means that c^* is a transporter for (u_t, v_t), contrary to assumption. Thus, b_t is a block for (u_t, v_t) in R', as required. □

Definition 6.3. Let R be a ring in $\mathcal{R}$, and let (u, v) be a test pair for R. We say that (u, v) is *good* if it has either a block or a transporter, and *bad* otherwise. We say that (u, v) is *nondegenerate* if $u_i \neq 0$ for all i. For any homogeneous element $x \in R$ we put $|x|_+ = \max(0, |x|)$. The *weight* of (u, v) is $\sum_i (1 + |u_i|_+ + |v_i|_+)$.

Lemma 6.4. *Let R be a ring in $\mathcal{R}$, and suppose that all nondegenerate test pairs are good. Then R is self-injective.*

Proof. Consider an arbitrary test pair $(u, v) \in R^r \times R^r$. If there exists i such that $u_i = 0$ but $v_i \neq 0$, then the basis vector $e_i \in R^r$ is a block for (u, v). Otherwise, let (u', v') be the test pair obtained by removing all zeros from u and the corresponding zeros from v. This is nondegenerate, so it has a block or a transporter. If b' is a block for (u', v'), then we can construct a block for (u, v) by inserting some zeros. If m' is a transporter for (u', v'), then it is also a transporter for (u, v). We therefore see that all test pairs for R are good, so R is self-injective. □

Lemma 6.5. *There are only finitely many nondegenerate bad test pairs of any given weight.*

Proof. Consider an integer $N \geq 0$. Any nondegenerate bad test pair (u, v) of weight N must have length at most N. Moreover, as (u, v) is nondegenerate we must have $u_i \neq 0$ for all i, and as $R \in \mathcal{R}$ this means that $|u_i| \geq 0$. We also have $\sum_i |u_i| \leq \text{weight}(u, v) = N$. It is clear from this (and the finiteness of R_k) that there are only finitely many possibilities for u. Next, let d be the degree of (u, v), so $|v_i| = |u_i| + d$. From this it is clear that $d \leq N$. If d is sufficiently negative then we will have $v_i = 0$ for all i, so 0 is a transporter for (u, v), contradicting the assumption that (u, v) is bad. We therefore see that there are only finitely many possibilities for d. Given u and d, it is clear that there are only finitely many possibilities for v. □

Theorem 6.6. *Suppose that $R \in \mathcal{R}$, and that $m \geq 0$. Then there is an extension $R' \geq R$ such that*:

(a) *R' is also in R.*

(b) *$R'_k = R_k$ for all $k < m$.*

(c) *R' is self-injective.*

Proof. We define rings $R'(0) \leq R'(1) \leq \cdots$ as follows. We start with $R'(0) = R$. For each $k \geq 0$, we let $R'(k+1)$ be an extension of $R'(k)$ that agrees with $R'(k)$ in degrees less than $k+m$, such that every nondegenerate bad test pair of weight at most k in $R'(k)$ has a block in $R'(k+1)$. This can be constructed by Proposition 6.2 and Lemma 6.5. Now take R' to be the colimit of the rings $R'(k)$. By construction we have $R'_i = R'(k)_i$ for sufficiently large k, and using this it is clear that $R' \in \mathcal{R}$. Consider a nondegenerate test pair $(u, v) \in R'$. For sufficiently large k we can assume that $k \geq \mathrm{weight}(u, v)$ and that $u_i, v_i \in R'(k)$ for all i. If (u, v) is good in $R'(k)$ then it is good in R'. If it is bad in $R'(k)$ then by construction it becomes good in $R'(k+1)$ and therefore in R'. $\square$

7. The cube algebra

Recall that in the statement of Theorem 1.6 we introduced the ring

$$C = \mathbb{F}[y_0, y_1, \dots]/(y_i^3 + y_i y_{i+1} \mid i \geq 0),$$

with the grading given by $|y_i| = 2^i$. We now investigate the structure of this ring (which we call the *cube algebra*).

Definition 7.1. We also put

$$\begin{aligned} C[n,\infty] &= \mathbb{F}[y_n, y_{n+1}, \dots]/(y_i^3 + y_i y_{i+1} | n \leq i < \infty), \\ C[n,m] &= \mathbb{F}[y_n, \dots, y_m]/(y_i^3 + y_i y_{i+1} | n \leq i < m), \\ \bar{C}[n,m] &= C[n,m]/y_m. \end{aligned}$$

Lemma 7.2. *The evident maps*

$$\begin{tikzcd} C[n+1,m] \arrow[r] \arrow[d] & C[n+1,m+1] \arrow[r] \arrow[d] & C[n+1,\infty] \arrow[d] \\ C[n,m] \arrow[r] \arrow[d] & C[n,m+1] \arrow[r] \arrow[d] & C[n,\infty] \arrow[d] \\ C[0,m] \arrow[r] & C[0,m+1] \arrow[r] & C[0,\infty] = C \end{tikzcd}$$

are all split injective, so all the rings mentioned can be considered as subrings of C.

Proof. There is a graded ring map $\tau_0 \colon \mathbb{F}[y_0, y_1, \dots] \to C[n,m]$ given by

$$\tau_0(y_i) = \begin{cases} 0 & \text{if } i < n, \\ y_i & \text{if } n \leq i \leq m, \\ y_m^{2^{i-m}} & \text{if } m \leq i. \end{cases}$$

It is straightforward to check that $\tau_0(y_i^3 + y_i y_{i+1}) = 0$ for all $i \geq 0$, so there is an induced map $\tau\colon C \to C[n, m]$. Clearly the composite $C[n, m] \to C \xrightarrow{\tau} C[n, m]$ is the identity, so the map $C[n, m] \to C$ is injective for all m and n. The other claims follow from this. $\square$

Definition 7.3. We write P for the polynomial ring $\mathbb{F}[y_0, y_1, \dots]$, so that C is a quotient of P. A *multiindex* is a sequence $\alpha = (\alpha_0, \alpha_1, \dots)$ of natural numbers with $\alpha_i = 0$ for $i \gg 0$. We write MP for the set of all multiindices. Given $\alpha \in MP$ we write $y^\alpha = \prod_i y_i^{\alpha_i}$ and $|\alpha| = |y^\alpha| = \sum_i \alpha_i 2^i$. It is clear that the set $BP = \{y^\alpha \mid \alpha \in MP\}$ is a basis for P over $\mathbb{F}$.

Definition 7.4. We put

$$\begin{aligned}
M'C[n, m] &= \{\alpha \in MP \mid \alpha_i = 0 \text{ for } i < n \text{ or } i > m \text{ and } \alpha_i < 3 \text{ for } n \leq i < m\},\\
M\bar{C}[n, m] &= \{\alpha \in MP \mid \alpha_i = 0 \text{ for } i < n \text{ or } i \geq m\},\\
B'C[n, m] &= \{y^\alpha \mid \alpha \in M'C[n, m]\},\\
B\bar{C}[n, m] &= \{y^\alpha \mid \alpha \in M\bar{C}[n, m]\}.
\end{aligned}$$

Note that in the definition of $M'C[n, m]$ the constraint $\alpha_i < 3$ does not apply when $i = m$, so in particular $M'C[n, m]$ is infinite.

Proposition 7.5. *$B'C[n, m]$ is a basis for $C[n, m]$, and $B\bar{C}[n, m]$ is a basis for $\bar{C}[n, m]$. Moreover, $\bar{C}[n, m]$ is a Poincaré duality algebra over $\mathbb{F}$.*

The proof depends on the following result:

Lemma 7.6. *Let A be a commutative algebra over $\mathbb{F}$, let $f(t) \in A[t]$ be a monic polynomial of degree d, and put $B = A[x]/f(x)$. Then $\{1, x, \dots, x^{d-1}\}$ is a basis for B over A. Moreover, if A is finite-dimensional over $\mathbb{F}$ and has Poincaré duality, then the same is true of B.*

Proof. We first claim that any polynomial $g(x) \in A[x]$ can be expressed uniquely in the form $g(x) = q(x) f(x) + r(x)$ with $\deg(r(x)) < d$. This can easily be proved by induction on the degree of $g(x)$, and it follows directly that $\{1, \dots, x^{d-1}\}$ is a basis for B over A. Now suppose that A has Poincaré duality, so there is a linear map $\theta\colon A \to \mathbb{F}$ such that the bilinear form $(u, v) \mapsto \theta(u, v)$ is perfect. This means that there exist bases $\{u_0, \dots, u_{n-1}\}$ and $\{v_0, \dots, v_{n-1}\}$ for A such that $\theta(u_i v_j) = \delta_{ij}$. Now define $\phi\colon B \to \mathbb{F}$ by $\phi\big(\sum_{i=0}^{d-1} a_i x^i\big) = \theta(a_{d-1})$. We define bases $\{s_0, \dots, s_{nd-1}\}$ and $\{t_0, \dots, t_{nd-1}\}$ for B by $s_{ni+j} = x^i u_j$ and $t_{ni+j} = x^{d-1-i} v_j$ for $0 \leq i < d$ and $0 \leq j < n$. It is clear that $\phi(s_k t_k) = 1$. Suppose we have $0 \leq k < k' < nd$. Write $k = ni + j$ and $k' = ni' + j'$ as before; we must have either $i < i'$, or ($i = i'$ and $j < j'$). In either case, we find that $\phi(s_i t_j) = 0$. Thus, the matrix of ϕ with respect to our bases is triangular, with ones on the diagonal, proving that ϕ gives a perfect pairing on B. $\square$

Proof of Proposition 7.5. From the definitions we have $C[m,m]=\mathbb{F}[y_n]$ and $B'C[m,m]=\{y_n^{\alpha_n} \mid \alpha_n\in\mathbb{N}\}$ so clearly $B'C[m,m]$ is a basis for $C[m,m]$. Similarly, it is clear that the set $\bar{C}[m,m]=\{1\}$ is a basis for the ring $\bar{C}[m,m]=C[m,m]/y_m=\mathbb{F}$, and that this has Poincaré duality.

Next, $C[n,m]$ can be described as

$$C[n+1,m][y_n]/f(y_n),$$

where $f(t)=t^3+y_{n+1}t$ is a monic polynomial of degree three with coefficients in $C[n+1,m]$. It also follows that $\bar{C}[n,m]=\bar{C}[n+1,m][y_n]/f(y_n)$. All claims in the proposition now follow by downwards induction on n using Lemma 7.6. $\square$

Remark 7.7. Note that the algebra

$$\bar{C}[n,m]=\frac{\mathbb{F}[y_n,y_{n+1},\dots,y_{m-1}]}{(y_n^3+y_ny_{n+1},\dots,y_{m-1}^3)}$$

has the same number of relations as generators, and has finite dimension over $\mathbb{F}$. It is known that in this situation the sequence of relations is necessarily regular, and that the algebra automatically has Poincaré duality. (This can be extracted from [Matsumura 1980, Section 17], for example.) This would give another approach to Proposition 7.5.

Definition 7.8. Let α be a multiindex. We say that

(a) α is *flat* if $\alpha_i<3$ for all i;

(b) α is *n-truncated* if $\alpha_i=0$ for all $i<n$;

(c) α is *m-solid* if it is flat and whenever $m\le p\le q$ and $\alpha_q>0$ we also have $\alpha_p>0$.

We consider all flat multiindices to be ∞-solid. For $0\le n\le m\le\infty$ we put

$$MC[n,m]=\{\alpha\in MP \mid \alpha \text{ is } n\text{-truncated and } m\text{-solid}\},$$

and $BC[n,m]=\{y^\alpha \mid \alpha\in MC[n,m]\}$. We also write MC for the set $MC[0,\infty]$ of all flat multiindices.

Proposition 7.9. *$BC[n,\infty]$ is a basis for $C[n,\infty]$.*

Proof. We must show that for each degree $d\in\mathbb{N}$, the set $BC[n,\infty]_d$ is a basis for $C[n,\infty]_d$. Choose $m>n$ such that $2^m>d$. Then clearly $BC[n,\infty]_d=B'C[n,m]_d$ and $C[n,\infty]_d=C[n,m]_d$ so the claim follows from Proposition 7.5. $\square$

It is also true that $BC[n,m]$ is a basis for $C[n,m]$ when $m<\infty$, but it is convenient to leave the proof until later.

Proposition 7.10. *For any multiindex $\alpha\in MP$, there is a multiindex $\beta\in MC$ such that $y^\alpha=y^\beta$.*

Proof. If $\alpha \notin MC$, we let k denote the smallest index such that $\alpha_k > 2$, and define $\alpha' \in MP$ by

$$\alpha'_i = \begin{cases} \alpha_i & \text{if } i < k, \\ \alpha_k - 2 & \text{if } i = k, \\ \alpha_{k+1} + 1 & \text{if } i = k+1, \\ \alpha_i & \text{if } i > k+1. \end{cases}$$

Because $y_k^3 = y_k y_{k+1}$ we have $y^\alpha = y^{\alpha'}$. Moreover, α' has the same degree as α, and is lexicographically lower than α. There are only finitely many monomials of any given degree, so the claim follows by induction over the lexicographic order. $\square$

Definition 7.11. (a) We put $x_0 = y_0$, and $x_n = y_n + y_{n-1}^2$ for all $n > 0$.

(b) For $n \geq m \geq 0$ we put $x_{[m,n]} = \prod_{i=n}^m x_i$ and $y_{[m,n]} = \prod_{i=n}^m y_i$.

Proposition 7.12. *For all $n \geq 0$ we have $y_n = \sum_{i=0}^n x_{n-i}^{2^i}$ and $y_n x_{n+1} = 0$. Thus, the ring C can also be presented as*

$$C = \mathbb{F}[x_0, x_1, x_2, \dots] \Big/ \Big(x_{n+1} \sum_{i=0}^n x_{n-i}^{2^i} \;\Big|\; n \geq 0 \Big).$$

Proof. Once we recall that $(a+b)^2 = a^2 + b^2 \pmod 2$, the equation $y_n = \sum_{i=0}^n x_{n-i}^{2^i}$ is easily checked by induction. Note that this already holds in the polynomial ring P. As the elements x_i can be expressed in terms of the y_j and vice versa, we see that $P = \mathbb{F}[x_0, x_1, \dots]$. The defining relations $y_n^3 + y_n y_{n+1} = 0$ for C can clearly be rewritten as $y_n x_{n+1} = 0$ and thus as $x_{n+1} \sum_{i=0}^n x_{n-i}^{2^i} = 0$. $\square$

Lemma 7.13. *Whenever $m \leq n$ we have $y_m y_{[m,n]}^2 = y_{[m,n+1]}$.*

Proof. The inductive step is

$$\begin{aligned} y_m y_{[m,n+1]}^2 &= y_m y_{[m,n]}^2 y_{n+1}^2 = y_{[m,n+1]} y_{n+1}^2 \\ &= y_{[m,n]} y_{n+1}^3 = y_{[m,n]} y_{n+1} y_{n+2} = y_{[m,n+2]}. \end{aligned}$$ $\square$

Corollary 7.14. *For $k \geq 0$ we have $y_m^{2^k - 1} = y_{[m,m+k-1]}$.*

Proof. The induction step is

$$y_m^{2^{k+1}-1} = y_m (y_m^{2^k - 1})^2 = y_m y_{[m,m+k-1]}^2 = y_{[m,m+k]}.$$ $\square$

Lemma 7.15. *Fix $m \in \mathbb{N}$, and put*

$$U = \{\alpha \in MC \mid \alpha \text{ is } m\text{-solid and } \alpha_i = 0 \text{ for } i < m\}.$$

Then there is a bijection $\mathbb{N} \to U$ written as $k \mapsto \theta[m,k]$ such that $y^{\theta[m,k]} = y_m^k$ in C.

Proof. First, if $\alpha \in U$ it is clear that $|\alpha|$ is divisible by 2^m, so we can define $\delta\colon U \to \mathbb{N}$ by $\delta(\alpha) = |\alpha|/2^m$.

Now consider $k \in \mathbb{N}$. There is a unique $r \in \mathbb{N}$ such that $2^r - 1 \leq k < 2^{r+1} - 1$. This means that $0 \leq k - (2^r - 1) < 2^r$, so there is a unique set $J \subseteq \{0, 1, \ldots, r-1\}$ with $k - (2^r - 1) = \sum_{j \in J} 2^j$. We put

$$\theta[m,k]_i = \begin{cases} 0 & \text{if } i < m, \\ 1 & \text{if } m \leq i < m+r \text{ and } i - m \notin J, \\ 2 & \text{if } m \leq i < m+r \text{ and } i - m \in J, \\ 0 & \text{if } m + r \leq i. \end{cases}$$

This is clearly in U. Next, we claim that $y^{\theta[m,k]} = y_m^k$. To see this, put $z = y_m^{2^r - 1}$, which is the same as $y_{[m,m+r-1]}$ by Corollary 7.14. We have

$$y^{\theta[m,k]} = y_{[m,m+r-1]} \prod_{j \in J} y_{m+j} = z \prod_{j \in J} y_{m+j},$$
$$y_m^k = y_m^{2^r - 1 + \sum_{j \in J} 2^j} = z \prod_{j \in J} y_m^{2^j}.$$

Now, for $0 \leq j < r$ we have $y_{m+j}(y_{m+j}^2 + y_{m+j+1}) = 0$ and z is divisible by y_{m+j} so $z(y_{m+j}^2 + y_{m+j+1}) = 0$, and so $y_{m+j+1} = y_{m+j}^2$ modulo $\operatorname{ann}(z)$. It follows inductively that $y_{m+j} = y_m^{2^j} \pmod{\operatorname{ann}(z)}$, so $\prod_{j \in J} y_{m+j} = \prod_{j \in J} y_m^{2^j} \pmod{\operatorname{ann}(z)}$, so $y^{\theta[m,k]} = y_m^k$ as claimed. It also follows that $\delta(\theta[m,k]) = |y^{\theta[m,k]}|/2^m = |y_m^k|/2^m = k$.

Now let α be an arbitrary element of U. By the definition of solidity, there is an integer $s \geq 0$ such that when $m \leq i < m+s$ we have $\alpha_i \in \{1, 2\}$ and for $i \geq m+s$ we have $\alpha_i = 0$. It is then clear that

$$\sum_{m \leq i < m+s} 2^i \leq |\alpha| \leq 2 \sum_{m \leq i < m+s} 2^i,$$

or in other words $2^s - 1 \leq \delta(\alpha) < 2^{s+1} - 1$. It follows easily that $\alpha = \theta[m, \delta(\alpha)]$, so we have a bijection as claimed. □

Proposition 7.16. *For* $0 \leq n \leq m \leq \infty$, *the set* $BC[n,m]$ *is a basis for* $C[n,m]$.

Proof. The case $m = \infty$ was covered by Proposition 7.9, so we may assume that $m < \infty$, so $B'C[n,m]$ is a basis for $C[n,m]$ by Proposition 7.5. However, Lemma 7.15 implies that $B'C[n,m]$, considered as a system of elements in $C[n,m]$, is just the same as $BC[n,m]$. □

Proposition 7.17. *Suppose that* $0 \leq n < k \leq m \leq \infty$ *and* $k < \infty$. *Then*

$$\operatorname{ann}_{C[n,m]}(x_k) = C[n,m] y_{k-1}.$$

Proof. The $m = \infty$ case will follow from the $m < \infty$ case, as $C[n, m]_d = C[n, \infty]_d$ when m is large relative to d. We will thus assume that $m < \infty$.

We have already observed that $x_k y_{k-1} = 0$, so $\mathrm{ann}_{C[n,m]}(x_k) \geq C[n, m]y_{k-1}$, and multiplying by x_k gives a well-defined map $f\colon C[n, m]/(C[n, m]y_{k-1}) \to C[n, m]$. It will suffice to show that f is injective.

For this, we put

$$\begin{aligned} N &= \{\alpha \in MC[n, m] \mid \alpha_{k-1} = 0\}, \\ A &= \{y^\alpha \mid \alpha \in N\} \subseteq C[n, m], \\ Z &= \mathrm{span}(A) \leq C[n, m]. \end{aligned}$$

By inspecting the generators and relations on both sides, we see that

$$C[n, m]/(C[n, m]y_{k-1}) = \bar{C}[n, k-1] \otimes C[k, m].$$

Propositions 7.5 and 7.9 show that A also gives a basis for $C[n, m]/(C[n, m]y_{k-1})$, so $C[n, m] = Z \oplus (C[n, m]y_{k-1})$. Now let g denote the composite

$$Z \xrightarrow{\simeq} C[n, m]/(C[n, m]y_{k-1}) \xrightarrow{f} C[n, m] \xrightarrow{\text{proj}} C[n, m]/Z.$$

It will certainly be enough to show that g is injective. It is not hard to see that $y_k Z \leq Z$, and $x_k = y_{k-1}^2 + y_k$, so $g(z) = x_k z + Z = y_{k-1}^2 z + Z$, and so g gives an injective map from A to $BC[n, m] \setminus A$. These sets are bases for the domain and codomain of g, so g is injective as required. □

Proposition 7.18. *C is self-injective.*

Proof. As C is finite in each degree, it will suffice (by Propositions 3.3 and 3.6) to show that every test pair (u, v) in C has either a block or a transporter. Let d be the degree of (u, v), so $|v_i| = |u_i| + d$. Note that some of the entries u_i and v_i may be zero, in which case $|u_i|$ or $|v_i|$ can be negative. Choose m such that $2^m > d$ and also $2^m > |u_i|$ and $2^m > |v_i|$ for all i. Now (u, v) can be regarded as a test pair in $C[n, m]$. Let π be the projection $C[n, m] \to \bar{C}[n, m] = C[n, m]/y_m$. As $\bar{C}[n, m]$ has Poincaré duality, it is self-injective, so the test pair $(\pi(u), \pi(v))$ has either a block or a transporter. First, suppose that there is a transporter $\pi(t)$, so $\pi(v_i) = \pi(tu_i)$ for all i. This is an equation between elements of degree $|v_i| < 2^m$, and $\pi : C[n, m] \to \bar{C}[n, m]$ is an isomorphism in this degree, so $v_i = tu_i$, so we have a transporter for the original pair (u, v).

Suppose instead that there is a block for $(\pi(u), \pi(v))$, say $\pi(b)$. This means that $\pi(b.u) = 0$ but $\pi(b.v) \neq 0$, so $b.u \in C[n, m]y_m$ but $b.v \notin C[n, m]y_m$. Using our bases for the various rings under consideration, we see that $C[n, m]y_m = (Cy_m) \cap C[n, m]$, and thus that $b.v \notin Cy_m$. It now follows from Proposition 7.17 that $(x_{m+1}b).u = 0$ and $(x_{m+1}b).v \neq 0$, so $x_{m+1}b$ is a block for the original pair (u, v). □

We now wish to prove that C is coherent, which turns out to involve substantial work. It will be convenient to regard the set $B\bar{C}[n,m] = \{y^\alpha \mid \alpha \in M\bar{C}[n,m]\}$ as a subset of $C[n,m]$ rather than a subset of $\bar{C}[n,m]$. We write $\tilde{C}[n,m]$ for the span of this set, so the projection $C[n,m] \to \bar{C}[n,m]$ restricts to give an isomorphism $\tilde{C}[n,m] \to \bar{C}[n,m]$.

Lemma 7.19. *For $p \geq 3$ we have*

$$y_{[0,p-3]}^2 y_{[0,p-1]}^2 y_1 y_{p-1} y_p = y_{[0,p]}^2$$

(*and in particular*, *this is nonzero modulo* y_{p+1}).

Proof. Put $A = C[0,p]/\operatorname{ann}(y_{[0,p]})$. We claim that in A we have

$$y_{[0,p-3]}^2 y_{[0,p-1]} y_1 y_{p-1} = y_{[0,p]}.$$

Assuming this, we can just multiply by $y_{[0,p]}$ to recover the statement in the lemma.

For $0 \leq i < p$ we have $y_i(y_i^2 + y_{i+1}) = 0$ so $y_{[0,p]}(y_i^2 + y_{i+1}) = 0$ and so $y_{i+1} = y_i^2$ in A. We thus have $y_k = y_0^{2^k}$ in A for $0 \leq k \leq p$, and so $A = \mathbb{F}[y_0]$. It is thus enough to show that the two sides of the claimed equation have the same degree, which is a straightforward calculation. □

Lemma 7.20. *For any $p \geq 3$ we have*

$$B\bar{C}[0,p-2]\, B\bar{C}[0,p] \subseteq \coprod_{i=0}^{3} B\bar{C}[0,p-1] y_{p-1}^i.$$

Proof. Consider $\alpha \in M\bar{C}[0,p-2]$ and $\beta \in M\bar{C}[0,p]$. We note that

$$y^\alpha, y^\beta \in C[0,p-1]$$

so we can rewrite $y^{\alpha+\beta}$ as an element of the basis $B'C[0,p-1]$, which means $y^{\alpha+\beta} = y^\gamma$ for some $\gamma \in M'C[0,p-1]$. It will be enough to show that $\gamma_{p-1} \leq 3$.

Note that y^α divides $y_{[0,p-3]}^2$ and y^β divides $y_{[0,p-1]}^2$ so y^γ divides $y_{[0,p-3]}^2 y_{[0,p-1]}^2$. It follows using Lemma 7.19 that $y^\gamma y_{p-1} y_p \neq 0 \pmod{y_{p+1}}$. However,

$$y_{p-1}^4 y_{p-1} y_p = y_{p-1}^5 y_p = y_{p-1}^3 y_p^2 = y_{p-1} y_p^3 = y_{p-1} y_p y_{p+1} = 0 \pmod{y_{p+1}},$$

so y^γ cannot be divisible by y_{p-1}^4, as required. □

Definition 7.21. For any vector $u \in C^n$ and $p \geq 0$, we put

$$K(u,p) = \{v \in C[0,p]^n \mid u.v = 0\},$$
$$\bar{K}(u,p) = \{\bar{v} \in \bar{C}[0,p]^n \mid \pi(u).\bar{v} = 0\}.$$

More precisely, $K(u, p)$ is the graded group where

$$K(u, p)_d = \Big\{ v \in C[0, p]^n \;\Big|\; |v_i| = d - |u_i| \text{ for all } i \text{ and } \sum_i u_i v_i = 0 \Big\},$$

and $\overline{K}(u, p)$ is graded in a similar way.

Lemma 7.22. *If $u_i \in \tilde{C}[0, p-2]$ for all i, then the map*

$$\pi : K(u, p+1) \to \overline{K}(u, p+1)$$

is surjective.

Proof. Consider an element $\bar{v} \in \overline{K}(u, p+1)$. This can be written as $\pi(v)$ for a unique element $v \in \tilde{C}[0, p+1]^n$, which must satisfy $u.v = 0 \pmod{y_{p+1}}$. We can write v as $\sum_{k=0}^{2} v_k y_p^k$ with $v_k \in \tilde{C}[0, p]^n$. Using Lemma 7.20 we see that $u.v_k$ can be written as $\sum_{j=0}^{3} w_{jk} y_{p-1}^j$ for some elements $w_{jk} \in \tilde{C}[0, p-1]$. This gives $u.v = \sum_{j=0}^{3} \sum_{k=0}^{2} w_{jk} y_{p-1}^j y_p^k$. After reducing the terms $y_{p-1}^j y_p^k$ using the defining relations for C, we obtain

$$\begin{aligned} u.v = w_{00} + w_{01} y_p + w_{02} y_p^2 + w_{10} y_{p-1} + (w_{11} + w_{30}) y_{p-1} y_p + (w_{12} + w_{31}) y_{p-1} y_p^2 \\ + w_{20} y_{p-1}^2 + w_{21} y_{p-1}^2 y_p + w_{22} y_{p-1}^2 y_p^2 + w_{32} y_{p-1} y_p y_{p+1}. \end{aligned}$$

By hypothesis, this maps to zero in $\bar{C}[0, p+1] = C[0, p+1]/y_{p+1}$. However, $\bar{C}[0, p+1]$ splits as the direct sum of subgroups $\tilde{C}[0, p-1] y_{p-1}^i y_p^j$ for $0 \le i, j < 3$, so we must have

$$w_{00} = w_{01} = w_{02} = w_{10} = w_{20} = w_{21} = w_{22} = 0$$

and $w_{11} = w_{30}$ and $w_{12} = w_{31}$, so $u.v = w_{32} y_{p-1} y_p y_{p+1}$.

Now put $d = |u.v|$, so $|w_{jk}| = d - j2^{p-1} - k2^p$. In particular, we have

$$|w_{32}| = d - 2^{p-1} - 2^p - 2^{p+1}.$$

If $d < 2^{p-1} + 2^p + 2^{p+1}$ then $|w_{32}| < 0$ so $w_{32} = 0$, and so $u.v = 0$. This means that $v \in K(u, p+1)$ with $\pi(v) = \bar{v}$, as required. Suppose instead that $d \ge 2^{p-1} + 2^p + 2^{p+1}$. We have

$$\begin{aligned} |w_{11}| = |w_{30}| &= d - 2^{p-1} - 2^p \ge 2^{p+1}, \\ |w_{12}| = |w_{31}| &= d - 2^{p-1} - 2^{p+1} \ge 2^p. \end{aligned}$$

However, the elements w_{jk} lie in $\tilde{C}[0, p-1]$, which is zero in degrees larger than $2^p - 2$. We therefore have $w_{11} = w_{12} = w_{30} = w_{31} = 0$, which means that $u.v_0 = 0$ and $u.v_1 = 0$ and $u.v_2 = w_{32} y_{p-1}^3 = w_{32} y_{p-1} y_p$. Put

$$v' = v_0 + v_1 y_p + v_2 (y_p^2 + y_{p+1}),$$

so $\pi(v') = \pi(v) = \bar{v}$ and

$$u.v' = u.v_0 + u.v_1 y_p + u.v_2(y_p^2 + y_{p+1}) = w_{32} y_{p-1} y_p (y_p^2 + y_{p+1}) = 0.$$

Thus, v' is the required lift of $\bar{v}$ in $\overline{K}(u, p+1)$. □

Lemma 7.23. *For all $p \geq 0$ we have a splitting*

$$C[0, p+1] = C[0, p] \oplus \bigoplus_{k>0} \bar{C}[0, p] x_{p+1}^k.$$

Proof. By definition, we have $C[0, p+1] = C[0, p][y_{p+1}]/(x_{p+1} y_p)$, where $x_{p+1} = y_{p+1} + y_p^2$ as usual. From this it is clear that

$$C[0, p][y_{p+1}] = C[0, p][x_{p+1}] = C[0, p] \oplus \bigoplus_{k>0} C[0, p] x_{p+1}^k.$$

The ideal generated by $y_p x_{p+1}$ in this ring clearly has a compatible splitting

$$C[0, p][y_{p+1}].y_p x_{p+1} = \bigoplus_{k>0} C[0, p] y_p x_{p+1}^k.$$

We can thus pass to the quotient to get

$$C[0, p+1] = C[0, p] \oplus \bigoplus_{k>0} \frac{C[0, p]}{C[0, p] y_p} x_{p+1}^k = C[0, p] \oplus \bigoplus_{k>0} \bar{C}[0, p] x_{p+1}^k$$

as claimed. □

Corollary 7.24. *If $u_i \in \tilde{C}[0, p-2]$ for $i = 0, \ldots, n-1$, then*

$$K(u, p+1) = C[0, p+1].K(u, p).$$

Proof. It is clear that $C[0, p+1].K(u, p) \leq K(u, p+1)$. For the converse, consider an element $v \in K(u, p+1) \leq C[0, p+1]^n$. Using Lemma 7.23, we can write v as $v_0 + \sum_{k>0} \bar{v}_k x_{p+1}^k$, with $v_0 \in C[0, p]^n$ and $\bar{v}_k \in \bar{C}[0, p]^n$ (with $\bar{v}_k = 0$ for $k \gg 0$). It follows that $u.v_0 \in C[0, p]$ and $u.\bar{v}_k \in \bar{C}[0, p]$ and

$$u.v_0 + \sum_{k>0} (u.\bar{v}_k) x_{p+1}^k = u.v = 0.$$

As the sum in Lemma 7.23 is direct, we must have $u.v_0 = 0$ and $u.\bar{v}_k = 0$, so $v_0 \in K(u, p)$ and $\bar{v}_k \in \overline{K}(u, p)$. By Lemma 7.22, we can choose $v_k \in K(u, p)$ for $k > 0$ lifting $\bar{v}_k$. If $\bar{v}_k = 0$ we choose $v_k = 0$; this ensures that $v_k = 0$ for $k \gg 0$. We now have $v = \sum_{k \geq 0} v_k x_{p+1}^k \in C[0, p+1].K(u, p)$, as required. □

Proposition 7.25. *The ring C is coherent.*

Proof. Let $I \leq C$ be a finitely generated ideal. Choose elements $u_0, \dots, u_{n-1}$ generating I. These give an epimorphism $g\colon \bigoplus_i \Sigma^{|u_i|}C \to I$, with $\ker(g) = K(u, \infty)$, so it will suffice to show that $K(u, \infty)$ is finitely generated as a C-module. Now choose p large enough that $u_i \in \tilde{C}[0, p-2]$ for all i. As $C[0, p]$ is Noetherian, we can choose a finite subset $T \subseteq C[0, p]^n$ that generates $K(u, p)$ as a $C[0, p]$-module. Corollary 7.24 tells us that T also generates $K(u, p+1)$ as a $C[0, p+1]$-module. In fact, we can apply the same corollary inductively to see that T generates $K(u, q)$ as a $C[0, q]$-module for all $q \geq p$. As $C = \bigcup_q C[0, q]$ we conclude that T generates $K(u, \infty)$ as required. □

Proposition 7.26. *The reduced quotient of C is*

$$C/\sqrt{0} = \mathbb{F}[x_i \mid i \geq 0]/(x_i x_j \mid i \neq j).$$

Proof. Put $C' = C/\sqrt{0}$. We first claim that for all p, q with $0 \leq p < q$ we have $x_p x_q = 0$ in C'. We may assume inductively that $x_i x_j = 0$ in C' whenever $0 \leq i < j < q$. By a nested downward induction over p, we may assume that $x_k x_q = 0$ in C' whenever $p < k < q$. As in Proposition 7.12, we have $x_q \sum_{k=0}^{q-1} x_k^{2^{q-1-k}} = 0$. We can multiply this by x_p and use the inner and outer inductive assumptions to see that $x_p x_q x_p^{2^{q-1-p}} = 0$, or in other words $x_p^m x_q = 0$ for some $m > 0$. This gives $(x_p x_q)^m = 0$ in C', but C' is reduced by construction so $x_p x_q = 0$ in C' as claimed.

Now put

$$C'' = C/(x_i x_j \mid i, j,\ i < j) = \mathbb{F}[x_i \mid i \geq 0]/(x_i x_j \mid i, j \geq 0,\ i < j).$$

We now see that C'' is a quotient of C by nilpotent elements, so C' can also be described as $C''/\sqrt{0}$. However, there is an obvious splitting

$$C'' = \mathbb{F} \oplus \bigoplus_{i \geq 0} x_i \mathbb{F}[x_i],$$

and using this we see that C'' is reduced. It follows that $C' = C''$ as claimed. □

8. Pontrjagin self-dual rings

Let R be a Pontrjagin self-dual ring, as in Definition 1.7. Thus, R is a graded $\mathbb{Z}_p$-algebra R equipped with an isomorphism $\zeta : R_d \to \mathbb{Q}_p/\mathbb{Z}_p$ (for some d) such that the resulting maps

$$\zeta^{\#}\colon R_{d-k} \to R_k^{\vee} = \operatorname{Hom}_{\mathbb{Z}_p}(R_k, \mathbb{Q}_p/\mathbb{Z}_p)$$

are isomorphisms.

Lemma 8.1. *For graded R-modules M there is a natural isomorphism*

$$\operatorname{Hom}_R(M, R) \simeq \operatorname{Hom}_{\mathbb{Z}_p}(M_d, \mathbb{Q}_p/\mathbb{Z}_p) = M_d^{\vee}.$$

Proof. Given $\phi \in \mathrm{Hom}_R(M, R)$, we put

$$\tau(\phi) = \zeta \circ \phi_d \colon M_d \to \mathbb{Q}_p/\mathbb{Z}_p.$$

This defines a map $\tau\colon \mathrm{Hom}_R(M, R) \to M_d^\vee$.

Now suppose we have a map $\psi\colon M_d \to \mathbb{Q}_p/\mathbb{Z}_p$. For any $k \in \mathbb{Z}$ we have a map

$$\phi'_k\colon M_k \to \mathrm{Hom}_{\mathbb{Z}_p}(R_{d-k}, \mathbb{Q}_p/\mathbb{Z}_p)$$

given by $\phi'_k(m)(a) = (-1)^{k(d-k)}\psi(am)$. As R is assumed to be Pontrjagin self-dual, there is a unique element $\phi_k(m) \in R_k$ such that

$$\phi'_k(m)(a) = \zeta(\phi_k(m)a)$$

for all $a \in R_{d-k}$. We leave it to the reader to check that this gives a map $\phi\colon M \to R$ of R-modules, and that this is the unique such map with $\tau(\phi) = \psi$. □

Proposition 8.2. *Any Pontrjagin self-dual ring is self-injective.*

Proof. We need to show that the functor $M \mapsto \mathrm{Hom}_R(M, R)$ is exact, but it is isomorphic to the functor $M \mapsto \mathrm{Hom}_{\mathbb{Z}_p}(M_d, \mathbb{Q}_p/\mathbb{Z}_p)$, which is exact because $\mathbb{Q}_p/\mathbb{Z}_p$ is divisible and therefore injective as a $\mathbb{Z}_p$-module. □

We now study the graded ring J described by Definition 1.9, and the tensor product $\hat{J} = \mathbb{Z}_p \otimes J$. It is standard that $\mathbb{Z}_p \otimes \mathbb{Z}/p^r = \mathbb{Z}/p^r$. Moreover, the group $\mathbb{Q}_p/\mathbb{Z}_p$ can be written as the colimit of the evident sequence

$$\mathbb{Z}/p \to \mathbb{Z}/p^2 \to \mathbb{Z}/p^3 \to \cdots,$$

and we can tensor with $\mathbb{Z}_p$ to get $\mathbb{Z}_p \otimes (\mathbb{Q}_p/\mathbb{Z}_p) = \mathbb{Q}_p/\mathbb{Z}_p$. Thus, the only difference between J and $\hat{J}$ is that $J_0 = \mathbb{Z}_{(p)}$ whereas $\hat{J}_0 = \mathbb{Z}_p$.

Lemma 8.3. *The ring $\hat{J}$ is Pontrjagin self-dual, so $\hat{J}_{-2-k} \simeq \hat{J}_k^\vee$.*

Proof. For $k \neq -2$ this is a straightforward calculation. For $k = -2$ we use the description $\mathbb{Q}_p/\mathbb{Z}_p = \varinjlim_j \mathbb{Z}/p^j$ to get

$$\mathrm{Hom}(\mathbb{Q}_p/\mathbb{Z}_p, \mathbb{Q}_p/\mathbb{Z}_p) = \varprojlim_j \mathrm{Hom}(\mathbb{Z}/p^j, \mathbb{Q}_p/\mathbb{Z}_p) = \varprojlim_j \mathbb{Z}/p^j = \mathbb{Z}_p,$$

as required. □

Corollary 8.4. *The ring $\hat{J}$ is self-injective.* □

Remark 8.5. The ring J itself is not self-injective. To see this, note that J_{-2} is an ideal in J and is a module over $\mathbb{Z}_p$. Choose any element $a \in \mathbb{Z}_p \setminus \mathbb{Z}_{(p)}$ and define $u\colon J_{-2} \to J$ by $u(x) = ax$. This cannot be extended to give a J-linear endomorphism of J.

Lemma 8.6. *The ring $\hat{J}$ is local (in the graded sense). The unique maximal graded ideal is given by $\mathfrak{m}_0 = p\mathbb{Z}_p$ and $\mathfrak{m}_k = \hat{J}_k$ for all $k \neq 0$. Moreover, the elements α_k together with the element p give a basis for $\mathfrak{m}/\mathfrak{m}^2$ over $\mathbb{Z}/p$.*

Proof. It is straightforward to check that the graded group $\mathfrak{m}$ described above is an ideal in $\hat{J}$, and the quotient $\hat{J}/\mathfrak{m}$ is the field $\mathbb{Z}/p$, so it is a maximal ideal. Let $\mathfrak{m}'$ be an arbitrary maximal graded ideal. Put $\mathfrak{a} = \bigoplus_{k\neq 0} \hat{J}_k$. Every homogeneous element $a \in \mathfrak{a}$ satisfies $a^2 = 0$, and it follows that $\mathfrak{a} \leq \mathfrak{m}'$ This means that $\mathfrak{m}'$ corresponds to a maximal ideal in the quotient $\hat{J}/\mathfrak{a} \simeq \mathbb{Z}_p$, and the only such ideal is $p\mathbb{Z}_p$. It follows that $\mathfrak{m}' = \mathfrak{m}$ as claimed. The description of $\mathfrak{m}/\mathfrak{m}^2$ is a straightforward calculation. □

Proposition 8.7. *The ring $\hat{J}$ is totally incoherent.*

Proof. Put $V = \{\alpha_k \mid k \neq 0 \pmod{p}\} \subset J$, so V is infinite, pV vanishes and V remains linearly independent in $\mathfrak{m}/\mathfrak{m}^2$. By inspecting the multiplication rules, we see that every noninvertible element of $\hat{J}$ annihilates all elements of V with at most one exception. It follows using Corollary 5.6 that $\hat{J}$ is totally incoherent. □

9. The infinite root algebra

In this section we fix a field K and study the infinite root algebra P over K, which was introduced in Definition 1.12. We first recall the details.

Definition 9.1. We say that a subset $U \subseteq [0, 1]$ is *well-ordered* if the usual order inherited from $\mathbb{R}$ is a well-ordering, so every nonempty subset of U has a smallest element. It is equivalent to say that every infinite nonincreasing sequence in U is eventually constant, or that there are no infinite, strictly decreasing sequences.

An *infinite root series* is a function $a\colon [0, 1] \to K$ such that the set $\mathrm{supp}(a) = \{q \mid a(q) \neq 0\}$ is well-ordered. The infinite root algebra is the set P of all infinite root series. We regard this as an ungraded object, or equivalently as a graded object concentrated in degree zero.

Remark 9.2. It is clear that any subset of a well-ordered set is well-ordered, and that the union of any two well-ordered sets is well-ordered. Now if $a, b \in P$ we have $\mathrm{supp}(a + b) \subseteq \mathrm{supp}(a) \cup \mathrm{supp}(b)$, so P is closed under addition. It is clearly also closed under multiplication by elements of K.

Lemma 9.3. *Any well-ordered subset of $[0, 1]$ is countable. Moreover, for any countable ordinal α, there is a well-ordered subset $U \subseteq [0, 1]$ that is order-isomorphic to α.*

Proof. Firstly, we can regard rational numbers in $[0, 1]$ as coprime pairs of integers and this gives a lexicographic ordering on $\mathbb{Q} \cap [0, 1]$, which is a well-ordering.

Next, let U be a well-ordered subset of $[0, 1]$. We define $f : U \to \mathbb{Q}$ as follows. If u is maximal in U, we put $f(u) = 1$. Otherwise, the set $\{v \in U \mid v > u\}$ has a smallest element v_0, and we define $f(u)$ to be the lexicographically smallest element of $\mathbb{Q} \cap [u, v_0)$. It is clear that f is injective, so U is countable.

Let α be any countable ordinal; we claim that there is an order-embedding $g : \alpha \to [0, 1]$. To see this, choose an injective map $p : \alpha \to \mathbb{N}$ and then put

$$g(\beta) = \sum_{\gamma<\beta} 2^{-p(\gamma)-1}.$$

It is clear that this has the required properties. □

Lemma 9.4. *If* $U, V \subseteq [0, 1]$ *are well-ordered and* $w \in [0, 1]$, *then*

$$\{(u, v) \in U \times V \mid u + v = w\}$$

is finite.

Proof. Put $U' = \{u \in U \mid w - u \in V\}$. This is well-ordered (because it is a subset of U) and it will suffice to show that it is finite. If not, we can define an infinite sequence $u_0 < u_1 < u_2 < \cdots$ in U' as follows: we take u_0 to be the smallest element in U', then take u_1 to be the smallest element in $U' \setminus \{u_0\}$, and so on. We then note that $w - u_0, w - u_1, w - u_2, \ldots$ is an infinite strictly decreasing sequence in V, contradicting the assumption that V is well-ordered. □

Lemma 9.5. *Let* U *be a well-ordered subset of* $[0, 1]$, *and let* (u_n) *be a sequence in* U. *Then there exists an infinite nondecreasing subsequence.*

Proof. Put $v_0 = \min\{u_j \mid j \geq 0\}$ (which is meaningful because U is well-ordered) and then $n_0 = \min\{j \mid u_j = v_0\}$. For $i > 0$, we define recursively

$$v_i = \min\{u_j \mid j > n_{i-1}\} \quad \text{and} \quad n_i = \min\{j > n_{i-1} \mid u_j = v_i\}.$$

We find that $n_0 < n_1 < n_2 < \cdots$ and $v_0 \leq v_1 \leq v_2 \leq \cdots$, or equivalently that $u_{n_0} \leq u_{n_1} \leq u_{n_2} \leq \cdots$ as required. □

Lemma 9.6. *Let* U *and* V *be well-ordered subsets of* $[0, 1]$, *and put*

$$U * V = \{u + v \mid u \in U \text{ and } v \in V\}.$$

Then $U * V$ *is also well-ordered.*

Proof. Suppose not. We can then find an infinite strictly descending chain in $U * V$, so we can choose a sequence (u_n, v_n) in $U \times V$ with $u_i + v_i > u_{i+1} + v_{i+1}$ for all i. Lemma 9.5 tells us that after passing to a subsequence, we may assume that $u_j \leq u_{j+1}$ for all j. After passing again to a sparser subsequence, we may also assume that $v_k \leq v_{k+1}$ for all k. This is clearly impossible. □

Proposition 9.7. *We can make P into a commutative ring by the rule*

$$ab(w) = \sum_{w=u+v} a(u)b(v).$$

Proof. Lemma 9.4 shows that the sum is essentially finite, so there is no problem with convergence. It is clear that $\mathrm{supp}(ab) \subseteq \mathrm{supp}(a) * \mathrm{supp}(b)$, and Lemma 9.6 shows that $\mathrm{supp}(a) * \mathrm{supp}(b)$ is well-ordered, so $ab \in P$. It is straightforward to check that the multiplication operation is commutative, associative and bilinear. Moreover, if we define $e(0) = 1$ and $e(q) = 0$ for $q \neq 0$, then e is a multiplicative identity element for P. $\square$

Definition 9.8. For $a \in P \setminus \{0\}$, we put $\delta(a) = \min(\mathrm{supp}(a))$. We also put $\delta(0) = \infty$.

Remark 9.9. If $\delta(a) + \delta(b) \leq 1$ we have

$$(ab)(\delta(a) + \delta(b)) = a(\delta(a))\, b(\delta(b)) \neq 0,$$

so $ab \neq 0$ and $\delta(ab) = \delta(a) + \delta(b)$. On the other hand, if $\delta(a) + \delta(b) > 1$ then $ab = 0$.

Definition 9.10. For $q \in \mathbb{R} \cup \{\infty\}$ with $q \geq 0$, we define $x^q \in P$ by

$$x^q(u) = \begin{cases} 1 & \text{if } u = q, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 9.11. (a) x^0 is the multiplicative identity element e.

(b) If $q > 1$ then $x^q = 0$.

(c) If $0 \leq q \leq 1$ then $\delta(x^q) = q$.

(d) For all $q, r \geq 0$ we have $x^q x^r = x^{q+r}$.

Lemma 9.12. *Consider an element $a \in P \setminus \{0\}$. If $a(0) = 0$ (or equivalently, $\delta(a) > 0$) then a is nilpotent, but if $\delta(a) = 0$ then a is invertible.*

Proof. If $\delta(a) > 0$ then we can find a positive integer n with $\delta(a) > 1/n$, and using Remark 9.9 we see that $a^n = 0$. Suppose instead that $\delta(a) = 0$. We can then write $a = ue + b = u(e + b/u)$, where $u \in K \setminus 0$ and $e = x^0$ is the multiplicative identity of P and $\delta(b) > 0$, so $b^n = 0$ for some n. Now a has inverse $\sum_{i=0}^{n-1} u^{-1}(-b/u)^i$. $\square$

Corollary 9.13. *The map $a \mapsto a(0)$ induces an isomorphism $P/\sqrt{0} \to K$.*

Proof. Clear. $\square$

Definition 9.14. For $a \in P$ with $\delta(a) \geq t$, we define $\lambda_t(a) \in P$ by

$$\lambda_t(a)(r) = \begin{cases} a(r+t) & \text{if } 0 \leq r \leq 1 - t, \\ 0 & \text{if } 1 - t < r \leq 1. \end{cases}$$

Corollary 9.15. *If $\delta(a) \geq t$ then $a = x^t \lambda_t(a)$ and $\delta(\lambda_t(a)) = \delta(a) - t$. Moreover, if $\delta(a) = t$ then $\lambda_t(a)$ is invertible, so $Pa = Px^t$.*

Proof. The first two claims are clear from the definitions, and the third then follows using Lemma 9.12. □

Definition 9.16. For $t \in [0, 1]$ we put

$$J_t = \{a \in P \mid \delta(a) > t\},$$
$$\bar{J}_t = \{a \in P \mid \delta(a) \geq t\} = Px^t.$$

Proposition 9.17. *Every ideal in P has the form J_t or $\bar{J}_t$.*

Proof. Let I be an ideal in P. If $I = 0$ then $I = J_1$. Otherwise, we put $t = \inf\{\delta(a) \mid a \in I\}$. If $t = \delta(a)$ for some $a \in I$ then Corollary 9.15 shows that $x^t \in I$, and it follows easily that $I = \bar{J}_t$. Suppose instead that there is no element $a \in I$ with $\delta(a) = t$. It is then clear that $I \leq J_t$. Moreover, if $b \in J_t$ then $\delta(b) > t$, so (by the infimum condition) there exists $a \in I$ with $\delta(b) > \delta(a) > t$. After applying Corollary 9.15 to a and b, we see that b is a multiple of a, and so $b \in I$. We now see that $I = J_t$, as required. □

Proposition 9.18. *For all $t \in [0, 1]$ we have $\mathrm{ann}_P(J_t) = \bar{J}_{1-t}$ and $\mathrm{ann}_P(\bar{J}_t) = J_{1-t}$.*

Proof. This follows easily from the fact that $ab = 0$ if and only if $\delta(a) + \delta(b) > 1$. □

Corollary 9.19. *For any ideal $I \leq P$ we have $\mathrm{ann}_P^2(I) = I$.*

Proof. Immediate from the last two propositions. □

Proposition 9.20. *P is self-injective.*

Proof. As we have classified all ideals in P, we can use Baer's criterion. Consider a number $t \in [0, 1]$ and a P-module map $f : \bar{J}_t = (x^t) \to P$. If $f(x^t) = a$ then we must have $J_{1-t}a = f(J_{1-t}x^t) = f(0) = 0$, so $a \in \mathrm{ann}(J_{1-t}) = \bar{J}_t$, so $a = x^t \lambda_t(a)$. We can now define $f' : P \to P$ extending f by $f'(p) = p\,\lambda_t(a)$, so Baer's criterion is satisfied in this case.

Now consider instead a P-module map $f : J_t \to P$. If $t = 1$ then $J_t = 0$ and the zero map $P \to P$ extends f. We suppose instead that $t < 1$. For $s \in (t, 1]$ we put $a_s = \lambda_s(f(x^s))$, so the first case shows that $f(p) = pa_s$ for all $p \in \bar{J}_s < J_t$. Now suppose that $t < r \leq s \leq 1$. As $x^s \in \bar{J}_s \leq \bar{J}_r$ we have $x^s(a_r - a_s) = f(x^s) - f(x^s) = 0$, so $a_r(q) = a_s(q)$ for all $q \leq 1 - s$. Moreover, from the definition of the λ operation we have $a_s(q) = 0$ for $q > 1 - s$, and thus certainly for $q \geq 1 - t$. We now see that there is a unique map $a : [0, 1] \to K$ with $a = a_s$ on $[0, 1 - s]$ (for all $s \in (t, 1]$) and $a = 0$ on $[1 - t, 1]$. It follows easily from these properties that supp(a) is well-ordered, so $a \in P$. We also see from the first property that f agrees with multiplication by a on $\bar{J}_s$ for all $s \in (t, 1]$. It follows that the same is true on $\bigcup_{s \in (t,1]} \bar{J}_s = J_t$, as required. □

Proposition 9.21. *P is totally incoherent.*

Proof. Let I be a finitely generated ideal, say $I = (a_1, \dots, a_r)$, where we can assume that the generators a_i are nonzero. If $r = 0$ then $I = 0$, and this is finitely presented. If $r > 0$ we can use Corollary 9.15 to see that $I = \bar{J}_t$, where $t = \min(\delta(a_1), \dots, \delta(a_r))$.

Now suppose that I is nonzero and finitely presented. We must have $I = \bar{J}_t$ for some t, so we have an epimorphism $g \colon P \to I$ given by $g(a) = ax^t$. Definition 5.1 tells us that $\ker(g)$ must also be finitely generated, but $\ker(g) = \mathrm{ann}_P(x^t) = J_{1-t}$, and this is only finitely generated when $t = 0$ and so $\ker(g) = J_1 = 0$ and $I = \bar{J}_0 = P$. □

Remark 9.22. Put $P' = \{a \in P \mid \mathrm{supp}(a) \subseteq \mathbb{Q}\}$. This is a subring of P, and one can adapt the above arguments to show that it is again self-injective and totally incoherent. Every ideal in P' has the form $J_t \cap P'$ or $\bar{J}_t \cap P'$ for some $t \in [0, 1]$, and these are all distinct except for the fact that $J_t \cap P' = \bar{J}_t \cap P'$ when t is irrational.

10. The Rado algebra

In this section we study the Rado algebra Q, which was defined in Definition 1.16. We will write Γ for the Rado graph.

We first clarify the kinds of graphs that we will consider.

Definition 10.1. A *graph* is a pair (V, E), where V is a set and E is a subset of $V \times V$ such that:

(a) For all $v \in V$ we have $(v, v) \notin E$.

(b) For all $v, w \in V$ we have $((v, w) \in E$ if and only if $(w, v) \in E)$.

Definition 10.2. Let $G = (V, E)$ and $G' = (V', E')$ be graphs. A *full embedding* of G in G' is an injective map $f \colon V \to V'$ such that $E = (f \times f)^{-1}(E')$ (so vertices $v_0, v_1 \in V$ are linked by an edge in G if and only if the images $f(v_0)$ and $f(v_1)$ are linked by an edge in G'). Similarly, a *full subgraph* of G' is a graph of the form $G = G'|_V = (V, E' \cap V^2)$ for some subset $V \subseteq V'$, so the inclusion map gives a full embedding $G \to G'$.

Lemma 10.3. *Suppose we have a finite graph G', a full subgraph G, and a full embedding $f \colon G \to \Gamma$. Then there is a full embedding $f' \colon G' \to \Gamma$ extending f.*

Proof. It is easy to reduce to the case where G' has only one more vertex than G, say $V' = V \amalg \{x\}$. Put $A = \{v \in V \mid (v, x) \in E'\}$ and $N = \max\{f(v) \mid v \in V\} + 1$, then let $f' \colon V' \to \mathbb{N}$ be the map extending f with $f'(x) = 2^N + \sum_{v \in A} 2^{f(v)}$. It is straightforward to check that this has the required properties. □

Remark 10.4. As we mentioned in Example 4.7, each group E_k (for $k \geq 0$) is isomorphic to $\mathbb{F}$. The generator is the element $y_k = x_{B(k)} = \prod_{i \in B(k)} x_i$. We say that a finite subset $I \subseteq \mathbb{N}$ is Γ-*complete* if the full subgraph $\Gamma|_I$ is a complete graph (so every two distinct points are linked by an edge). We say that a natural number n is *$B\Gamma$-complete* if $B(n)$ is Γ-complete. It is clear that the set

$$\{y_n \mid n \text{ is not } B\Gamma\text{-complete}\}$$

is a basis for the Rado ideal, and thus that the set

$$\{y_n \mid n \text{ is } B\Gamma\text{-complete}\}$$

gives a basis for Q.

Proposition 10.5. *For any finitely generated ideal $I \leq Q$, we have* $\operatorname{ann}^2(I) = I$. (*In other words, Q satisfies the double annihilator condition.*)

Proof. Let $I \leq Q$ be a finitely generated ideal. Because of Remark 10.4, the ideal I must be generated by a finite list of monomials, say $I = (x_{A_1}, \dots, x_{A_r})$, where each A_i is a finite Γ-complete subset of $\mathbb{N}$. Similarly, $\operatorname{ann}^2(I)$ is generated by the monomials that it contains.

Let T be another Γ-complete subset of $\mathbb{N}$. If T contains A_i for some i, it is clear that $x_T \in I$. Suppose instead that T does not contain any of the A_i. Let N be strictly larger than any of the elements of $T \cup \bigcup_i A_i$, and put $n = 2^N + \sum_{t \in T} 2^t$, so $B(n) = \{N\} \cup T$. It is clear that $n \notin T$ and $T \cup \{n\}$ is Γ-complete so $x_n x_T \neq 0$. However, we claim that $x_n x_{A_i} = 0$ for all i. Indeed, as $T \not\supseteq A_i$ we can choose $k \in A_i \setminus T$. As N is so large we cannot have $n \in B(k)$, and also $k \notin \{N\} \cup T = B(n)$, so $x_n x_k = 0$, and so $x_n x_{A_i} = 0$ as claimed. We now see that $x_n \in \operatorname{ann}(I)$, but $x_n x_T \neq 0$, so $x_T \notin \operatorname{ann}^2(I)$. It follows that $\operatorname{ann}^2(I) = I$ as claimed. □

Proposition 10.6. *Q is not self-injective.*

Proof. Take any pair $p, q \in \mathbb{N}$ with $p \neq q$ and $x_p x_q = 0$ (say $p = 0$ and $q = 2$). Put $u = (x_p, x_q)$ and $v = (0, x_q)$, and consider the test pair (u, v). Any transporter would have to be an element $t \in Q_0 = \{0, 1\}$ with $t x_p = 0$ and $t x_q = x_q$. It is clear from this that there is no transporter. A block would be a pair (a, b) with $b x_q \neq 0$ but $a x_p + b x_q = 0$ (so $a x_p = b x_q \neq 0$). This means that a and b are nonzero homogeneous elements, say $a = x_A$ and $b = x_B$ for some Γ-complete sets A and B. As $a x_p \neq 0$ we see that $p \notin A$, and that $A \cup \{p\}$ is again Γ-complete. Similarly, we have $q \notin B$ and $B \cup \{q\}$ is Γ-complete. The equation $a x_p = b x_q$ means that $A \cup \{p\} = B \cup \{q\}$, so we have $A = C \cup \{q\}$ and $B = C \cup \{p\}$ for some set C. This now gives $b x_q = x_C x_p x_q$, but $x_p x_q = 0$ so $b x_q = 0$, contrary to assumption. This shows that we have neither a block nor a transporter, so Q is not self-injective. □

Remark 10.7. We could give Q a different grading such that there are some pairs (i, j) with $i \neq j$ but $|x_i| = |x_j|$, so $x_i + x_j$ becomes homogeneous. One can check that if $x_i x_j = 0$ then $\operatorname{ann}^2(x_i + x_j) = (x_i, x_j) \neq (x_i + x_j)$, so the double annihilator condition no longer holds. We will discuss a similar situation with more details in

Lemma 11.18. We believe that the self-injectivity condition is similarly sensitive to the choice of grading, but we do not have an example to prove this.

Proposition 10.8. *Q is totally incoherent.*

Proof. First, it is clear that Q is local, with maximal ideal $\mathfrak{m} = (x_i \mid i \in \mathbb{N}) = \bigoplus_{k>0} Q_k$. The generators x_i form a basis for $\mathfrak{m}/\mathfrak{m}^2$. Note that if $A \subset \mathbb{N}$ is nonempty and Γ-complete, then infinitely many of the variables x_i will satisfy $x_i x_A = 0$, so the image of $\operatorname{ann}(x_A)$ in $\mathfrak{m}/\mathfrak{m}^2$ will have infinite dimension. The claim therefore follows by Corollary 5.6. □

11. The ϵ_0-algebra

The ϵ_0-algebra A was introduced in Definition 1.19. We now explain the definition in more detail, and prove some properties.

Definition 11.1. Suppose we have a sequence $\underline{\beta} = (\beta_1 > \beta_2 > \cdots > \beta_r)$ of ordinals, and a sequence $\underline{n} = (n_1, \ldots, n_r)$ of positive integers. We write

$$C(\underline{\beta}, \underline{n}) = \omega^{\beta_1} n_1 + \cdots + \omega^{\beta_r} n_r.$$

Note that this uses ordinal exponentiation, defined in the usual recursive way by $\alpha^{\beta+1} = \alpha\alpha^{\beta}$ and $\alpha^{\lambda} = \bigcup_{\beta<\lambda} \alpha^{\beta}$ when λ is a limit ordinal.

The following fact is standard (and not hard to prove by transfinite induction).

Proposition 11.2. *For any ordinal α, there is a unique pair $(\underline{\beta}, \underline{n})$ such that $\alpha = C(\underline{\beta}, \underline{n})$. (This is the* Cantor normal form *for α.)*

Proof. See [Johnstone 1987, Exercise 6.10], for example. □

Definition 11.3. We put $\pi_0 = \omega$ and define π_n recursively by $\pi_{n+1} = \omega^{\pi_n}$, and then put $\epsilon_0 = \bigcup_n \pi_n$.

One can check that $\epsilon_0 = \omega^{\epsilon_0}$, and that ϵ_0 is the smallest ordinal with this property. Note that the expression $\epsilon_0 = \omega^{\epsilon_0}$ is the Cantor normal form of ϵ_0. For $\alpha < \epsilon_0$ we find that the exponents β_t in the Cantor normal form of α are strictly less than α, so in this case one can do induction or recursion based on the Cantor normal form.

Definition 11.4. We define $\delta\colon \epsilon_0 \to \mathbb{N}$ recursively by $\delta(0) = 1$, and $\delta(\alpha) = \left(\sum_t (\delta(\beta_t) + 2)n_t\right) - 1$ if $\alpha = \omega^{\beta_1} n_1 + \cdots + \omega^{\beta_r} n_r$.

We will give enough examples to show that δ is not injective, which will be needed later.

Example 11.5.

$$\begin{aligned}
\delta(1) &= \delta(\omega^0) = (\delta(0)+2)-1 = 2, \\
\delta(2) &= \delta(\omega^0\,2) = (\delta(0)+2)2-1 = 5, \\
\delta(\omega) &= \delta(\omega^1) = (\delta(1)+2)-1 = 3, \\
\delta(\omega+1) &= \delta(\omega^1+\omega^0) = (\delta(1)+2)+(\delta(0)+2)-1 = 6, \\
\delta(\omega^2) &= (\delta(2)+2)-1 = 6.
\end{aligned}$$

In order to analyse δ, it is helpful to modify the Cantor normal form slightly.

Lemma 11.6. *If $\alpha < \epsilon_0$ then there is a unique way to write*

$$\alpha = \omega^{\beta_1} + \omega^{\beta_2} + \cdots + \omega^{\beta_m}$$

with $\alpha > \beta_1 \geq \beta_2 \geq \cdots \geq \beta_m$. (This is the expanded Cantor normal form.*)*

Proof. Just take the ordinary Cantor normal form and replace $\omega^{\beta_t} n_t$ by n_t copies of ω^{β_t}. $\square$

Lemma 11.7. *For any $d \in \mathbb{N}$ there are only finitely many ordinals $\alpha \in \epsilon_0$ with $\delta(\alpha) = d$.*

Proof. Let A denote the alphabet $\{0, \pi, +\}$. For each $\alpha < \epsilon_0$ we define a word $\phi(\alpha)$ in A as follows. We start with $\phi(0) = 0$. If $\theta > 0$ has expanded Cantor normal form $\theta = \omega^{\beta_1} + \cdots + \omega^{\beta_m}$ we put

$$\phi(\theta) = \phi(\beta_1)\pi\phi(\beta_2)\pi \cdots \phi(\beta_m)\pi + \cdots +$$

(with $m-1$ plusses at the end). For example we have

$$\begin{aligned}
&\phi(3) = \phi(\omega^0+\omega^0+\omega^0) = 0\pi 0\pi 0\pi{+}{+}, \\
&\phi(\omega^\omega+\omega) = 0\pi\pi\pi 0\pi\pi{+}.
\end{aligned}$$

It is clear from the definitions that $\delta(\theta)$ is the length of $\phi(\theta)$, and there are only 3^d words in A of length d, so it will suffice to show that ϕ is injective. If we interpret π as the operator $x \mapsto \omega^x$ then $\phi(\theta)$ is a reverse Polish expression that evaluates to θ, and this implies injectivity. $\square$

Corollary 11.8. *ϵ_0 is countable.* $\square$

Definition 11.9. Let $\tilde{A}$ be the graded polynomial algebra over $\mathbb{F}$ generated by elements x_α for each ordinal $\alpha < \epsilon_0$, with $|x_\alpha| = \delta(\alpha)$.

Using Lemma 11.7 we see that $\tilde{A}_d$ is finite for all d.

Definition 11.10. For ordinals $\alpha, \beta < \epsilon_0$ with $\alpha \neq \beta$ we define $\mu_0(\alpha, \beta)$ to be the coefficient of ω^β in α. More explicitly, if the Cantor normal form of α involves a term $\omega^\beta n$, then $\mu_0(\alpha, \beta) = n$; if there is no such term then $\mu_0(\alpha, \beta) = 0$. One can check that if $\mu_0(\alpha, \beta) > 0$ then $\mu_0(\beta, \alpha) = 0$. We put $\mu(\alpha, \beta) = \max(\mu_0(\alpha, \beta), \mu_0(\beta, \alpha))$.

Proposition 11.11. *For any finite set $J \subset \epsilon_0$ and map $\nu\colon J \to \mathbb{N}$ there exists $\alpha \in \epsilon_0 \setminus J$ such that $\mu(\alpha, \beta) = \nu(\beta)$ for all $\beta \in J$.* (*We will call this the* extension property.)

Proof. Write J in order as $J = \{\beta_1 > \beta_2 > \cdots > \beta_r\}$ and then take

$$\alpha = \omega^{\beta_1+1} + \omega^{\beta_1}.\nu(\beta_1) + \cdots + \omega^{\beta_r}.\nu(\beta_r).$$

It is visible that $\mu_0(\alpha, \beta_t) = \nu(\beta_t)$ for all t. Also, because of the initial term ω^{β_1+1} we have $\omega^\alpha > \alpha > \beta_t$ for all t and so $\mu_0(\beta_t, \alpha) = 0$. It follows that $\mu(\alpha, \beta_t) = \nu(\beta_t)$ for all t, as required. □

From now on we will only need the fact that our index set ϵ_0 is countable and that the extension property holds. It will therefore be notationally convenient to write $I = \epsilon_0$ and ignore the fact that the elements of I are ordinals, and to write i instead of α for a typical element of I. We also put $I_2 = \{(i, j) \in I^2 \mid i \neq j\}$.

Definition 11.12. For each $(i, j) \in I_2$ we put $\rho(i, j) = x_i x_j^{\mu(i,j)+1}$. We then let A be the quotient of $\tilde{A}$ by all such elements $\rho(i, j)$. We call this the ϵ_0-*algebra*.

Definition 11.13. Given a map $\alpha\colon I \to \mathbb{N}$, we write $\mathrm{supp}(\alpha) = \{i \mid \alpha(i) > 0\}$. Let $M\tilde{A}$ be the set of all such maps α for which $\mathrm{supp}(\alpha)$ is finite. For $\alpha \in M\tilde{A}$ we put $x^\alpha = \prod_i x_i^{\alpha(i)} \in \tilde{A}$. We write $B\tilde{A}$ for the set of all such monomials x^α, so $B\tilde{A}$ is a basis for $\tilde{A}$. Next, put

$$MA = \{\alpha \in M\tilde{A} \mid \forall i \neq j\ \alpha(i) > 0 \Rightarrow \alpha(j) \leq \mu(i, j)\}$$

and $BA = \{x^\alpha \mid \alpha \in MA\}$. One can check that BA gives a basis for A.

Definition 11.14. A *monomial ideal* is just an ideal in A that is generated by some subset of BA.

Remark 11.15. Let P be a monomial ideal, generated by $\{x^\alpha \mid \alpha \in U\}$ for some subset $U \subseteq MA$. Put

$$U^+ = \{\alpha \in MA \mid \alpha \geq \beta \text{ for some } \beta \in U\}.$$

It is easy to see that $\{x^\alpha \mid \alpha \in U^+\}$ is then a basis for P over $\mathbb{F}$. It follows easily that sums, products, intersections and annihilators of monomial ideals are again monomial ideals.

Lemma 11.16. *If P is a monomial ideal then it is finitely generated if and only if there is a finite list of monomials that generate it.*

Proof. Suppose that P is generated by $a_1, \dots, a_m$, where the elements a_t need not be monomials. We can write $a_t = \sum_{\alpha \in U_t} a_{t,\alpha} x^\alpha$, for some finite set $U_t \subset MA$ and some nonzero coefficients $a_{t,\alpha}$. Using Remark 11.15 we see that the terms x^α (for $\alpha \in U_t$) lie in P. Put $U = \bigcup_t U_t$ (which is finite) and put $P' = (x^\alpha \mid \alpha \in U) \leq P$. Clearly $a_t \in (x^\alpha \mid \alpha \in U_t) \leq P'$ and the elements a_t generate P so $P \leq P'$ and so $P = P'$. Thus, P is generated by a finite list of monomials. □

Proposition 11.17. *Let $P \leq A$ be a finitely generated monomial ideal. Then* $\mathrm{ann}^2(P) = P$.

Proof. It is automatic that $P \leq \mathrm{ann}^2(P)$, so it will suffice to prove the opposite inclusion. Note that both P and $\mathrm{ann}^2(P)$ are monomial ideals, so it will suffice to show that they contain the same monomials. Suppose that x^β is a nonzero monomial that does not lie in P; we must find $y \in \mathrm{ann}(P)$ such that $x^\beta y \neq 0$.

We can choose a finite list $\alpha_1, \dots, \alpha_r \in M$ such that $P = (x^{\alpha_1}, \dots, x^{\alpha_r})$. Put $J = \mathrm{supp}(\beta) \cup \bigcup_i \mathrm{supp}(\alpha_i)$, which is a finite subset of I. Put $N = \max\{\beta(j) \mid j \in J\}$.

Next, for each t we note that x^β cannot be divisible by x^{α_t}, so we can choose $i_t \in J$ such that $\alpha_t(i_t) > \beta(i_t)$. Using the extension property we can recursively define distinct elements $k_1, \dots, k_r \in I \setminus J$ such that

(a) $\mu(k_t, i_t) = \alpha_t(i_t) - 1$,

(b) $\mu(k_t, j) = N$ for $j \in J \setminus \{i_t\}$,

(c) $\mu(k_t, k_s) = 1$ for $s < t$.

Put $y = \prod_t x_{k_t}$. This is nonzero by property (c). Property (a) tells us that $x_{j_t} x^{\alpha_t} = 0$ for all t, which implies that $y \in \mathrm{ann}(A)$. On the other hand:

- Clause (a) above tells us that yx^β is not divisible by any relator $\rho(k_t, i_t)$.
- Clause (b) tells us that yx^β is not divisible by any relator $\rho(k_t, j)$ with $j \in J \setminus \{i_t\}$.
- Clause (c) tells us that yx^β is not divisible by any relator $\rho(k_t, k_s)$.
- Our original assumption $x^\beta \neq 0$ implies that yx^β is not divisible by any relator $\rho(j, j')$ with $j, j' \in J$.

This shows that $yx^\beta \neq 0$, but $y \in \mathrm{ann}(P)$, so $x^\beta \notin \mathrm{ann}^2(P)$, as claimed. □

Lemma 11.18. *Let i and j be any two distinct indices in I with $|x_i| = |x_j|$ and $\mu(i, j) = 0$. Then* $\mathrm{ann}^2(x_i + x_j) = (x_i, x_j) > (x_i + x_j)$.

Proof. As $\mu(i, j) = 0$ we have $x_i x_j = 0$ and so (using monomial bases) $(x_i) \cap (x_j) = 0$. If $u(x_i + x_j) = 0$ then we have $ux_i = -ux_j$, with the left hand side in (x_i) and the right hand side in (x_j). As $(x_i) \cap (x_j) = 0$ this gives $ux_i = ux_j = 0$. It now follows that $\operatorname{ann}(x_i+x_j)=\operatorname{ann}(x_i, x_j)$, and so $\operatorname{ann}^2(x_i+x_j)=\operatorname{ann}^2(x_i, x_j)$. As (x_i, x_j) is a monomial ideal we also have $\operatorname{ann}^2(x_i, x_j) = (x_i, x_j)$, so $\operatorname{ann}^2(x_i+x_j) = (x_i, x_j) > (x_i+x_j)$ as claimed. □

Corollary 11.19. *Example 11.5 shows that the lemma applies to the pair* $(\omega^2, \omega+1)$, *so* A *does not satisfy the double annihilator condition. Thus, Remark 2.4 shows that* A *cannot be self-injective.* □

Remark 11.20. We could choose a different grading such that all the generators had different degrees, which would eliminate any examples as in Lemma 11.18. However, we cannot ensure that A_d has dimension at most one for all d, because when $i \neq j$ the elements $x_i^{|x_j|}$ and $x_j^{|x_i|}$ have the same degree and are linearly independent. Thus, there will always be ideals that are not monomial ideals. We suspect that there is no grading for which A satisfies the full double annihilator condition, but we have not proved this.

Proposition 11.21. *A is totally incoherent.*

Proof. Put $\mathfrak{m}_0 = 0$ and $\mathfrak{m}_k = A_k$ for all $k > 0$, so $A/\mathfrak{m} = \mathbb{F}$. It is clear that $\mathfrak{m}$ is an ideal, and that the (homogeneous) elements of $\mathfrak{m}$ are precisely the elements of A that are not invertible. Given this, it follows that $\mathfrak{m}$ is the unique maximal ideal in A, so A is local. From the form of the relations in A we see that $\{x_i \mid i \in I\}$ is a basis for $\mathfrak{m}/\mathfrak{m}^2$.

Now consider an element $a \in A_d$ for some $d > 0$. Put

$$U = \{i \in I \mid \delta(i) \leq d\},$$
$$V = \{\omega^i \mid i \in I \setminus U\}.$$

We find that $x_i x_j = 0$ for all $i \in U$ and $j \in V$. Moreover, we have $a \in (x_i \mid i \in U)$, so $ax_j = 0$ for all $j \in V$, so the image of $\operatorname{ann}(a)$ in $\mathfrak{m}/\mathfrak{m}^2$ has infinite dimension.

Now let P be a finitely presented ideal in A. If $P = \mathfrak{m}P$ then $P = 0$ by Nakayama's lemma. Otherwise, we can choose $a \in P \setminus \mathfrak{m}P$, and Lemma 5.5 tells us that $\operatorname{ann}(a)$ has finite image in $\mathfrak{m}/\mathfrak{m}^2$. The above remarks show that we must have $|a| = 0$, and $a \notin \mathfrak{m}P$ so $a \neq 0$. Thus a is invertible, so $P = A$. □

Proposition 11.22. *The reduced quotient is*

$$A/\sqrt{0} = \mathbb{F}[x_i \mid i \in I]/(x_i x_j \mid i \neq j).$$

Proof. In A we have $x_i x_j^{\mu(i,j)+1} = 0$, so $(x_i x_j)^{\mu(i,j)+1} = 0$, and so $x_i x_j$ is nilpotent. If we put

$$A' = A/(x_i x_j \mid i \neq j) = \mathbb{F}[x_i \mid i \in I]/(x_i x_j \mid i \neq j),$$

we deduce that $A/\sqrt{0} = A'/\sqrt{0}$. However, it is easy to see that A' is already reduced, so $A/\sqrt{0} = A'$ as claimed. □

12. Triangulation

Recall that a *triangulated category* is a triple $(\mathscr{C}, \Sigma, \Delta)$, where $\mathscr{C}$ is an additive category, and $\Sigma\colon \mathscr{C} \to \mathscr{C}$ is an equivalence, and Δ is a class of diagrams of shape

$$X \to Y \to Z \to \Sigma X$$

(called *distinguished triangles*), subject to certain axioms that we will not list here.

Definition 12.1. Let R be a self-injective graded ring, let Mod_R be the category of R-modules, and let $\Sigma\colon\ \mathrm{Mod}_R \to \mathrm{Mod}_R$ be the usual suspension functor so that $(\Sigma M)_i = M_{i-1}$. Let InjMod_R be the full subcategory of injective modules. A *triangulation structure* for R is a pair $(\mathscr{N}, \Delta)$, where:

(a) $\mathscr{N}$ is a full subcategory of InjMod_R containing R.

(b) $\mathscr{N}$ is closed under finite direct sums, retracts, suspensions and desuspensions.

(c) Δ is a class of distinguished triangles making $(\mathscr{N}, \Sigma, \Delta)$ into a triangulated category.

We can also make a similar definition for ungraded rings.

Definition 12.2. Let R be a self-injective ungraded ring. An *ungraded triangulation structure* for R is a pair $(\mathscr{N}, \Delta)$, where:

(a) $\mathscr{N}$ is a full subcategory of InjMod_R containing R.

(b) $\mathscr{N}$ is closed under finite direct sums, retracts, suspensions and desuspensions.

(c) Δ is a class of distinguished triangles making $(\mathscr{N}, 1, \Delta)$ into a triangulated category.

In [Muro et al. 2007] we constructed ungraded triangulation structures for $\mathbb{Z}/4$ and for $K[\epsilon]/\epsilon^2$ (where K is any field of characteristic two). If Freyd's generating hypothesis is true, then the image of the functor π_* gives a graded triangulation structure for the ring $\pi_*(S)^\wedge_p$. We have not succeeded in constructing any examples of graded triangulation structures by pure algebra. Here we offer only some rather limited and negative results.

Lemma 12.3. *If $(\mathscr{N}, \Delta)$ is a triangulation structure (in the graded or ungraded context) then all distinguished triangles in Δ are exact sequences.*

Proof. The general theory of triangulated categories tells us that all functors of the form $\mathscr{N}(X, -)$ send distinguished triangles to long exact sequences. By assumption we have $R \in \mathscr{N}$, and we can take $X = R$ to prove the claim. $\square$

Lemma 12.4. *If $(\mathscr{N}, \Delta)$ is a triangulation structure then all surjective maps in $\mathscr{N}$ are split.*

Proof. Let $M \xrightarrow{f} N$ be a surjective map in $\mathcal{N}$. This must fit into a distinguished triangle $L \xrightarrow{e} M \xrightarrow{f} N \xrightarrow{g} \Sigma L$. Here $gf = 0$, but f is surjective so $g = 0$. It is standard that the functor $\mathcal{N}(N, -)$ converts our distinguished triangle to an exact sequence, so $f_*\colon \mathcal{N}(N, M) \to \mathcal{N}(N, N)$ is surjective. We can thus find $h\colon N \to M$ with $fh = 1$, so h splits f. □

Corollary 12.5. *If* $(\mathcal{N}, \Delta)$ *is a triangulation structure then all finitely generated modules in* $\mathcal{N}$ *are projective. Thus, if* R *is local then all such modules are free.*

Proof. Let N be a finitely generated module in $\mathcal{N}$. This means that there is a surjective homomorphism $f\colon F \to N$ for some finitely generated free module F. As $\mathcal{N}$ is standard we see that $F \in \mathcal{N}$, so the lemma tells us that N is a retract of F, so it is projective. It is well-known that finitely generated projective modules over local rings are free. □

Proposition 12.6. *Suppose that* R *is a local graded ring with* $R_i = 0$ *for* $i < 0$, *and suppose that* R *admits a triangulation structure. Then* R *is totally incoherent.*

Proof. Let $\mathfrak{m}$ be the unique maximal ideal, and let $(\mathcal{N}, \Delta)$ be a triangulation structure. It is not hard to see that $\mathfrak{m}_0$ is the unique maximal ideal in R_0, so R_0 is a local ring in the ungraded sense.

Let J be any finitely generated ideal. We can then find a finitely generated free module Q and an epimorphism $Q \to J$ such that $Q/\mathfrak{m}Q \to J/\mathfrak{m}J$ is an isomorphism. We will write g for the composite map $Q \to J \to R$, so that $J = \mathrm{image}(g)$. If J is finitely presented then $\ker(g)$ is again finitely generated, so we can find a finitely generated free module P and a map $f\colon P \to Q$ with $\mathrm{image}(f) = \ker(g)$ and $P/\mathfrak{m}P \xrightarrow{\simeq} \ker(g)/\mathfrak{m}\ker(g)$. With these minimal choices for P and Q, it is clear that $P_i = Q_i = 0$ when $i < 0$. Next, we can fit g into a distinguished triangle $\Sigma^{-1}R \xrightarrow{d} K \xrightarrow{i} Q \xrightarrow{g} R$. As $gf = 0$, we can find a lift $\tilde{f}\colon P \to K$ with $i\tilde{f} = f$. We can combine this with d to give a map $P \oplus \Sigma^{-1}R \to K$, and a diagram chase shows that this is surjective. Using Lemma 12.4 we deduce that this map is split epi and that K is a finitely generated free module. It follows that $K_i = 0$ for $i < -1$ and that K_{-1} is a retract of R_0. As R_0 is local we must have either $K_{-1} = 0$ or $K_{-1} = R_0$. If $K_{-1} = 0$ then $d\colon \Sigma^{-1}R \to K$ must be zero, which implies that $g\colon Q \to R$ is split epi, which means that $J = R$. If $K_{-1} \neq 0$ then we find that d must induce a monomorphism $\Sigma^{-1}R/\mathfrak{m} \to K$, and as R is local this implies that d is a split monomorphism, and thus that $g = 0$ and so $J = 0$. □

Remark 12.7. As mentioned previously, there is an ungraded triangulation structure for the ring $\mathbb{Z}/4$. The ideal $(2) < \mathbb{Z}/4$ is finitely presented and is neither 0 nor $\mathbb{Z}/4$. It follows that our grading assumptions are playing an essential role in the proof of the above proposition.

Corollary 12.8. *Neither the infinite exterior algebra* (*as in Example 4.7*) *nor the cube algebra* (*as in Section 7*) *admits a triangulation structure.*

Proof. Both rings are coherent, by Propositions 5.4 and 7.25. □

References

[Baer 1940] R. Baer, "Abelian groups that are direct summands of every containing abelian group", *Bull. Amer. Math. Soc.* **46** (1940), 800–806. MR 2,126i Zbl 0024.14902

[Benson et al. 2007] D. J. Benson, S. K. Chebolu, J. D. Christensen, and J. Mináč, "The generating hypothesis for the stable module category of a p-group", *J. Algebra* **310**:1 (2007), 428–433. MR 2007k:16011 Zbl 1120.20002

[Bruns and Herzog 1993] W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics **39**, Cambridge University Press, 1993. MR 95h:13020 Zbl 0788.13005

[Cameron 2001] P. J. Cameron, "The random graph revisited", pp. 267–274 in *European Congress of Mathematics* (Barcelona, 2000), vol. 1, edited by C. Casacuberta et al., Progr. Math. **201**, Birkhäuser, Basel, 2001. MR 2003j:05110 Zbl 1023.05124

[Carlson et al. 2009] J. F. Carlson, S. K. Chebolu, and J. Mináč, "Freyd's generating hypothesis with almost split sequences", *Proc. Amer. Math. Soc.* **137**:8 (2009), 2575–2580. MR 2010c:20007 Zbl 1188.20006

[Devinatz 1990] E. S. Devinatz, "K-theory and the generating hypothesis", *Amer. J. Math.* **112**:5 (1990), 787–804. MR 91i:55011 Zbl 0783.55008

[Freyd 1966] P. Freyd, "Stable homotopy", pp. 121–172 in *Proceedings of the Conference on Categorical Algebra* (La Jolla, CA, 1965), edited by S. Eilenberg et al., Springer, New York, 1966. MR 35 #2280 Zbl 0195.52901

[Hovey 2007] M. Hovey, "On Freyd's generating hypothesis", *Q. J. Math.* **58**:1 (2007), 31–45. MR 2008b:55014 Zbl 1173.55007

[Hovey et al. 2007] M. Hovey, K. Lockridge, and G. Puninski, "The generating hypothesis in the derived category of a ring", *Math. Z.* **256**:4 (2007), 789–800. MR 2008b:18017 Zbl 1137.16016

[Johnstone 1987] P. T. Johnstone, *Notes on logic and set theory*, Cambridge University Press, 1987. MR 89e:03003 Zbl 0642.03001

[Lam 1999] T. Y. Lam, *Lectures on modules and rings*, Graduate Texts in Mathematics **189**, Springer, New York, 1999. MR 99i:16001 Zbl 0911.16001

[Lockridge 2007] K. H. Lockridge, "The generating hypothesis in the derived category of R-modules", *J. Pure Appl. Algebra* **208**:2 (2007), 485–495. MR 2007i:55009 Zbl 1112.55011

[Matsumura 1980] H. Matsumura, *Commutative ring theory*, Kyoritsu Shuppan, Tokyo, 1980. In Japanese; translated in Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, 1986. MR 88h:13001 Zbl 0603.13001

[Muro et al. 2007] F. Muro, S. Schwede, and N. Strickland, "Triangulated categories without models", *Invent. Math.* **170**:2 (2007), 231–241. MR 2008g:18016 Zbl 1125.18009

[Rado 1964] R. Rado, "Universal graphs and universal functions", *Acta Arith.* **9** (1964), 331–340. MR 30 #2488 Zbl 0139.17303

[Ravenel 1984] D. C. Ravenel, "Localization with respect to certain periodic homology theories", *Amer. J. Math.* **106**:2 (1984), 351–414. MR 85k:55009 Zbl 0586.55003

[Ravenel 1992] D. C. Ravenel, *Nilpotence and periodicity in stable homotopy theory*, Annals of Mathematics Studies **128**, Princeton University Press, 1992. MR 94b:55015 Zbl 0774.55001

Communicated by David Benson
Received 2012-06-06 Revised 2013-06-29 Accepted 2013-07-30

lshepperson@rm.com *Dorset Software, 10 the Linnets, Cottenham, Cambridge CB24 8XZ, United Kingdom*

n.p.strickland@sheffield.ac.uk *School of Mathematics and Statistics, University of Sheffield, The Hicks Building, Hounsfield Road, Sheffield S3 7RH, United Kingdom*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.303

On lower ramification subgroups and canonical subgroups

Shin Hattori

Let p be a rational prime, k be a perfect field of characteristic p and K be a finite totally ramified extension of the fraction field of the Witt ring of k. Let $\mathcal{G}$ be a finite flat commutative group scheme over $\mathcal{O}_K$ killed by some p-power. In this paper, we prove a description of ramification subgroups of $\mathcal{G}$ via the Breuil–Kisin classification, generalizing the author's previous result on the case where $\mathcal{G}$ is killed by $p \geq 3$. As an application, we also prove that the higher canonical subgroup of a level n truncated Barsotti–Tate group $\mathcal{G}$ over $\mathcal{O}_K$ coincides with lower ramification subgroups of $\mathcal{G}$ if the Hodge height of $\mathcal{G}$ is less than $(p-1)/p^n$, and the existence of a family of higher canonical subgroups improving a previous result of the author.

1. Introduction

Let p be a rational prime, k be a perfect field of characteristic p and $W = W(k)$ be the Witt ring of k. The natural Frobenius endomorphism of the ring W lifting the p-th power Frobenius of k is denoted by φ. Let K be a finite extension of $K_0 = \mathrm{Frac}(W)$ with integer ring $\mathcal{O}_K$, uniformizer π and absolute ramification index e. We fix an algebraic closure $\bar{K}$ of K and extend the valuation v_p of K satisfying $v_p(p) = 1$ to $\bar{K}$. Let $\hat{\mathcal{O}}_{\bar{K}}$ be the completion of the integer ring $\mathcal{O}_{\bar{K}}$. We also fix a system $\{\pi_n\}_{n\geq 0}$ of p-power roots of π in $\bar{K}$ satisfying $\pi_0 = \pi$ and $\pi_{n+1}^p = \pi_n$ and put $K_\infty = \bigcup_n K(\pi_n)$. The absolute Galois groups of K and K_∞ are denoted by G_K and G_{K_∞}, respectively. For any positive rational number i, put $m_K^{\geqslant i} = \{x \in \mathcal{O}_K \mid v_p(x) \geq i\}$ and $\mathcal{O}_{K,i} = \mathcal{O}_K/m_K^{\geqslant i}$. For any valuation ring V of height one, we define $m_V^{\geqslant i}$ and V_i similarly. We also put $\mathcal{S}_i = \mathrm{Spec}(\mathcal{O}_{K,i})$, $\mathcal{S}_{L,i} = \mathrm{Spec}(\mathcal{O}_{L,i})$ for any finite extension L/K, and $\bar{\mathcal{S}}_i = \mathrm{Spec}(\mathcal{O}_{\bar{K},i})$.

Breuil conjectured a classification of finite flat (commutative) group schemes over $\mathcal{O}_K$ killed by some p-power via φ-modules over the formal power series ring $\mathfrak{S} = W[[u]]$ and obtained such a classification for the case where groups are killed

MSC2010: primary 11S23; secondary 14L05, 14L15.
Keywords: finite flat group scheme, Breuil–Kisin module, canonical subgroup.

by $p \geq 3$ [Breuil 2002]. It is often referred to as the Breuil–Kisin classification, since Kisin showed the conjecture for $p \geq 3$ [Kisin 2006] and for the case where $p = 2$ and groups are connected [Kisin 2009]. The conjecture was proved for any p independently in [Kim 2012; Lau 2010; Liu 2013]. In particular, we have an exact category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$ of such φ-modules over $\mathfrak{S}$ killed by some p-power (for the definition, see Section 2) and an anti-equivalence of exact categories $\mathfrak{M}^*(-)$ from the category of finite flat group schemes over $\mathcal{O}_K$ killed by some p-power to the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$. Moreover, we can recover the G_{K_∞}-module $\mathcal{G}(\mathcal{O}_{\bar{K}})$ via this classification: Let R be the valuation ring defined as the projective limit of p-th power maps

$$R = \varprojlim(\mathcal{O}_{\bar{K},1} \leftarrow \mathcal{O}_{\bar{K},1} \leftarrow \cdots)$$

and $\underline{\pi}$ be the element of the ring R defined by $\underline{\pi} = (\pi_0, \pi_1, \ldots)$. We normalize the valuation v_R by $v_R(\underline{\pi}) = 1/e$ and define R_i similarly to $\mathcal{O}_{K,i}$, using v_R in place of v_p. For any positive integer n, let $W_n(R)$ be the Witt ring of length n of R, which is considered as an $\mathfrak{S}$-algebra by the map $u \mapsto [\underline{\pi}]$. The ring $W_n(R)$ admits a natural G_K-action. Then, by the Breuil–Kisin classification, we also have an isomorphism of G_{K_∞}-modules

$$\varepsilon_{\mathcal{G}} : \mathcal{G}(\mathcal{O}_{\bar{K}}) \to T^*_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G})) = \mathrm{Hom}_{\mathfrak{S},\varphi}\big(\mathfrak{M}^*(\mathcal{G}), W_n(R)\big).$$

On the other hand, for any positive rational number i, we have a finite flat closed subgroup scheme $\mathcal{G}_i$ of $\mathcal{G}$ over $\mathcal{O}_K$, the i-th lower ramification subgroup of $\mathcal{G}$, whose index is adapted to the valuation v_p. Namely, it is defined as the unique finite flat closed subgroup scheme of $\mathcal{G}$ over $\mathcal{O}_K$ satisfying

$$\mathcal{G}_i(\mathcal{O}_{\bar{K}}) = \mathrm{Ker}\big(\mathcal{G}(\mathcal{O}_{\bar{K}}) \to \mathcal{G}(\mathcal{O}_{\bar{K},i})\big).$$

The lower ramification subgroups, which are named as such because of their similarity to the lower numbering ramification groups in algebraic number theory, have similar properties to the upper ramification subgroups [Abbes and Mokrane 2004, §2.3] such as the functoriality and the compatibility with base extension. While this upper variant is used to construct canonical subgroups of abelian varieties [Abbes and Mokrane 2004], the lower ramification subgroups have been also studied and used to construct canonical subgroups [Hattori 2013; 2014; Rabinoff 2012], as explained later.

If $\mathcal{G}$ is killed by $p \geq 3$, then [Hattori 2012, Theorem 1.1] shows that the isomorphism $\varepsilon_{\mathcal{G}}$ induces an isomorphism

$$\mathcal{G}_i(\mathcal{O}_{\bar{K}}) \simeq \mathrm{Ker}\big(T^*_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G})) \to \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}^*(\mathcal{G}), R_i)\big)$$

for any i. This description of the lower ramification subgroups of $\mathcal{G}$ via the Breuil–Kisin classification is used in [Hattori 2013] to deduce various properties of canonical subgroups. In this paper, we prove the following theorem, which generalizes this description.

Theorem 1.1. *Let i be a positive rational number satisfying $i \leq 1$ and $W_n^{\mathrm{DP}}(R)_i$ be the divided power envelope of the natural surjection*

$$W_n(R) \to \mathscr{O}_{\bar{K},i}, \quad (r_0, \dots, r_{n-1}) \mapsto \mathrm{pr}_0(r_0) \bmod m_{\bar{K}}^{\geqslant i}.$$

Let $I_{n,i}$ be the kernel of the map $W_n(R) \xrightarrow{\varphi} W_n^{\mathrm{DP}}(R)_i$ induced by the Frobenius map

$$\varphi : (r_0, \dots, r_{n-1}) \mapsto (r_0^p, \dots, r_{n-1}^p).$$

Let $\mathscr{G}$ be a finite flat group scheme over $\mathscr{O}_K$ killed by p^n and $\mathfrak{M} = \mathfrak{M}^*(\mathscr{G})$ be the corresponding object of the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$. Then the natural isomorphism

$$\varepsilon_{\mathscr{G}} : \mathscr{G}(\mathscr{O}_{\bar{K}}) \to T^*_{\mathfrak{S}}(\mathfrak{M}) = \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, W_n(R))$$

induces an isomorphism

$$\mathscr{G}_i(\mathscr{O}_{\bar{K}}) \simeq \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, I_{n,i}).$$

For the case of $n = 1$, Theorem 1.1 can be interpreted as a correspondence of both upper and lower ramification between $\mathscr{G}$ and a finite flat group scheme $\mathscr{H}(\mathfrak{M}^*(\mathscr{G}))$ over $k[\![u]\!]$ (Corollary 3.3), generalizing [Hattori 2012, Theorem 1.1]. Indeed, by a theorem of Tian and Fargues, Theorem 3.3 of [Hattori 2012], and the compatibility of the Breuil–Kisin classification with Cartier duality, Theorem 1.1 for $n = 1$ also implies the assertion of the corollary on upper ramification subgroups. However, the author does not know if a description of upper ramification subgroups via the Breuil–Kisin classification for $n > 1$ can be obtained from Theorem 1.1, since we do not have a comparison result between upper and lower ramification subgroups similar to the theorem of Tian and Fargues for $n > 1$.

In [Hattori 2012], the proof of Theorem 1.1 for the case where $\mathscr{G}$ is killed by $p \geq 3$ is reduced to showing a congruence of the defining equations of $\mathscr{G}$ and $\mathscr{H}(\mathfrak{M}^*(\mathscr{G}))$ with respect to the identification $k[\![u]\!]/(u^e) \simeq \mathscr{O}_{K,1}$ sending u to π. This congruence is a consequence of an explicit description of the affine algebra of $\mathscr{G}$ in terms of $\mathfrak{M}^*(\mathscr{G})$ due to Breuil [2000, Proposition 3.1.2], which is known only for the case where $\mathscr{G}$ is killed by $p \geq 3$. Here, instead, we study a relationship between the groups

$$\mathscr{G}(\mathscr{O}_{\bar{K},i}) \text{ and } \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}^*(\mathscr{G}), W_n(R)/I_{n,i})$$

by using the faithfulness of the crystalline Dieudonné functor [de Jong and Messing 1999], from which Theorem 1.1 follows easily.

As an application of Theorem 1.1 and an explicit description of the ideal $I_{n,i}$ (Lemma 4.3), we also prove the coincidence with canonical subgroups with lower ramification subgroups, and the existence of a family of canonical subgroups improving Corollary 1.2 of [Hattori 2014]. Before stating the results, we briefly explain a background of this application.

Let $K/\mathbb{Q}_p$ be an extension of complete discrete valuation fields, $\mathfrak{X}$ be an admissible formal scheme over $\mathrm{Spf}(\mathcal{O}_K)$ and $\mathfrak{G}$ be a truncated Barsotti–Tate group of level n over $\mathfrak{X}$. Consider their Raynaud generic fibers X and G. For any point $x \in X$, the fiber $\mathfrak{G}_x$ is a truncated Barsotti–Tate group of level n over the ring of integers of a finite extension of K. If $\mathfrak{G}_x$ is ordinary, then the unit component $\mathfrak{G}_x^0$ satisfies $\mathfrak{G}_x^0(\mathcal{O}_{\bar{K}}) \simeq (\mathbb{Z}/p^n\mathbb{Z})^{\dim \mathfrak{G}_x}$ and its special fiber is equal to the Frobenius kernel of the special fiber of $\mathfrak{G}_x$. We refer to a finite flat closed subgroup scheme of $\mathfrak{G}_x$ as a canonical subgroup if it has these properties. What we want to construct here is a family of canonical subgroups for G: namely, an admissible open subgroup C of G over a strict neighborhood U of the ordinary locus $X^{\mathrm{ord}} \subseteq X$ for $\mathfrak{G}$ such that for any $x \in U$, the fiber C_x is the generic fiber of a canonical subgroup of $\mathfrak{G}_x$. The existence of a family of canonical subgroups is one of the key ingredients in the theory of p-adic Siegel modular forms, and for such arithmetic applications, we also need a precise understanding of C_x. This leads us to construct such a family by first constructing and studying a canonical subgroup of $\mathfrak{G}_x$ fiberwise, and then patching them into a family.

For each fiber $\mathfrak{G}_x$, the method of lifting the conjugate Hodge filtration to the Breuil–Kisin module [Hattori 2013; 2014] gives a sharp result on the existence of a canonical subgroup of $\mathfrak{G}_x$, which is stronger than other methods such as the one using the Hodge–Tate map. Namely, it shows that a canonical subgroup $\mathcal{C}_n$ of $\mathfrak{G}_x$ exists if the Hodge height of $\mathfrak{G}_x$ is less than $1/(p^{n-2}(p+1))$ and $\mathcal{C}_n$ has various properties needed for arithmetic applications.

To obtain a family of canonical subgroups (from any of such fiberwise constructions), we typically need to show the coincidence of canonical subgroups with a specific series of subgroups of $\mathfrak{G}_x$ which can be patched into a family when varying x, and this step often requires us to restrict to a smaller admissible open subset than the locus of x such that a canonical subgroup of $\mathfrak{G}_x$ exists. We have at least three series of such subgroups: Harder–Narasimhan filtrations, upper ramification subgroups and lower ramification subgroups, where the former two were mainly used in preceding works; see [Abbes and Mokrane 2004, Fargues 2011, Hattori 2013; 2014, Tian 2010; 2012].

For $n = 1$, the canonical subgroup $\mathcal{C}_1$ constructed in [Hattori 2013; 2014] was shown to coincide with both upper and lower ramification subgroups, and this again gives a sharp result, namely the existence of a family of canonical subgroups over the locus of Hodge height less than $p/(p+1)$. For $n \geq 2$, it was also shown that $\mathcal{C}_n$ coincides with upper ramification subgroups under a condition on the Hodge height, and this yields a family over the locus of Hodge height less than $1/(2p^{n-1})$ [Hattori

2013; 2014] A weaker result can be obtained also by the Harder–Narasimhan method [Fargues 2011].

In this paper, to obtain a stronger existence theorem of a family of canonical subgroups, we also prove the coincidence of the canonical subgroup constructed in [Hattori 2013; 2014] with lower ramification subgroups, as follows.

Theorem 1.2. *Let $K/\mathbb{Q}_p$ be an extension of complete discrete valuation fields. Let $\mathcal{G}$ be a truncated Barsotti–Tate group of level n, height h and dimension d over $\mathcal{O}_K$ with $0 < d < h$ and Hodge height $w < (p-1)/p^n$. Then the level n canonical subgroup $\mathcal{C}_n$ of $\mathcal{G}$* [Hattori 2014, Theorem 1.1] *satisfies $\mathcal{C}_n = \mathcal{G}_{i_n} = \mathcal{G}_{i'_n}$ for*

$$i_n = \frac{1}{p^{n-1}(p-1)} - \frac{w}{p-1}, \quad i'_n = \frac{1}{p^n(p-1)}.$$

Note that by our assumption and [Hattori 2014, Theorem 1.1], we have an isomorphism of groups

$$\mathcal{C}_n(\mathcal{O}_{\bar{K}}) \simeq (\mathbb{Z}/p^n\mathbb{Z})^d.$$

The fact that the lower ramification subgroup $\mathcal{G}_{i_n}(\mathcal{O}_{\bar{K}})$ is isomorphic to $(\mathbb{Z}/p^n\mathbb{Z})^d$ for $w < (p-1)/p^n$ was proved by Rabinoff [2012, Theorem 1.9] for the case where $K/\mathbb{Q}_p$ is an extension of (not necessarily discrete) complete valuation fields of height one, by a different method. Theorem 1.2 reproves this result of Rabinoff for the case where the base field K is a complete discrete valuation field, and also shows that the subgroup considered by Rabinoff coincides with $\mathcal{C}_n$. In particular, we show that his subgroup has standard properties as a canonical subgroup as in [Hattori 2014, Theorem 1.1], such as the coincidence with a lift of the Frobenius kernel.

Using Theorem 1.2, we also prove the following theorem on a family construction of canonical subgroups, which is stronger than [Hattori 2014, Corollary 1.2] for $n \geq 2$.

Theorem 1.3. *Let $K/\mathbb{Q}_p$ be an extension of complete discrete valuation fields. Let $\mathfrak{X}$ be an admissible formal scheme over $\mathrm{Spf}(\mathcal{O}_K)$ and $\mathfrak{G}$ be a truncated Barsotti–Tate group of level n over $\mathfrak{X}$ of constant height h and dimension d with $0 < d < h$. We let X and G denote the Raynaud generic fibers of the formal schemes $\mathfrak{X}$ and $\mathfrak{G}$, respectively. Put $r_n = (p-1)/p^n$ and let $X(r_n)$ be the admissible open subset of X defined by*

$$X(r_n)(\bar{K}) = \{x \in X(\bar{K}) \mid \mathrm{Hdg}(\mathfrak{G}_x) < r_n\}.$$

Then there exists an admissible open subgroup C_n of $G|_{X(r_n)}$ over $X(r_n)$ such that, etale locally on $X(r_n)$, the rigid-analytic group C_n is isomorphic to the constant group $(\mathbb{Z}/p^n\mathbb{Z})^d$ and, for any finite extension L/K and $x \in X(L)$, the fiber $(C_n)_x$ coincides with the generic fiber of the level n canonical subgroup of $\mathfrak{G}_x$.

2. The Breuil–Kisin classification

In this section, we briefly recall the classification of finite flat group schemes and Barsotti–Tate groups over $\mathcal{O}_K$ due to Kisin ([2006] for $p \geq 3$ and [2009] for $p = 2$ and connected group schemes) and to Kim [2012], Lau [2010] and Liu [2013] for $p = 2$. We basically follow the presentation of [Kim 2012].

We let the continuous φ-semilinear endomorphism of $\mathfrak{S}$ defined by $u \mapsto u^p$ be denoted also by φ. Put $\mathfrak{S}_n = \mathfrak{S}/p^n\mathfrak{S}$. Let $E(u) \in W[u]$ be the (monic) Eisenstein polynomial of the uniformizer π. Then a Kisin module (of E-height ≤ 1) is an $\mathfrak{S}$-module endowed with a φ-semilinear map $\varphi_{\mathfrak{M}} : \mathfrak{M} \to \mathfrak{M}$, which we also write abusively as φ, such that the cokernel of the map

$$1 \otimes \varphi : \varphi^*\mathfrak{M} = \mathfrak{S} \otimes_{\varphi,\mathfrak{S}} \mathfrak{M} \to \mathfrak{M}$$

is killed by $E(u)$. The Kisin modules form an exact category in an obvious manner, and its full subcategory consisting of $\mathfrak{M}$ such that $\mathfrak{M}$ is free of finite rank over $\mathfrak{S}$ (resp. free of finite rank over $\mathfrak{S}_1$, resp. finitely generated, p-power torsion and u-torsion free) is denoted by $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}}$ (resp. $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_1}$, resp. $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$).

We also have categories of Breuil modules $\mathrm{Mod}^{1,\varphi}_{/S}$, $\mathrm{Mod}^{1,\varphi}_{/S_1}$ and $\mathrm{Mod}^{1,\varphi}_{/S_\infty}$ defined as follows (for more precise definitions, see for example [Hattori 2012, §2.1], where the definitions are valid also for $p = 2$). Let S be the p-adic completion of the divided power envelope of $W[u]$ with respect to the ideal $(E(u))$ and put $S_n = S/p^nS$. The ring S has a natural divided power ideal $\mathrm{Fil}^1 S$, a continuous φ-semilinear endomorphism defined by $u \mapsto u^p$ which is also denoted by φ and a differential operator $N : S \to S$ defined by $N(u) = -u$. We can also define a φ-semilinear map $\varphi_1 = p^{-1}\varphi : \mathrm{Fil}^1 S \to S$. Then a Breuil module (of Hodge–Tate weights in [0, 1]) is an S-module endowed with an S-submodule $\mathrm{Fil}^1\mathcal{M}$ containing $(\mathrm{Fil}^1 S)\mathcal{M}$ and a φ-semilinear map $\varphi_{1,\mathcal{M}} : \mathrm{Fil}^1\mathcal{M} \to \mathcal{M}$ satisfying some conditions. We also define $\varphi_{\mathcal{M}} : \mathcal{M} \to \mathcal{M}$ by $\varphi_{\mathcal{M}}(x) = \varphi_1(E(u))^{-1}\varphi_{1,\mathcal{M}}(E(u)x)$. We drop the subscript $\mathcal{M}$ if there is no risk of confusion. The Breuil modules also form an exact category. Its full subcategory $\mathrm{Mod}^{1,\varphi}_{/S}$ (resp. $\mathrm{Mod}^{1,\varphi}_{/S_1}$) is defined to be the one consisting of $\mathcal{M}$ such that $\mathcal{M}$ is free of finite rank over S and $\mathcal{M}/\mathrm{Fil}^1\mathcal{M}$ is p-torsion free (resp. $\mathcal{M}$ is free of finite rank over S_1). The category $\mathrm{Mod}^{1,\varphi}_{/S_\infty}$ is defined as the smallest full subcategory containing $\mathrm{Mod}^{1,\varphi}_{/S_1}$ and closed under extensions. Then the functor $\mathfrak{M} \mapsto S \otimes_{\varphi,\mathfrak{S}} \mathfrak{M}$ induces exact functors

$$\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}} \to \mathrm{Mod}^{1,\varphi}_{/S}, \quad \mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_1} \to \mathrm{Mod}^{1,\varphi}_{/S_1}, \quad \mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty} \to \mathrm{Mod}^{1,\varphi}_{/S_\infty}$$

which are all denoted by $\mathcal{M}_{\mathfrak{S}}(-)$, by putting

$$\mathrm{Fil}^1\mathcal{M}_{\mathfrak{S}}(\mathfrak{M}) = \mathrm{Ker}(S \otimes_{\varphi,\mathfrak{S}} \mathfrak{M} \overset{1\otimes\varphi}{\to} S/\mathrm{Fil}^1 S \otimes_{\mathfrak{S}} \mathfrak{M}).$$

Put $\underline{\pi} = (\pi_0, \pi_1, \dots) \in R$ as before and consider the Witt ring $W(R)$ as an $\mathfrak{S}$-algebra by the map $u \mapsto [\underline{\pi}]$. The p-adic period ring A_{crys} is defined as the p-adic completion of the divided power envelope of $W(R)$ with respect to the ideal $E(u)W(R)$ and the ring $A_{\mathrm{crys}}[1/p]$ is denoted by B^+_{crys}. For any $r = (r_0, r_1, \dots) \in R$ with $r_l \in \mathcal{O}_{\bar{K},1}$, choose a lift $\hat{r}_l$ of r_l in $\mathcal{O}_{\bar{K}}$ and put $r^{(m)} = \lim_{l\to\infty} \hat{r}_{l+m}^{p^l} \in \hat{\mathcal{O}}_{\bar{K}}$. Consider the surjection $\theta_n : W_n(R) \to \mathcal{O}_{\bar{K},n}$ sending $(r_0, r_1, \dots, r_{n-1})$ to $\sum_{l=0}^{n-1} p^l r_l^{(l)}$. Then the quotient $A_{\mathrm{crys}}/p^n A_{\mathrm{crys}}$ can be identified with the divided power envelope $W_n^{\mathrm{DP}}(R)$ of the surjection θ_n compatible with the canonical divided power structure on the ideal $pW_n(R)$. For any objects $\mathfrak{M} \in \mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}}$ and $\mathcal{M} \in \mathrm{Mod}^{1,\varphi}_{/S}$, we have the associated G_{K_∞}-modules

$$T^*_{\mathfrak{S}}(\mathfrak{M}) = \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, W(R)), \quad T^*_{\mathrm{crys}}(\mathcal{M}) = \mathrm{Hom}_{S,\varphi,\mathrm{Fil}^1}(\mathcal{M}, A_{\mathrm{crys}}),$$

which are related by the injection

$$T^*_{\mathfrak{S}}(\mathfrak{M}) \to T^*_{\mathrm{crys}}(\mathcal{M}_{\mathfrak{S}}(\mathfrak{M}))$$

defined by $f \mapsto 1 \otimes (\varphi \circ f)$. Similarly, for any object $\mathfrak{M} \in \mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$, we have the associated G_{K_∞}-module

$$T^*_{\mathfrak{S}}(\mathfrak{M}) = \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, \mathbb{Q}_p/\mathbb{Z}_p \otimes_{\mathbb{Z}_p} W(R)).$$

Let D be an admissible filtered φ-module over K such that $\mathrm{gr}^i D_K = 0$ unless $i = 0, 1$. Put $S_{K_0} = S \otimes_W K_0$ and $\mathcal{D} = S_{K_0} \otimes_{K_0} D$. The S_{K_0}-module $\mathcal{D}$ is endowed with a natural Frobenius map $\varphi_{\mathcal{D}} : \mathcal{D} \to \mathcal{D}$ induced by the Frobenius of D, a derivation $N_{\mathcal{D}} = N \otimes 1 : \mathcal{D} \to \mathcal{D}$ and an S_{K_0}-submodule $\mathrm{Fil}^1\mathcal{D}$ defined as the inverse image of $\mathrm{Fil}^1 D_K$ by the map $\mathcal{D} \to \mathcal{D}/(\mathrm{Fil}^1 S)\mathcal{D} = D_K$. Then a strongly divisible lattice in $\mathcal{D}$ is an S-submodule $\mathcal{M}$ of $\mathcal{D}$ which satisfies the following:

- $\mathcal{M}$ is a free S-module of finite rank and $\mathcal{D} = \mathcal{M}[1/p]$.
- $\mathcal{M}$ is stable under $\varphi_{\mathcal{D}}$ and $N_{\mathcal{D}}$.
- $\varphi_{\mathcal{D}}(\mathrm{Fil}^1\mathcal{M}) \subseteq p\mathcal{M}$, where $\mathrm{Fil}^1\mathcal{M} = \mathcal{M} \cap \mathrm{Fil}^1\mathcal{D}$.

We put $V^*_{\mathrm{crys}}(\mathcal{D}) = \mathrm{Hom}_{S_{K_0},\varphi,\mathrm{Fil}^1}(\mathcal{D}, B^+_{\mathrm{crys}})$. If $\mathcal{M}$ is a strongly divisible lattice in $\mathcal{D}$, then the natural G_{K_∞}-actions on $T^*_{\mathrm{crys}}(\mathcal{M})$ and $V^*_{\mathrm{crys}}(\mathcal{D}) = T^*_{\mathrm{crys}}(\mathcal{M})[1/p]$ extend to G_K-actions and we have a natural isomorphism of G_K-modules

$$V^*_{\mathrm{crys}}(\mathcal{D}) \to V^*_{\mathrm{crys}}(D) = \mathrm{Hom}_{K_0,\varphi,\mathrm{Fil}}(D, B^+_{\mathrm{crys}})$$

[Breuil 2002, Proposition 2.2.5] and [Liu 2008, Lemma 5.2.1].

Let $(\mathrm{BT}/\mathcal{O}_K)$ (resp. $(p\text{-Gr}/\mathcal{O}_K)$) be the exact category of Barsotti–Tate groups (resp. finite flat group schemes killed by some p-power) over $\mathcal{O}_K$. For any Barsotti–Tate group Γ over $\mathcal{O}_K$, we let $T_p(\Gamma)$ denote its p-adic Tate module, $V_p(\Gamma) = \mathbb{Q}_p \otimes_{\mathbb{Z}_p} T_p(\Gamma)$ and $D^*(\Gamma)$ be the filtered φ-module over K associated to $V_p(\Gamma)$. We also let $\mathbb{D}^*(-)$ denote the contravariant crystalline Dieudonné functor [Berthelot et al. 1982] and consider its module of sections

$$\mathbb{D}^*(\Gamma)(S \to \mathcal{O}_K) = \varprojlim_n \mathbb{D}^*(\Gamma)(S_n \to \mathcal{O}_{K,n})$$

on the divided power thickening $S \to \mathcal{O}_K$ defined by $u \mapsto \pi$. Note that the S-module $\mathbb{D}^*(\Gamma)(S \to \mathcal{O}_K)$ can be considered as an object of the category $\mathrm{Mod}^{1,\varphi}_{/S}$ and also as a strongly divisible lattice in $\mathcal{D}^*(\Gamma) = S_{K_0} \otimes_{K_0} D^*(\Gamma)$ [Faltings 1999, §6]. For any finite flat group scheme $\mathcal{G}$ over $\mathcal{O}_K$ killed by some p-power, we define an object $\mathbb{D}^*(\mathcal{G})(S \to \mathcal{O}_K)$ of the category $\mathrm{Mod}^{1,\varphi}_{/S_\infty}$ similarly. Then we have the following classification theorem, whose first assertion (which is Theorem 2.2.7 of [Kisin 2006] for $p \geq 3$, and Theorem 4.1 and Proposition 4.2 of [Kim 2012] for $p = 2$) implies the second one (Theorem 2.3.5 of [Kisin 2006] for $p \geq 3$, and Corollary 4.3 of [Kim 2012] for $p = 2$) by an argument of taking a resolution.

Theorem 2.1 (Kisin). (1) *There exists an anti-equivalence of exact categories*

$$\mathfrak{M}^*(-) : (\mathrm{BT}/\mathcal{O}_K) \to \mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}}$$

with a natural isomorphism of G_{K_∞}-modules

$$\varepsilon_\Gamma : T_p(\Gamma) \to T^*_{\mathfrak{S}}(\mathfrak{M}^*(\Gamma)).$$

Moreover, the S-module $\mathcal{M}_{\mathfrak{S}}(\mathfrak{M}^*(\Gamma))$ can be considered as a strongly divisible lattice in $\mathcal{D}^*(\Gamma)$ and we also have a natural isomorphism of strongly divisible lattices in $\mathcal{D}^*(\Gamma)$

$$\mu_\Gamma : \mathcal{M}_{\mathfrak{S}}(\mathfrak{M}^*(\Gamma)) \to \mathbb{D}^*(\Gamma)(S \to \mathcal{O}_K).$$

(2) *There exists an anti-equivalence of exact categories*

$$\mathfrak{M}^*(-) : (p\text{-}\mathrm{Gr}/\mathcal{O}_K) \to \mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$$

with a natural isomorphism of G_{K_∞}-modules

$$\varepsilon_{\mathcal{G}} : \mathcal{G}(\mathcal{O}_{\bar{K}}) \to T^*_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G})).$$

Moreover, we also have a natural isomorphism of the category $\mathrm{Mod}^{1,\varphi}_{/S_\infty}$

$$\mu_{\mathcal{G}} : \mathcal{M}_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G})) \to \mathbb{D}^*(\mathcal{G})(S \to \mathcal{O}_K).$$

On the other hand, for any object $\mathfrak{M}$ of the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}}$ or $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$, we can define a dual object $\mathfrak{M}^\vee$ which is compatible with Cartier duality of Barsotti–Tate groups or finite flat group schemes. In particular, for any object $\mathfrak{M}$ of the category

$\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$ killed by p^n, we have a commutative diagram of G_{K_∞}-modules

$$\begin{tikzcd}
\mathcal{G}(\mathcal{O}_{\bar{K}}) \arrow[d, "\varepsilon_{\mathcal{G}}"', "\wr"] & \times & \mathcal{G}^\vee(\mathcal{O}_{\bar{K}}) \arrow[d, "\delta_{\mathcal{G}}"', "\wr"] \arrow[r] & \mathbb{Z}/p^n\mathbb{Z}(1) \arrow[d] \\
T^*_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G})) & \times & T^*_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G})^\vee) \arrow[r] & W_n(R)
\end{tikzcd}$$

where the upper horizontal arrow is the pairing of Cartier duality, the lower horizontal arrow is a natural perfect pairing, $\delta_{\mathcal{G}}$ is the composite

$$\mathcal{G}^\vee(\mathcal{O}_{\bar{K}}) \overset{\varepsilon_{\mathcal{G}^\vee}}{\simeq} T^*_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G}^\vee)) \simeq T^*_{\mathfrak{S}}(\mathfrak{M}^*(\mathcal{G})^\vee)$$

and the right vertical arrow is an injection (see [Kim 2012, §5.1], and also [Hattori 2012, Proposition 4.4]).

Let Γ be a Barsotti–Tate group over $\mathcal{O}_K$. We consider any element g of $T_p(\Gamma)$ as a homomorphism $g : \mathbb{Q}_p/\mathbb{Z}_p \to \Gamma \times \mathrm{Spec}(\hat{\mathcal{O}}_{\bar{K}})$. By evaluating the map

$$\mathbb{D}^*(g) : \mathbb{D}^*(\Gamma \times \mathrm{Spec}(\hat{\mathcal{O}}_{\bar{K}})) \to \mathbb{D}^*(\mathbb{Q}_p/\mathbb{Z}_p)$$

on the natural divided power thickening $A_{\mathrm{crys}} \to \hat{\mathcal{O}}_{\bar{K}}$, we obtain a homomorphism of G_{K_∞}-modules

$$\begin{aligned} T_p(\Gamma) \to \mathrm{Hom}_{S,\varphi,\mathrm{Fil}}\big(\mathbb{D}^*(\Gamma)(A_{\mathrm{crys}} \to \hat{\mathcal{O}}_{\bar{K}}), \mathbb{D}^*(\mathbb{Q}_p/\mathbb{Z}_p)(A_{\mathrm{crys}} \to \hat{\mathcal{O}}_{\bar{K}})\big) \\ = T^*_{\mathrm{crys}}(\mathbb{D}^*(\Gamma)(S \to \mathcal{O}_K)). \end{aligned}$$

This map is an injection, and an isomorphism after inverting p [Faltings 1999, Theorem 7]. Then we have the following compatibility of this map with the Breuil–Kisin classification.

Lemma 2.2. *Let Γ be a Barsotti–Tate group over $\mathcal{O}_K$. Then the following diagram is commutative*:

$$\begin{tikzcd}
T_p(\Gamma) \arrow[r, "\sim", "\varepsilon_\Gamma"'] \arrow[d, hook] & T_{\mathfrak{S}}(\mathfrak{M}^*(\Gamma)) \arrow[d, hook] \\
T^*_{\mathrm{crys}}(\mathbb{D}^*(\Gamma)(S \to \mathcal{O}_K)) \arrow[r, "\sim", "T^*_{\mathrm{crys}}(\mu_\Gamma)"'] & T^*_{\mathrm{crys}}(\mathcal{M}_{\mathfrak{S}}(\mathfrak{M}^*(\Gamma)))
\end{tikzcd}$$

Proof. Put $D = D^*(\Gamma)$ and $\mathfrak{M} = \mathfrak{M}^*(\Gamma)$. Consider the diagram

$$\begin{tikzcd}
T_p(\Gamma) \arrow[r] \arrow[dr] & T^*_{\mathrm{crys}}(\mathbb{D}^*(\Gamma)(S \to \mathcal{O}_K)) \arrow[r, "\sim"] \arrow[d] & T^*_{\mathrm{crys}}(\mathcal{M}_{\mathfrak{S}}(\mathfrak{M})) \arrow[dl] & T^*_{\mathfrak{S}}(\mathfrak{M}) \arrow[l] \arrow[dll] \\
& V^*_{\mathrm{crys}}(D) & &
\end{tikzcd}$$

where the left and middle triangles are commutative by [Kim 2012, Theorem 5.6.2] and Theorem 2.1 (1), respectively. The commutativity of the right one is remarked in [Kim 2012, footnote 11]. We briefly reproduce a proof of this remark for the convenience of the reader. We follow the notation of [Kisin 2006]. In particular, let $\mathcal{O} = \mathcal{O}_{[0,1)}$ be the ring of rigid-analytic functions on the open unit disc over K_0 and $M = \mathcal{O} \otimes_{\mathfrak{S}} \mathfrak{M}$ be the associated φ-module over the ring $\mathcal{O}$. We also put $\mathcal{D}_0 = (\mathcal{O}[l_u] \otimes_{K_0} D)^{N=0} = \mathcal{O} \otimes_{K_0} D$. Then the map $T^*_{\mathfrak{S}}(\mathfrak{M}) \to V^*_{\mathrm{crys}}(D)$ is defined as the composite

$$\mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, W(R)) \to \mathrm{Hom}_{\mathcal{O},\varphi}(M, B^+_{\mathrm{crys}}) \xrightarrow{(1\otimes\varphi)^*} \mathrm{Hom}_{\mathcal{O},\varphi}(\varphi^*M, B^+_{\mathrm{crys}})$$
$$\xrightarrow{(1\otimes\xi)^*} \mathrm{Hom}_{\mathcal{O},\varphi,\mathrm{Fil}}(\mathcal{D}_0, B^+_{\mathrm{crys}}) \to \mathrm{Hom}_{K_0,\varphi,\mathrm{Fil}}(D, B^+_{\mathrm{crys}}).$$

Here the map $\xi : D \to M$ is the unique φ-compatible section and the map $1 \otimes \xi : \mathcal{D}_0 = \mathcal{O} \otimes_{K_0} D \to M$ factors through the injection

$$1 \otimes \varphi : \varphi^* M = \mathcal{O} \otimes_{\varphi,\mathcal{O}} M \to M$$

[Kisin 2006, Lemma 1.2.6]. Put $\mathcal{D}_{\mathfrak{S}}(\mathfrak{M}) = \mathcal{M}_{\mathfrak{S}}(\mathfrak{M})[1/p] = S_{K_0} \otimes_{\mathcal{O}} \varphi^* M$. Then we have $K_0 \otimes_{S_{K_0}} \mathcal{D}_{\mathfrak{S}}(\mathfrak{M}) = K_0 \otimes_{\varphi,K_0} D$ and the composite

$$s_0 : K_0 \otimes_{\varphi,K_0} D \xrightarrow{1\otimes\varphi} D \xrightarrow{\xi} \varphi^* M \to \mathcal{D}_{\mathfrak{S}}(\mathfrak{M})$$

is the unique φ-compatible section. Using this, we can check that $K_0 \otimes_{\varphi,K_0} D \xrightarrow{1\otimes\varphi} D$ is an isomorphism of filtered φ-modules, where we consider on the left-hand side the induced filtration by the isomorphism

$$\mathcal{D}_{\mathfrak{S}}(\mathfrak{M})/(\mathrm{Fil}^1 S)\mathcal{D}_{\mathfrak{S}}(\mathfrak{M}) \to K \otimes_{\varphi,K_0} D,$$

and hence we can also check the above remark easily. Since the map ε_Γ is defined by identifying the images of $T_p(\Gamma)$ and $T^*_{\mathfrak{S}}(\mathfrak{M})$ in $V^*_{\mathrm{crys}}(D)$, the lemma follows. □

3. Lower ramification subgroups

In this section, we prove Theorem 1.1. We begin with the following lemma, which gives upper bounds of the lower ramification of finite flat group schemes. For any valuation ring V of height one with valuation v and any N-tuple $\underline{x} = (x_1, \dots, x_N)$ in V, we put $v(\underline{x}) = \min_{l=1,\dots,N} v(x_l)$.

Lemma 3.1. (1) *Let $\mathcal{K}/\mathbb{Q}_p$ be an extension of complete discrete valuation fields and $\mathcal{G}$ be a finite flat group scheme over $\mathcal{O}_{\mathcal{K}}$ killed by some p-power. Then we have $\mathcal{G}_i = 0$ for any $i > 1/(p-1)$.*

(2) *Let $\mathcal{K}$ be an extension of complete discrete valuation fields over $\mathbb{Q}_p$ or $k((u))$ with valuation v and $\mathcal{G}$ be a finite flat generically etale group scheme over $\mathcal{O}_{\mathcal{K}}$ killed by some p-power. Then we have the following.*

(a) $\mathcal{G}_i = (\mathcal{G}^0)_i$ *for any* $i > 0$.

(b) $\mathcal{G}_i = 0$ *for any* $i > \deg(\mathcal{G})/(p-1)$.

Here $\mathcal{G}_i$ *and* $\deg(\mathcal{G})$ *are defined using* v. *Namely, we extend* v *to a separable closure* $\mathcal{K}^{\mathrm{sep}}$ *of* $\mathcal{K}$, *write as* $\omega_{\mathcal{G}} \simeq \bigoplus_l \mathcal{O}_{\mathcal{K}}/(a_l)$ *and put*

$$\mathcal{G}_i(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}}}) = \mathrm{Ker}\big(\mathcal{G}(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}}}) \to \mathcal{G}(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}},i})\big), \quad \deg(\mathcal{G}) = \sum_l v(a_l).$$

Proof. For the assertion (1), we may replace $\mathcal{K}$ by its finite extension and assume $\mathcal{G}^\vee(\mathcal{O}_{\bar{\mathcal{K}}}) = \mathcal{G}^\vee(\mathcal{O}_{\mathcal{K}})$ for an algebraic closure $\bar{\mathcal{K}}$ of $\mathcal{K}$. By Cartier duality, there exists a generic isomorphism $\mathcal{G} \to \mathcal{G}' = \oplus_l \mu_{p^{n_l}}$ for some n_l. Then $\mathcal{G}'_i = 0$ for any $i > 1/(p-1)$ and the assertion follows from the commutative diagram

$$\begin{array}{ccc} \mathcal{G}(\mathcal{O}_{\bar{\mathcal{K}}}) & \xrightarrow[\sim]{} & \mathcal{G}'(\mathcal{O}_{\bar{\mathcal{K}}}) \\ \downarrow & & \downarrow \\ \mathcal{G}(\mathcal{O}_{\bar{\mathcal{K}},i}) & \longrightarrow & \mathcal{G}'(\mathcal{O}_{\bar{\mathcal{K}},i}) \end{array}$$

Let us consider the assertion (2). For any $i > 0$, we have a commutative diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{G}^0(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}}}) & \longrightarrow & \mathcal{G}(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}}}) & \longrightarrow & \mathcal{G}^{\mathrm{et}}(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}}}) & \longrightarrow & 0 \\ & & & & \downarrow & & \downarrow & & \\ & & & & \mathcal{G}(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}},i}) & \longrightarrow & \mathcal{G}^{\mathrm{et}}(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}},i}) & & \end{array}$$

where the upper row is the connected-etale sequence. Then the right vertical arrow is an isomorphism and the part (a) follows.

For the part (b), suppose $i > \deg(\mathcal{G})/(p-1)$. By part (a), we may assume that $\mathcal{G}$ is connected. By [Tian 2012, Proposition 1.5], we have a presentation of the affine algebra $\mathcal{O}_{\mathcal{G}}$ of $\mathcal{G}$

$$\mathcal{O}_{\mathcal{G}} \simeq \mathcal{O}_{\mathcal{K}}[[X_1, \ldots, X_d]]/(f_1, \ldots, f_d),$$
$$(f_1, \ldots, f_d) \equiv (X_1, \ldots, X_d)U \mod \deg p$$

with some $U \in M_d(\mathcal{O}_{\mathcal{K}})$ satisfying the equality $v(\det(U)) = \deg(\mathcal{G})$, where $X_1 = \cdots = X_d = 0$ gives the zero section. Let $\hat{U}$ be the matrix satisfying $U\hat{U} = \det(U)I_d$, where I_d is the identity matrix. For any element $\underline{x} = (x_1, \ldots, x_d)$ of $\mathcal{G}(\mathcal{O}_{\mathcal{K}^{\mathrm{sep}}})$, multiplying by $\hat{U}$ implies the inequality

$$v(\underline{x}) + v(\det(U)) \geq pv(\underline{x}).$$

Thus we obtain the inequality $v(\underline{x}) \leq \deg(\mathcal{G})/(p-1)$ unless $\underline{x} = 0$ and the assertion follows. $\square$

For any positive rational number $i \leq 1$, we let $W_n^{\mathrm{DP}}(R)_i$ denote the divided power envelope of the composite

$$\theta_{n,i} : W_n(R) \xrightarrow{\theta_n} \mathcal{O}_{\bar{K},n} \to \mathcal{O}_{\bar{K},i}, \ (r_0, \ldots, r_{n-1}) \mapsto \mathrm{pr}_0(r_0) \bmod m_{\bar{K}}^{\geqslant i}$$

compatible with the canonical divided power structure on the ideal $pW_n(R)$. Note that, by fixing a generator $\underline{p}^i$ of the principal ideal $m_R^{\geqslant i}$, we have an isomorphism of R-algebras

$$W_n(R)[Y_1, Y_2, \ldots]/([\underline{p}^i]^p - pY_1, Y_1^p - pY_2, Y_2^p - pY_3, \ldots) \to W_n^{\mathrm{DP}}(R)_i \quad (1)$$

sending Y_l to $\delta^l([\underline{p}^i])$, where we put $\delta(x) = (p-1)!\gamma_p(x)$ with the p-th divided power γ_p. The surjection $\theta_{n,i}$ defines a divided power thickening $W_n^{\mathrm{DP}}(R)_i \to \mathcal{O}_{\bar{K},i}$ over the thickening $S \to \mathcal{O}_K$, which is denoted by $A_{n,i}$. Put

$$I_{n,i} = \mathrm{Ker}(W_n(R) \xrightarrow{\varphi} W_n^{\mathrm{DP}}(R)_i).$$

From the definition, we see the inclusion $I_{n,i} \subseteq I_{n,i'}$ for any $i > i'$.

We show Theorem 1.1 by relating both sides of the isomorphism in its statement via Breuil modules using the lemma below.

Lemma 3.2. *Let $i \leq 1$ be a positive rational number and $\mathcal{G}$ be a finite flat group scheme over $\mathcal{O}_{K,i}$ killed by p^n. Then the map*

$$\begin{aligned}\mathcal{G}(\mathcal{O}_{\bar{K},i}) = \mathrm{Hom}_{\mathcal{O}_{\bar{K},i}}(\mathbb{Z}/p^n\mathbb{Z}, \mathcal{G} \times \bar{\mathcal{S}}_i) &\to \mathrm{Hom}(\mathbb{D}^*(\mathcal{G})(A_{n,i}), \mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z})(A_{n,i})) \\ &= \mathrm{Hom}(\mathbb{D}^*(\mathcal{G})(A_{n,i}), W_n^{\mathrm{DP}}(R)_i)\end{aligned}$$

defined by $g \mapsto \mathbb{D}^*(g)(A_{n,i})$ is an injection.

Proof. Suppose that a homomorphism $g : \mathbb{Z}/p^n\mathbb{Z} \to \mathcal{G} \times \bar{\mathcal{S}}_i$ satisfies $\mathbb{D}^*(g)(A_{n,i}) = 0$. We can take a finite extension L/K such that the map g is defined over $\mathrm{Spec}(\mathcal{O}_{L,i})$. Then we have the commutative diagram

$$\begin{array}{ccc} \mathrm{Hom}_{\mathcal{O}_{L,i}}(\mathbb{Z}/p^n\mathbb{Z}, \mathcal{G} \times \mathcal{S}_{L,i}) & \longrightarrow & \mathrm{Hom}(\mathbb{D}^*(\mathcal{G} \times \mathcal{S}_{L,i})(A_{n,i}), \mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z})(A_{n,i})) \\ \downarrow & & \wr\downarrow \\ \mathrm{Hom}_{\mathcal{O}_{\bar{K},i}}(\mathbb{Z}/p^n\mathbb{Z}, \mathcal{G} \times \bar{\mathcal{S}}_i) & \longrightarrow & \mathrm{Hom}(\mathbb{D}^*(\mathcal{G} \times \bar{\mathcal{S}}_i)(A_{n,i}), \mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z})(A_{n,i})) \end{array}$$

and thus we may assume $L = K$.

Put $\Sigma = \mathrm{Spec}(\mathbb{Z}_p)$ and $\Sigma_n = \mathrm{Spec}(\mathbb{Z}/p^n\mathbb{Z})$. Consider the big fppf crystalline site $\mathrm{CRYS}(\mathcal{S}_i/\Sigma)$ and its topos $(\mathcal{S}_i/\Sigma)_{\mathrm{CRYS}}$ [Berthelot et al. 1982]. Note that the local ring $\mathcal{O}_{K,i}$ is a Noetherian complete intersection ring and, for any finite extension L/K, the ring $\mathcal{O}_{L,i}$ is faithfully flat and of relative complete intersection over $\mathcal{O}_{K,i}$.

Thus, by [de Jong and Messing 1999, Proposition 1.2 and Lemma 4.1], we see that the composite

$$\begin{aligned}\mathrm{Hom}_{\mathcal{O}_{K,i}}(\mathbb{Z}/p^n\mathbb{Z}, \mathcal{G}) &\to \mathrm{Hom}_{(\mathcal{S}_i/\Sigma)_{\mathrm{CRYS}}}(\mathbb{D}^*(\mathcal{G}), \mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z}))\\ &\to \mathrm{Hom}_{(\bar{\mathcal{S}}_i/\Sigma)_{\mathrm{CRYS}}}(\mathbb{D}^*(\mathcal{G}), \mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z}))\end{aligned}$$

is an injection.

Consider the natural morphism of topoi

$$i_{n\mathrm{CRYS}} : (\bar{\mathcal{S}}_i/\Sigma_n)_{\mathrm{CRYS}} \to (\bar{\mathcal{S}}_i/\Sigma)_{\mathrm{CRYS}}.$$

Since the crystal $\mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z})$ is isomorphic to the quotient $\mathcal{O}_{\bar{\mathcal{S}}_i/\Sigma}/p^n\mathcal{O}_{\bar{\mathcal{S}}_i/\Sigma}$ of the structure sheaf $\mathcal{O}_{\bar{\mathcal{S}}_i/\Sigma}$ [Berthelot et al. 1982, Exemples 4.2.16] and this is equal to $i_{n\mathrm{CRYS}*}(\mathcal{O}_{\bar{\mathcal{S}}_i/\Sigma_n})$ [Berthelot et al. 1982, (4.2.17.4)], the natural map

$$\begin{aligned}i^*_{n\mathrm{CRYS}} : \mathrm{Hom}_{(\bar{\mathcal{S}}_i/\Sigma)_{\mathrm{CRYS}}}&(\mathbb{D}^*(\mathcal{G}), \mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z}))\\ &\to \mathrm{Hom}_{(\bar{\mathcal{S}}_i/\Sigma_n)_{\mathrm{CRYS}}}(i^*_{n\mathrm{CRYS}}(\mathbb{D}^*(\mathcal{G})), i^*_{n\mathrm{CRYS}}(\mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z})))\end{aligned}$$

is an isomorphism.

Finally, we claim that the thickening $A_{n,i}$ defines the final object of the big crystalline site $\mathrm{CRYS}(\bar{\mathcal{S}}_i/\Sigma_n)$. This follows as the proof of [Fontaine 1994, Théorème 1.2.1]. Indeed, it suffices to show that for any $\mathcal{O}_{\bar{K},i}$-algebra $\mathcal{O}_U$, any $\mathbb{Z}/p^n\mathbb{Z}$-algebra $\mathcal{O}_T$ and any surjection $\mathcal{O}_T \to \mathcal{O}_U$ defined by a divided power ideal J_T, the composite

$$W_n(R) \overset{\theta_{n,i}}{\to} \mathcal{O}_{\bar{K},i} \to \mathcal{O}_U$$

uniquely factors through $\mathcal{O}_T$. For this, we define the map $f : W_n(R) \to \mathcal{O}_T$ as follows: For any element $r = (r_0, \dots, r_{n-1})$ of the ring $W_n(R)$, choose a lift $\widehat{\mathrm{pr}_n(r_l)}$ in $\mathcal{O}_T$ of the element $\mathrm{pr}_n(r_l)$ for any $l = 0, \dots, n-1$ and put

$$f(r) = \sum_{l=0}^{n-1} p^l (\widehat{\mathrm{pr}_n(r_l)})^{p^{n-l}}.$$

This is independent of the choice of lifts and gives a ring homomorphism satisfying the condition. Conversely, suppose that a homomorphism $f' : W_n(R) \to \mathcal{O}_T$ satisfies the condition. Then, for any element $r = (r_0, \dots, r_{n-1})$ of the ring $W_n(R)$, we have $f'(r) = \sum_{l=0}^{n-1} p^l f'([r_l]^{1/p^n})^{p^{n-l}}$ and $f'([r_l]^{1/p^n}) \bmod J_T = \mathrm{pr}_n(r_l)$. Thus the uniqueness follows. Hence the evaluation map on the thickening $A_{n,i}$

$$\begin{aligned}\mathrm{Hom}_{(\bar{\mathcal{S}}_i/\Sigma_n)_{\mathrm{CRYS}}}&(i^*_{n\mathrm{CRYS}}(\mathbb{D}^*(\mathcal{G})), i^*_{n\mathrm{CRYS}}(\mathbb{D}^*(\mathbb{Z}/p^n\mathbb{Z})))\\ &\to \mathrm{Hom}(\mathbb{D}^*(\mathcal{G})(A_{n,i}), W_n^{\mathrm{DP}}(R)_i)\end{aligned}$$

is an injection. This concludes the proof of the lemma. □

Proof of Theorem 1.1. Take a resolution of $\mathcal{G}$ by Barsotti–Tate groups over $\mathcal{O}_K$

$$0 \to \mathcal{G} \to \Gamma_1 \to \Gamma_2 \to 0$$

and consider the associated exact sequence of Kisin modules

$$0 \to \mathfrak{N}_2 \to \mathfrak{N}_1 \to \mathfrak{M} \to 0.$$

Put $\mathcal{M} = \mathcal{M}_{\mathfrak{S}}(\mathfrak{M})$ and $\mathcal{N}_l = \mathcal{M}_{\mathfrak{S}}(\mathfrak{N}_l)$ for $l = 1, 2$. By Lemma 2.2 and the definition of the anti-equivalence $\mathfrak{M}^*(-)$, we have a diagram

$$\begin{tikzcd}
T_p(\Gamma_1) \arrow[r, hook] \arrow[d] \arrow[rr, dotted, bend left, hook, "\varepsilon_{\Gamma_1}"] & T^*_{\mathrm{crys}}(\mathcal{N}_1) \arrow[d] & T^*_{\mathfrak{S}}(\mathfrak{N}_1) \arrow[l, hook'] \arrow[d] \\
T_p(\Gamma_2) \arrow[r, hook] \arrow[d, "\pi_{\mathcal{G}}"'] \arrow[rr, dotted, bend left, hook, "\varepsilon_{\Gamma_2}"] & T^*_{\mathrm{crys}}(\mathcal{N}_2) \arrow[d, "\pi_{\mathcal{M}}"'] & T^*_{\mathfrak{S}}(\mathfrak{N}_2) \arrow[l, hook'] \arrow[d, "\pi_{\mathfrak{M}}"'] \\
\mathcal{G}(\mathcal{O}_{\bar{K}}) \arrow[r] \arrow[d] \arrow[rr, dotted, bend left, "\varepsilon_{\mathcal{G}}"] & \mathrm{Hom}_{S,\varphi}(\mathcal{M}, W_n^{\mathrm{DP}}(R)) \arrow[d] & T^*_{\mathfrak{S}}(\mathfrak{M}) \arrow[l] \arrow[d] \\
\mathcal{G}(\mathcal{O}_{\bar{K},i}) \arrow[r, hook] & \mathrm{Hom}_{S,\varphi}(\mathcal{M}, W_n^{\mathrm{DP}}(R)_i) & \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, W_n(R)/I_{n,i}) \arrow[l, hook']
\end{tikzcd}$$

where the left horizontal arrows are induced by $g \mapsto \mathbb{D}^*(g)$ and the right horizontal arrows are the maps sending f to $1 \otimes (\varphi \circ f)$. The middle left vertical arrow $\pi_{\mathcal{G}} : T_p(\Gamma_2) \to \mathcal{G}(\mathcal{O}_{\bar{K}})$ is defined as follows: For $g \in T_p(\Gamma_2)$, the element $p^n g$ is contained in the image of $T_p(\Gamma_1) = \varprojlim_l \Gamma_1[p^l](\mathcal{O}_{\bar{K}})$ and put $p^n g = h = (h_n)_{n>0}$. Then the element $h_n \in \Gamma_1[p^n](\mathcal{O}_{\bar{K}})$ is contained in the subgroup $\mathcal{G}(\mathcal{O}_{\bar{K}})$ and the map $\pi_{\mathcal{G}}$ is defined by $g \mapsto h_n$. We define the map $\pi_{\mathcal{M}} : T^*_{\mathrm{crys}}(\mathcal{N}_2) \to \mathrm{Hom}_{S,\varphi}(\mathcal{M}, W_n^{\mathrm{DP}}(R))$ similarly: For any map $f : \mathcal{N}_2 \to A_{\mathrm{crys}}$, the map $p^n f$ induces a map $\mathcal{N}_1 \to A_{\mathrm{crys}}$. Its composite with the natural map $A_{\mathrm{crys}} \to W_n^{\mathrm{DP}}(R)$ factors through $\mathcal{M}$ and defines the map $\pi_{\mathcal{M}}(f) : \mathcal{M} \to W_n^{\mathrm{DP}}(R)$. The map $\pi_{\mathfrak{M}}$ is defined in the same way. From these definitions, we see that the diagram is commutative. Note that the bottom left horizontal arrow is an injection by Lemma 3.2, and that the bottom right horizontal arrow is also an injection by the definition of the ideal $I_{n,i}$.

Thus, for any element $g \in \mathcal{G}(\mathcal{O}_{\bar{K}})$, its image in $\mathcal{G}(\mathcal{O}_{\bar{K},i})$ is zero if and only if the image of $\varepsilon_{\mathcal{G}}(g) \in T^*_{\mathfrak{S}}(\mathfrak{M})$ in $\mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, W_n(R)/I_{n,i})$ is zero. Hence the theorem follows. □

The special case of $n = 1$ of Theorem 1.1 can be interpreted as a correspondence of ramification for finite flat group schemes over $\mathcal{O}_K$ and $k[[u]]$ generalizing [Hattori 2012, Theorem 1.1], as follows. Recall that we have an anti-equivalence $\mathcal{H}(-)$ from the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_1}$ to an exact category of finite flat generically etale group schemes over $k[\![u]\!]$ whose Verschiebung is zero [Gabriel 1965, Théorème 7.4]. This gives the equality $T^*_{\mathfrak{S}}(\mathfrak{M}) = \mathcal{H}(\mathfrak{M})(R)$ for any object $\mathfrak{M}$ of the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_1}$. We normalize the indices of the upper and the lower ramification subgroups of finite flat generically etale group schemes $\mathcal{G}$ over $\mathcal{O}_K$ and $\mathcal{H}$ over $k[\![u]\!]$ to be adapted to v_p and v_R, respectively. In particular, we define the i-th lower ramification subgroup of $\mathcal{H}$ by

$$\mathcal{H}_i(R) = \mathrm{Ker}(\mathcal{H}(R) \to \mathcal{H}(R_i)).$$

Note that the field $\mathrm{Frac}(R)$ can be identified with the completion of an algebraic closure of $k((u))$.

Corollary 3.3. *Let p be a rational prime and $K/\mathbb{Q}_p$ be an extension of complete discrete valuation fields with perfect residue field k. Let $\mathcal{G}$ be a finite flat group scheme over $\mathcal{O}_K$ killed by p and consider the associated object $\mathfrak{M}^*(\mathcal{G})$ of the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_1}$. Then the map $\varepsilon_{\mathcal{G}} : \mathcal{G}(\mathcal{O}_{\bar{K}}) \simeq \mathcal{H}(\mathfrak{M}^*(\mathcal{G}))(R)$ induces the isomorphisms of G_{K_∞}-modules*

$$\mathcal{G}_i(\mathcal{O}_{\bar{K}}) \simeq \mathcal{H}(\mathfrak{M}^*(\mathcal{G}))_i(R), \quad \mathcal{G}^j(\mathcal{O}_{\bar{K}}) \simeq \mathcal{H}(\mathfrak{M}^*(\mathcal{G}))^j(R)$$

for any positive rational numbers i and j.

Proof. By Cartier duality, a theorem of Tian and Fargues [Tian 2010, Theorem 1.6; Fargues 2011, Proposition 6] and Theorem 3.3 of [Hattori 2012], it is enough to show the assertion of Corollary 3.3 on lower ramification subgroups. Moreover, since the i-th lower ramification subgroups of $\mathcal{G}$ and $\mathcal{H}(\mathfrak{M}^*(\mathcal{G}))$ vanish for any $i > 1/(p-1)$ [Hattori 2012, Corollary 3.5 and Remark 3.6], we may assume $i \leq 1$. Then the equality $I_{1,i} = m_R^{\geqslant i}$ and Theorem 1.1 imply Corollary 3.3. $\square$

4. Description of the ideal $I_{n,i}$

In this section, we give an explicit description of the ideal $I_{n,i}$. We identify the rings of both sides of the isomorphism (1).

Proposition 4.1. *Let $n_1, \dots, n_l$ be integers satisfying $0 \leq n_j \leq p-1$ for any j and r be an element of $W_n(R)$. If the element $rY_1^{n_1} \cdots Y_l^{n_l}$ is zero in the ring $W_n^{\mathrm{DP}}(R)_i$, then $[\underline{p}^i]^p \mid r$ in the ring $W_n(R)$. In particular, we have the inclusion $I_{n,i} \subseteq ([\underline{p}^i])$.*

Proof. By substituting $Y_j = 0$ for $j > l$, we reduce ourselves to showing that the equality in the ring $W_n(R)[Y_1, \dots, Y_l]$

$$rY_1^{n_1} \cdots Y_l^{n_l} = ([\underline{p}^i]^p - pY_1)f_0 + (Y_1^p - pY_2)f_1 + \cdots + (Y_{l-1}^p - pY_l)f_{l-1} + Y_l^p f_l \quad (2)$$

with $f_0, \dots, f_l$ in this ring implies $[\underline{p}^i]^p \mid r$. By replacing f_j's, we may assume the inequality

$$\deg_{j'}(f_j) < p \quad (j' = j+1, \dots, l), \quad (3)$$

where $\deg_{j'}$ means the degree with respect to $Y_{j'}$.

For any l-tuple $\underline{m} = (m_1, \dots, m_l)$, write $\underline{Y}^{\underline{m}} = Y_1^{m_1} \cdots Y_l^{m_l}$ and let $c_{j,\underline{m}}$ be the coefficient of $\underline{Y}^{\underline{m}}$ in f_j. Put $\underline{n} = (n_1, \dots, n_l)$ and $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 on the j-th entry. We consider a lexicographic order on the module $\mathbb{Z}^l$: we say $\underline{m} < \underline{m}'$ if there exists j with $1 \le j \le l$ such that $m_j < m'_j$ and $m_{j'} = m'_{j'}$ for any $j < j' \le l$. Taking the terms of scalar multiples of the monomial $\underline{Y}^{\underline{n}}$ in (2), we have the equality

$$r\underline{Y}^{\underline{n}} = [\underline{p}^i]^p c_{0,\underline{n}} \underline{Y}^{\underline{n}} + \sum_{j=0}^{l-1} (-pY_{j+1}) c_{j,\underline{n}-e_{j+1}} \underline{Y}^{\underline{n}-e_{j+1}}.$$

Now we claim that

$$c_{j,\underline{n}-e_{j+1}} = 0 \quad (j = 0, \dots, l-1). \tag{4}$$

Suppose the contrary. Choose j such that $0 \le j \le l-1$ and $c_{j,\underline{n}-e_{j+1}} \neq 0$. Consider the term $c_{j,\underline{n}-e_{j+1}} \underline{Y}^{\underline{n}-e_{j+1}}$ in f_j. The right-hand side of the equality (2) contains the term $c_{j,\underline{n}-e_{j+1}} \underline{Y}^{\underline{n}+pe_j-e_{j+1}}$ for $j \ge 1$ and $[\underline{p}^i]^p c_{0,\underline{n}-e_1} \underline{Y}^{\underline{n}-e_1}$ for $j = 0$. Note that, for $j' \le j-2$, the j-th entry of the l-tuple $\underline{n} + pe_j - e_{j+1} - e_{j'+1}$ is equal to $n_j + p$ and thus $f_{j'}$ does not contain any scalar multiple of $\underline{Y}^{\underline{n}+pe_j-e_{j+1}-e_{j'+1}}$ by assumption (3). Since $\underline{n} + pe_j - e_{j+1} < \underline{n}$ and $\underline{n} - e_1 < \underline{n}$, it follows from (2) that

$$c_{j,\underline{n}-e_{j+1}} \underline{Y}^{\underline{n}+pe_j-e_{j+1}} = -\sum_{j'=j-1}^{l-1} (-pY_{j'+1}) c_{j',\underline{n}+pe_j-e_{j+1}-e_{j'+1}} \underline{Y}^{\underline{n}+pe_j-e_{j+1}-e_{j'+1}}$$

for $j \ge 1$ and

$$[\underline{p}^i]^p c_{0,\underline{n}-e_1} \underline{Y}^{\underline{n}-e_1} = -\sum_{j'=0}^{l-1} (-pY_{j'+1}) c_{j',\underline{n}-e_1-e_{j'+1}} \underline{Y}^{\underline{n}-e_1-e_{j'+1}}$$

for $j = 0$.

We let Eq(1) denote this equation. Put $\underline{m}(1) = \underline{n} + pe_j - e_{j+1}$ for $j \ge 1$ and $\underline{m}(1) = \underline{n} - e_1$ for $j = 0$. Repeating this by arbitrarily choosing a term with nonzero coefficient $c_{j',\underline{m}'}$ on the right-hand side of the equation Eq(s), we obtain a series of equations Eq(1), Eq(2), ... and a sequence of l-tuples of non-negative integers $\underline{m}(1), \underline{m}(2), \dots$ such that Eq(s) is an equation of monomials of degree $\underline{m}(s)$ for any $s \ge 1$. Note that if there is no such term on the right-hand side of the equation Eq(s), the procedure stops. On the other hand, if the equation Eq(s) is either of the types

$$c\underline{Y}^{\underline{m}(s)} = \begin{cases} -\cdots - (Y_j^p) c_{j,\underline{m}(s)-pe_j} \underline{Y}^{\underline{m}(s)-pe_j} - \cdots & (1 \le j \le l-1), \\ -[\underline{p}^i]^p c_{0,\underline{m}(s)} \underline{Y}^{\underline{m}(s)} - \cdots & (j = 0), \end{cases}$$

with some $c \in W_n(R)$ such that the indicated term is chosen and that $c_{j,\underline{m}(s)-pe_j}$ (resp. $c_{0,\underline{m}(s)}$) is contained in the ideal $p^{n-1}W_n(R)$, then the equation Eq$(s+1)$ is empty and the procedure also stops. In the latter case, we put $\underline{m}(s+1) = \underline{m}(s) - pe_j + e_{j+1}$ for $1 \leq j \leq l-1$ and $\underline{m}(s+1) = \underline{m}(s) + e_1$ for $j = 0$.

Lemma 4.2. *The sequence $\underline{m}(s)$ is strictly decreasing with respect to the lexicographic order on $\mathbb{Z}^l$ defined as above.*

Proof. Note the inequalities $\underline{n} > \underline{m}(1) > \underline{m}(2)$. Suppose that we have $\underline{m}(1) > \underline{m}(2) > \cdots > \underline{m}(t) \leq \underline{m}(t+1)$ for some $t \geq 2$. Then the term $Y_l^p f_l$ in (2) does not affect Eq(s) for $1 \leq s \leq t$. Thus, by the construction, one of the following four cases holds for each $1 \leq s \leq t$:

(C_j) $\quad \underline{m}(s+1) = \underline{m}(s) + pe_j - e_{j+1}$ for some $1 \leq j \leq l-1$,

(C_j') $\quad \underline{m}(s+1) = \underline{m}(s) - pe_j + e_{j+1}$ for some $1 \leq j \leq l-1$,

(C_0) $\quad \underline{m}(s+1) = \underline{m}(s) - e_1$,

(C_0') $\quad \underline{m}(s+1) = \underline{m}(s) + e_1$.

Moreover, (C_j) and (C_j') do not occur consecutively for any j satisfying $0 \leq j \leq l-1$. Note that $\underline{m}(s) > \underline{m}(s+1)$ for (C_j) and $\underline{m}(s) < \underline{m}(s+1)$ for (C_j').

First we claim that (C_0') does not hold for $s = t$. Suppose the contrary. Then (C_j) holds for $s = t-1$ with some j satisfying $1 \leq j \leq l-1$. Hence the j-th entry $m(t)_j$ of the l-tuple $\underline{m}(t)$ is no less than p. The equation Eq(t)

$$c_{j,\underline{m}(t-1)-e_{j+1}}\underline{Y}^{\underline{m}(t)} = -[\underline{p}^i]^p c_{0,\underline{m}(t)}\underline{Y}^{\underline{m}(t)} - \cdots$$

implies $\deg_j(f_0) \geq p$. This contradicts (3).

Hence (C_j') holds for $s = t$ with some $1 \leq j \leq l-1$. From this we see that $m(t)_j \geq p$. Since $n_j < p$, there exists an integer t' with $1 \leq t' \leq t-2$ such that (C_j) holds for $s = t'$ and that it does not hold for any s satisfying $t' < s \leq t$.

Next we claim that $m(s)_j = m(t')_j + p$ for any s satisfying $t' < s \leq t$. Suppose the contrary and take the smallest integer t'' with $t' < t'' < t$ such that (C_{j-1}) holds for $s = t''$. Then $m(s)_j = m(t')_j + p$ for $t' < s \leq t''$ and $m(t''+1)_j = m(t')_j + p - 1$. By assumption, we also have $m(t''+1)_j \geq m(t)_j \geq p$. On the other hand, the equation Eq(t'') is

$$c\underline{Y}^{\underline{m}(t'')} = -\cdots - (-pY_j)c_{j-1,\underline{m}(t'')-e_j}\underline{Y}^{\underline{m}(t'')-e_j} - \cdots$$

with some $c \in W_n(R)$. Hence we obtain

$$\deg_j(f_{j-1}) \geq m(t'')_j - 1 = m(t')_j + p - 1 \geq p,$$

which contradicts (3).

Now let j_0 be the non-negative integer such that (C_{j_0}) holds for $s = t-1$. Then $j_0 \neq j, j-1$ by the constancy of $m(s)_j$ which we have just proved. The equation Eq$(t-1)$ is

$$c\underline{Y}^{\underline{m}(t-1)} = -\cdots - (-pY_{j_0+1})c_{j_0,\underline{m}(t-1)-e_{j_0+1}}\underline{Y}^{\underline{m}(t-1)-e_{j_0+1}} - \cdots$$

with some $c \in W_n(R)$ and thus $\deg_j(f_{j_0}) \geq m(t-1)_j = m(t')_j + p \geq p$. By assumption (3), we obtain $j_0 > j$. In particular, we have $j_0 \geq 1$ and $\underline{m}(t) = \underline{m}(t-1) + pe_{j_0} - e_{j_0+1}$. Therefore the equation Eq(t) is

$$c'\underline{Y}^{\underline{m}(t)} = -\cdots - (Y_j^p)c_{j,\underline{m}(t)-pe_j}\underline{Y}^{\underline{m}(t)-pe_j} - \cdots$$

with some $c' \in W_n(R)$ and $\deg_{j_0}(f_j) \geq m(t)_{j_0} \geq p$. This contradicts (3), and the lemma follows. □

By Lemma 4.2, the case (C'_j) does not occur in the procedure for any non-negative integer j. In particular, if there is no term with non-zero $c_{j',\underline{m}'}$ on the right-hand side of Eq(s) for some s, then the equation is

$$[\underline{p}^i]^{p\epsilon}c_{j'',\underline{m}''}\underline{Y}^{\underline{m}(s)} = 0,$$

where $c_{j'',\underline{m}''}\underline{Y}^{\underline{m}''}$ is the chosen term on the right-hand side of Eq$(s-1)$ and $\epsilon \in \{0, 1\}$. Note that this occurs for s satisfying $\underline{m}(s) = (0, \ldots, 0)$, since in this case (C_0) holds for $s-1$. Therefore, Lemma 4.2 implies that, for any choice of terms as above, we end up with an equation of this type for a sufficiently large s. Since the element $[\underline{p}^i]^p$ is a non-zero divisor in the ring $W_n(R)$, we see that $c_{j'',\underline{m}''} = 0$. This contradicts the choice of terms, and (4) follows.

Hence we obtain the equality

$$r\underline{Y}^{\underline{n}} = [\underline{p}^i]^p c_{0,\underline{n}}\underline{Y}^{\underline{n}}$$

and thus $[\underline{p}^i]^p \mid r$. This concludes the proof of Proposition 4.1. □

Lemma 4.3. *Put $\mathbf{n}(s) = v_p((ps)!)$ for any non-negative integer s. Then an element $r = (r_0, \ldots, r_{n-1})$ of the ring $W_n(R)$ is contained in the ideal $I_{n,i}$ if and only if the condition*

$$[\underline{p}^i]^s \mid (r_0, \ldots, r_{n-1-\mathbf{n}(s-1)}, 0, \ldots, 0) \tag{5}$$

holds for any $s \geq 1$.

Proof. Let r be an element of the ideal $I_{n,i}$ and show the condition (5) for r by induction on s. The case of $s = 1$ follows from Proposition 4.1. Suppose that the condition (5) holds for some $s \geq 1$. Let $r' = (r'_0, \ldots, r'_{n-1-\mathbf{n}(s-1)}, 0, \ldots, 0)$ be the element of $W_n(R)$ such that

$$(r_0, \ldots, r_{n-1-\mathbf{n}(s-1)}, 0, \ldots, 0) = [\underline{p}^i]^s r'.$$

We write the p-adic expansion of the integer s as

$$s = n_1 + pn_2 + \cdots + p^{l-1}n_l$$

with $0 \le n_j \le p-1$. Then in the ring $W_n^{\mathrm{DP}}(R)_i$ we have

$$\varphi(r) = p^{\mathbf{n}(s)}\varphi(r')Y_1^{n_1}\cdots Y_l^{n_l},$$

and Proposition 4.1 implies that $[\underline{p}^i]$ divides $p^{\mathbf{n}(s)}r'$. Hence the element $[\underline{p}^i]$ divides $(r'_0, \ldots, r'_{n-1-\mathbf{n}(s)}, 0, \ldots, 0)$ and thus

$$[\underline{p}^i]^{s+1} \mid (r_0, \ldots, r_{n-1-\mathbf{n}(s)}, 0, \ldots, 0).$$

Conversely, suppose that an element r of the ring $W_n(R)$ satisfies the condition (5) for any $s \ge 1$. Since we have $\mathbf{n}(s) \ge n$ for some s, a similar argument as above shows that $\varphi(r) = 0$ in the ring $W_n^{\mathrm{DP}}(R)_i$. This concludes the proof of the lemma. □

Remark 4.4. Lemma 4.3 enables us to compute the ideal $I_{n,i}$. For example, $I_{2,i} = (m_R^{\geqslant 2i}, m_R^{\geqslant pi}) \subseteq W_2(R)$ and

$$I_{3,i} = \begin{cases} (m_R^{\geqslant 2i}, m_R^{\geqslant 4i}, m_R^{\geqslant 4i}) & (p = 2), \\ (m_R^{\geqslant 3i}, m_R^{\geqslant 2pi}, m_R^{\geqslant p^2 i}) & (p \ge 3). \end{cases}$$

Finally we prove a relationship between the ideals $I_{n-1,pi}$ and $I_{n,i}$, which will be used in Section 5.

Lemma 4.5. *For any $r = (r_0, \ldots, r_{n-2}) \in I_{n-1,pi}$ and $r_{n-1} \in R$, we have*

$$\hat{r} = (r_0, \ldots, r_{n-2}, \underline{p}^{ip^{n-1}} r_{n-1}) \in I_{n,i}.$$

Proof. By Lemma 4.3, we have

$$[\underline{p}^{pi}]^s \mid (r_0, \ldots, r_{n-2-\mathbf{n}(s-1)}, 0, \ldots, 0)$$

in the ring $W_{n-1}(R)$ for any $s \ge 1$ satisfying $\mathbf{n}(s-1) < n-1$. Let us show that the element $\hat{r} = (\hat{r}_0, \ldots, \hat{r}_{n-1})$ satisfies the condition

$$[\underline{p}^i]^s \mid (\hat{r}_0, \ldots, \hat{r}_{n-1-\mathbf{n}(s-1)}, 0, \ldots, 0)$$

in the ring $W_n(R)$ for any $s \ge 1$ satisfying $\mathbf{n}(s-1) < n$. The case of $s = 1$ follows from the definition of $\hat{r}$. Suppose $s \ge 2$. Since $\mathbf{n}(s-2)+1 \le \mathbf{n}(s-1)$, we have $n-1-\mathbf{n}(s-1) \le n-2-\mathbf{n}(s-2)$ and $[\underline{p}^{pi}]^{s-1}$ divides $(\hat{r}_0, \ldots, \hat{r}_{n-1-\mathbf{n}(s-1)})$. Then the inequality $p(s-1) \ge s$ implies the condition. This concludes the proof of the lemma. □

5. Application to canonical subgroups

In this section, we prove Theorem 1.2 and Theorem 1.3. First we consider Theorem 1.2. Let $K/\mathbb{Q}_p$ be an extension of complete discrete valuation fields. Let $\mathcal{G}$ be a truncated Barsotti–Tate group of level n, height h and dimension d over $\mathcal{O}_K$ with $0 < d < h$ and Hodge height $w < (p-1)/p^n$. Let $\mathcal{C}_n$ be the level n canonical subgroup of $\mathcal{G}$ as in Theorem 1.1 of [Hattori 2014]. By a base change argument and the uniqueness of $\mathcal{C}_n$ (see Proposition 3.8 of the same reference), we may assume that the residue field k is perfect. Recall that we normalized the valuation v_R on the ring R as $v_R(\underline{\pi}) = 1/e$ in Section 1.

Let $\mathfrak{M} = \mathfrak{M}^*(\mathcal{G})$ be the corresponding object of the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_\infty}$. Then, by Remark 3.4 of [Hattori 2014], we can show as in the proof of [Hattori 2013, Lemma 3.3] that the object $\mathfrak{M}/p\mathfrak{M}$ has a basis $\bar{e}_1, \dots, \bar{e}_h$ such that

$$\varphi(\bar{e}_1, \dots, \bar{e}_h) = (\bar{e}_1, \dots, \bar{e}_h) \begin{pmatrix} P_1 & P_2 \\ u^e P_3 & u^e P_4 \end{pmatrix},$$

where the matrices P_i have entries in the ring $k[\![u]\!]$ with

$$P_1 \in M_{h-d}(k[\![u]\!]), \quad v_R(\det(P_1)) = w, \quad \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix} \in \mathrm{GL}_h(k[\![u]\!]).$$

Let $\hat{P}_1$ be the element of $M_{h-d}(k[\![u]\!])$ such that $P_1\hat{P}_1 = u^{ew} I_{h-d}$. Let B be the unique solution in $M_{d,h-d}(k[\![u]\!])$ of the equation

$$B = P_3\hat{P}_1 - u^{ep(1-w)-ew} B P_2 \varphi(B)\hat{P}_1 + u^{ep(1-w)} P_4\varphi(B)\hat{P}_1$$

and put $D = P_1 + u^{ep(1-w)} P_2\varphi(B)$, which also satisfies $v_R(\det(D)) = w$ (see the proof just cited). Moreover, put

$$(\bar{e}'_1, \dots, \bar{e}'_{h-d}) = (\bar{e}_1, \dots, \bar{e}_h) \begin{pmatrix} I_{h-d} \\ u^{e(1-w)} B \end{pmatrix}.$$

The elements $\bar{e}'_1, \dots, \bar{e}'_{h-d}, \bar{e}_{h-d+1}, \dots, \bar{e}_h$ form a basis of the $\mathfrak{S}_1$-module $\mathfrak{M}/p\mathfrak{M}$ satisfying

$$\varphi(\bar{e}'_1, \dots, \bar{e}'_{h-d}, \bar{e}_{h-d+1}, \dots, \bar{e}_h) = (\bar{e}'_1, \dots, \bar{e}'_{h-d}, \bar{e}_{h-d+1}, \dots, \bar{e}_h) \begin{pmatrix} D & P_2 \\ 0 & u^{e(1-w)} P'_4 \end{pmatrix}$$

for some matrix $P'_4 \in M_d(k[\![u]\!])$. Then we have the following description of the level one canonical subgroup $\mathcal{C}_1$ of $\mathcal{G}[p]$.

Lemma 5.1. *Let f be an element of the module* $\mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}/p\mathfrak{M}, R)$ *defined by*

$$(\bar{e}_1, \dots, \bar{e}_h) \mapsto (\underline{x}, \underline{y})$$

with an $(h-d)$-tuple $\underline{x}$ and a d-tuple $\underline{y}$ in R. Then f corresponds to an element of $\mathscr{C}_1(\mathcal{O}_{\bar{K}})$ by the isomorphism

$$\varepsilon_{\mathscr{G}[p]} : \mathscr{G}[p](\mathcal{O}_{\bar{K}}) \simeq \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}/p\mathfrak{M}, R)$$

if and only if $v_R(\underline{x} + u^{e(1-w)}\underline{y}B) > w/(p-1)$.

Proof. Let $\mathfrak{L}$ be the $\mathfrak{S}_1$-submodule of $\mathfrak{M}/p\mathfrak{M}$ generated by $\bar{e}'_1, \dots, \bar{e}'_{h-d}$. Then $\mathfrak{L}$ defines a subobject of $\mathfrak{M}/p\mathfrak{M}$ in the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_1}$. Put $\mathfrak{N} = (\mathfrak{M}/p\mathfrak{M})/\mathfrak{L}$. Lemma 3.2 of [Hattori 2014] also holds for our $\mathscr{G}[p]$ and its subgroup scheme corresponding to $\mathfrak{N}$, by Remark 3.4 of the same reference. By Lemma 3.2 and Theorem 3.5(1) of that reference, the level one canonical subgroup $\mathscr{C}_1$ is the closed subgroup scheme of $\mathscr{G}[p]$ corresponding to the object $\mathfrak{N}$. We have the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathscr{C}_1(\mathcal{O}_{\bar{K}}) \arrow[r] \arrow[d, "\varepsilon_{\mathscr{C}_1}", "\wr"'] & \mathscr{G}[p](\mathcal{O}_{\bar{K}}) \arrow[r] \arrow[d, "\varepsilon_{\mathscr{G}[p]}", "\wr"'] & (\mathscr{G}[p]/\mathscr{C}_1)(\mathcal{O}_{\bar{K}}) \arrow[r] \arrow[d, "\varepsilon_{\mathscr{G}[p]/\mathscr{C}_1}", "\wr"'] & 0 \\
0 \arrow[r] & \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{N}, R) \arrow[r] & \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}/p\mathfrak{M}, R) \arrow[r, "\iota^*"] & \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{L}, R) \arrow[r] & 0
\end{tikzcd}$$

where the rows are exact and the vertical arrows are isomorphisms. The element f corresponds to an element of $\mathscr{C}_1(\mathcal{O}_{\bar{K}})$ if and only if $\iota^*(f) = 0$. The map $\iota^*(f) : \mathfrak{L} \to R$ is defined by

$$(\bar{e}'_1, \dots, \bar{e}'_{h-d}) \mapsto \underline{x} + u^{e(1-w)}\underline{y}B,$$

which we consider as an element of $\mathscr{H}(\mathfrak{L})(R)$. Since $\deg(\mathscr{H}(\mathfrak{L})) = w$, the lemma follows from [Hattori 2013, Lemma 2.4]. □

Recall that we put

$$i_n = 1/(p^{n-1}(p-1)) - w/(p-1), \quad i'_n = 1/(p^n(p-1)).$$

Lemma 5.2. *If $w < (p-1)/p^n$, then we have $\mathscr{C}_1 = \mathscr{G}[p]_{i_m} = \mathscr{G}[p]_{i'_m}$ for any integer m satisfying $1 \le m \le n$.*

Proof. By [Hattori 2014, Theorem 1.1(c)], the equality $\mathscr{C}_1 = \mathscr{G}[p]_{i_1}$ holds. From the inequalities

$$i'_n < i_n \le i'_{n-1} < \cdots < i_2 \le i'_1 < i_1,$$

we have the inclusions

$$\mathscr{C}_1 \subseteq \mathscr{G}[p]_{i'_1} \subseteq \mathscr{G}[p]_{i_2} \subseteq \cdots \subseteq \mathscr{G}[p]_{i_n} \subseteq \mathscr{G}[p]_{i'_n}.$$

Let us show the reverse inclusion. Let $\mathfrak{N}$ be the quotient of $\mathfrak{M}/p\mathfrak{M}$ in the category $\mathrm{Mod}^{1,\varphi}_{/\mathfrak{S}_1}$ corresponding to the closed subgroup scheme $\mathscr{C}_1 \subseteq \mathscr{G}$. By Corollary 3.3, it is enough to show that

$$\mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}/p\mathfrak{M}, m_R^{\geqslant i'_n}) \subseteq \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{N}, R).$$

Consider a φ-compatible homomorphism of $\mathfrak{S}$-modules $\mathfrak{M}/p\mathfrak{M} \to R$ defined by

$$(\bar{e}_1, \dots, \bar{e}_h) \mapsto (\underline{x}, \underline{y}) = \underline{p}^{i'_n}(\underline{a}, \underline{b})$$

with an $(h-d)$-tuple $\underline{a}$ and a d-tuple $\underline{b}$ in R. Then we have

$$\underline{p}^{pi'_n}(\underline{a}^p, \underline{b}^p) = \underline{p}^{i'_n}(\underline{a}, \underline{b}) \begin{pmatrix} I_{h-d} & 0 \\ 0 & u^e I_d \end{pmatrix} \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix},$$

where $\underline{a}^p = (a_1^p, \dots, a_{h-d}^p)$ and similarly for $\underline{b}^p$. Multiplying this by $\begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}^{-1} \in \mathrm{GL}_h(k[\![u]\!])$, we obtain the equality

$$(\underline{a}, u^e \underline{b}) = \underline{p}^{1/p^n}(\underline{a}^p, \underline{b}^p) \begin{pmatrix} P_1 & P_2 \\ P_3 & P_4 \end{pmatrix}^{-1},$$

and we can write $\underline{a} = \underline{p}^{1/p^n}\underline{a}'$. The $(h-d)$-tuple $\underline{a}'$ satisfies

$$\underline{a}' = \underline{p}^{1/p^{n-1}-w}(\underline{a}')^p \hat{P}_1 - \underline{p}^{(p^n-1)/p^n-w}\underline{b}P_3\hat{P}_1.$$

Hence $v_R(\underline{a}') \geq \min\{1/p^{n-1}, (p^n-1)/p^n\} - w$ and

$$v_R(\underline{x}) \geq \min\{1/(p^{n-2}(p-1)) - w, 1 + 1/(p^n(p-1)) - w\} > w/(p-1).$$

Since $1 - w > w/(p-1)$, we obtain

$$v_R(\underline{x} + u^{e(1-w)}\underline{y}B) > w/(p-1).$$

Then Lemma 5.1 implies the reverse inclusion, and the lemma follows. $\square$

To show Theorem 1.2, we proceed by induction on n. The case of $n = 1$ follows from Lemma 5.2. Put $n \geq 2$ and suppose that the theorem holds for any truncated Barsotti–Tate groups of level $n-1$ over $\mathcal{O}_K$. Consider a truncated Barsotti–Tate group $\mathcal{G}$ of level n over $\mathcal{O}_K$ with Hodge height $w < (p-1)/p^n$, as in Theorem 1.2. In particular, we have $\mathcal{C}_{n-1} = \mathcal{G}[p^{n-1}]_{i_{n-1}} = \mathcal{G}[p^{n-1}]_{i'_{n-1}}$, and thus the inclusions $\mathcal{C}_{n-1} \subseteq \mathcal{G}_{i_n} \subseteq \mathcal{G}_{i'_n}$ also hold.

Lemma 5.3. *For any positive rational number i satisfying $i \leq 1/(p-1)$, multiplication by p induces the map $\mathcal{G}_i(\mathcal{O}_{\bar{K}}) \to \mathcal{G}[p^{n-1}]_{pi}(\mathcal{O}_{\bar{K}})$.*

Proof. By Lemma 3.1(2), we may assume that $\mathcal{G}$ is connected. By [Illusie 1985, Théorème 4.4(e)], there exists a p-divisible formal Lie group Γ over $\mathcal{O}_K$ such that $\mathcal{G}$ is isomorphic to $\Gamma[p^n]$. By [Rabinoff 2012, Lemma 11.3], we can choose formal parameters $X_1, \dots, X_d$ of the formal Lie group Γ such that the multiplication-by-p map of Γ is written as

$$[p](\mathbb{X}) \equiv p\mathbb{X} + (X_1^p, \dots, X_d^p)U + pf(\mathbb{X}) \mod \deg p^2,$$

where $\mathbb{X} = (X_1, \ldots, X_d)$, $f(\mathbb{X}) = (f_1(\mathbb{X}), \ldots, f_d(\mathbb{X}))$ such that every f_l contains no monomial of degree less than p and $U \in M_d(\mathcal{O}_K)$. Let $\underline{x} = (x_1, \ldots, x_d)$ be a d-tuple in $\mathcal{O}_{\bar{K}}$ satisfying $[p^n](\underline{x}) = 0$ and $v_p(\underline{x}) \geq i$. Since $1 + i \geq pi$, we have $1 + v_p(\underline{x}) \geq pi$ and $pv_p(\underline{x}) \geq pi$. Hence $v_p([p](\underline{x})) \geq pi$ and the lemma follows. □

Lemma 5.4. *We have the inclusion* $\mathcal{G}_{i'_n} \subseteq \mathcal{C}_n$.

Proof. By Lemma 5.2 and Lemma 5.3, multiplication by p^{n-1} induces a homomorphism $\mathcal{G}_{i'_n}(\mathcal{O}_{\bar{K}}) \to \mathcal{G}[p]_{i'_1}(\mathcal{O}_{\bar{K}}) = \mathcal{C}_1(\mathcal{O}_{\bar{K}})$. Hence we have the inclusion

$$\mathcal{G}_{i'_n} \subseteq p^{-(n-1)}\mathcal{C}_1.$$

Consider the natural map $\mathcal{G} \to \mathcal{G}/\mathcal{C}_1$. By [Hattori 2014, Theorem 1.1], the subgroup scheme $\mathcal{C}_1 \times \mathcal{S}_{1-w}$ coincides with the kernel of the Frobenius of $\mathcal{G} \times \mathcal{S}_{1-w}$. Put $\bar{\mathcal{G}} = \mathcal{G} \times \mathcal{S}_{1-w}$ and similarly for $\overline{\mathcal{G}/\mathcal{C}_1}$. Note that $pi'_n = i'_{n-1} < 1 - w$. Then we have a commutative diagram

$$\begin{tikzcd}
\mathcal{G}(\mathcal{O}_{\bar{K}}) \arrow[r] \arrow[d] & (\mathcal{G}/\mathcal{C}_1)(\mathcal{O}_{\bar{K}}) \arrow[d] & \\
\bar{\mathcal{G}}(\mathcal{O}_{\bar{K},1-w}) \arrow[r] & \overline{\mathcal{G}/\mathcal{C}_1}(\mathcal{O}_{\bar{K},1-w}) \arrow[r, hook] \arrow[d] & \bar{\mathcal{G}}^{(p)}(\mathcal{O}_{\bar{K},1-w}) \arrow[d] \\
& \overline{\mathcal{G}/\mathcal{C}_1}(\mathcal{O}_{\bar{K},pi'_n}) \arrow[r, hook] & \bar{\mathcal{G}}^{(p)}(\mathcal{O}_{\bar{K},pi'_n})
\end{tikzcd}$$

where the composite of the middle row is the Frobenius map and the right horizontal arrows are injections. From this diagram, we see that the map $\mathcal{G} \to \mathcal{G}/\mathcal{C}_1$ induces a map

$$\mathcal{G}_{i'_n}(\mathcal{O}_{\bar{K}}) \to (\mathcal{G}/\mathcal{C}_1)_{i'_{n-1}}(\mathcal{O}_{\bar{K}}).$$

This implies the inclusion $\mathcal{G}_{i'_n}/\mathcal{C}_1 \subseteq (p^{-(n-1)}\mathcal{C}_1/\mathcal{C}_1)_{i'_{n-1}}$. Note that the group scheme $p^{-(n-1)}\mathcal{C}_1/\mathcal{C}_1$ is a truncated Barsotti–Tate group of level $n-1$, height h and dimension d with Hodge height pw and that the subgroup scheme $\mathcal{C}_n/\mathcal{C}_1$ is its level $n-1$ canonical subgroup (see the proof of [Hattori 2013, Theorem 1.1] and [Hattori 2014, Theorem 1.1]). From the induction hypothesis, we see that

$$(p^{-(n-1)}\mathcal{C}_1/\mathcal{C}_1)_{i'_{n-1}} = \mathcal{C}_n/\mathcal{C}_1.$$

This implies the inclusion $\mathcal{G}_{i'_n} \subseteq \mathcal{C}_n$, and the lemma follows. □

Proposition 5.5. *The image of the map* $\mathcal{G}_{i_n}(\mathcal{O}_{\bar{K}}) \to \mathcal{G}[p^{n-1}]_{pi_n}(\mathcal{O}_{\bar{K}})$ *induced by the multiplication by* p *contains the subgroup* $\mathcal{G}[p^{n-1}]_{i_{n-1}}(\mathcal{O}_{\bar{K}})$.

Proof. By Theorem 1.1 and Lemma 5.3, we have a commutative diagram

$$\begin{tikzcd}
\mathcal{G}_{i_n}(\mathcal{O}_{\bar{K}}) \arrow[r, "\sim"'] \arrow[d, "\times p"'] & \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, I_{n,i_n}) \arrow[d, "\mathrm{pr}"] \\
\mathcal{G}[p^{n-1}]_{pi_n}(\mathcal{O}_{\bar{K}}) \arrow[r, "\sim"'] & \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, I_{n-1,pi_n}) \\
\mathcal{G}[p^{n-1}]_{i_{n-1}}(\mathcal{O}_{\bar{K}}) \arrow[r, "\sim"'] \arrow[u, hook] & \mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, I_{n-1,i_{n-1}}), \arrow[u, hook]
\end{tikzcd}$$

where the horizontal arrows are isomorphisms and the map pr is induced by the natural projection $W_n(R) \to W_{n-1}(R)$. It suffices to show that the image of the map pr contains the subgroup $\mathrm{Hom}_{\mathfrak{S},\varphi}(\mathfrak{M}, I_{n-1,i_{n-1}})$.

Let $e_1, \dots, e_h$ be a basis of the $\mathfrak{S}_n$-module $\mathfrak{M}$ lifting $\bar{e}_1, \dots, \bar{e}_h$ and $e'_1, \dots, e'_{h-d}$ be lifts of $\bar{e}'_1, \dots, \bar{e}'_{h-d}$ in $\mathfrak{M}$, respectively. Then $e'_1, \dots, e'_{h-d}, e_{h-d+1}, \dots, e_h$ also form a basis of the $\mathfrak{S}_n$-module $\mathfrak{M}$. Take a φ-compatible homomorphism of $\mathfrak{S}$-modules $\mathfrak{M} \to I_{n-1,i_{n-1}}$ defined by

$$(e'_1, \dots, e'_{h-d}, e_{h-d+1}, \dots, e_h) \mapsto (\underline{x}, \underline{y}),$$

where $\underline{x} = (x_1, \dots, x_{h-d})$ and $\underline{y}$ are an $(h-d)$-tuple and a d-tuple in the ideal $I_{n-1,i_{n-1}}$, respectively. Put $\hat{x}_l = (x_l, 0) \in W_n(R)$, $\hat{\underline{x}} = (\hat{x}_1, \dots, \hat{x}_{h-d})$ and similarly for $\hat{\underline{y}}$. Let A be the matrix in $M_h(\mathfrak{S}_n)$ satisfying

$$\varphi(e'_1, \dots, e'_{h-d}, e_{h-d+1}, \dots, e_h) = (e'_1, \dots, e'_{h-d}, e_{h-d+1}, \dots, e_h)A.$$

Define an $(h-d)$-tuple $\underline{\xi} = (\xi_1, \dots, \xi_{h-d})$ and a d-tuple $\underline{\eta}$ in R by

$$p^{n-1}([\underline{\xi}], [\underline{\eta}]) = \varphi(\hat{\underline{x}}, \hat{\underline{y}}) - (\hat{\underline{x}}, \hat{\underline{y}})A,$$

where we put $[\underline{\xi}] = ([\xi_1], \dots, [\xi_{h-d}])$ and similarly for $[\underline{\eta}]$. By Proposition 4.1, the elements $\hat{x}$ and $\hat{y}$ are divisible by $[\underline{p}^{i_{n-1}}]$ and thus we can write

$$(\underline{\xi}, \underline{\eta}) = \underline{p}^{i_{n-1}}(\underline{\xi}', \underline{\eta}').$$

Since $i_{n-1} = pi_n + w \geq pi_n$, Lemma 4.5 implies that, for any h-tuple $\underline{z}$ in R, the element $(\hat{\underline{x}}, \hat{\underline{y}}) + p^{n-1}[\underline{p}^{i_n}\underline{z}]$ is contained in the ideal I_{n,i_n}. It is enough to show that there exists an h-tuple $\underline{z}$ in R satisfying

$$\varphi((\hat{\underline{x}}, \hat{\underline{y}}) + p^{n-1}[\underline{p}^{i_n}\underline{z}]) = ((\hat{\underline{x}}, \hat{\underline{y}}) + p^{n-1}[\underline{p}^{i_n}\underline{z}])A.$$

Put $\underline{z} = (\underline{\zeta}, \underline{\omega})$ with an $(h-d)$-tuple $\underline{\zeta}$ and a d-tuple $\underline{\omega}$. Then this is equivalent to the equation

$$(\underline{\xi}, \underline{\eta}) + \underline{p}^{pi_n}(\underline{\zeta}^p, \underline{\omega}^p) = \underline{p}^{i_n}(\underline{\zeta}, \underline{\omega}) \begin{pmatrix} D & P_2 \\ 0 & u^{e(1-w)}P'_4 \end{pmatrix}.$$

We claim that the equation

$$\underline{\xi} + \underline{p}^{pi_n} \underline{\zeta}^p = \underline{p}^{i_n} \underline{\zeta} D$$

for the first entry has a solution $\underline{\zeta} = \underline{p}^{(p-1)i_n} \underline{\zeta}'$ with an $(h-d)$-tuple $\underline{\zeta}'$ in R. Indeed, let $\hat{D} \in M_{h-d}(k[\![u]\!])$ be the matrix satisfying $D\hat{D} = u^{ew} I_{h-d}$. Then this is equivalent to the equation

$$\underline{\zeta}' = \underline{\xi}' \hat{D} + \underline{p}^{p(p-1)i_n - w} (\underline{\zeta}')^p \hat{D}.$$

Since $p(p-1)i_n > w$, we can find a solution $\underline{\zeta}'$ of the equation by recursion.

For the second entry, we have the equation

$$\underline{p}^{pi_n + w} \underline{\eta}' + \underline{p}^{pi_n} \underline{\omega}^p = \underline{p}^{i_n} (\underline{\zeta} P_2 + \underline{p}^{1-w} \underline{\omega} P_4').$$

This is equivalent to the equation

$$\underline{\omega}^p = \underline{p}^{1-w-(p-1)i_n} \underline{\omega} P_4' + \underline{\zeta}' P_2 - \underline{p}^w \underline{\eta}'.$$

Note that $1 - w \geq (p-1)i_n$. Write this equation as

$$(\omega_1^p, \ldots, \omega_d^p) + (\omega_1, \ldots, \omega_d) C + (c_1', \ldots, c_d') = 0$$

with some $C = (c_{i,j}) \in M_d(R)$ and $c_i' \in R$. Then the R-algebra

$$R[\omega_1, \ldots, \omega_d] \Big/ \left(\omega_1^p + \sum_{j=1}^{d} c_{j,1} \omega_j + c_1', \ldots, \omega_d^p + \sum_{j=1}^{d} c_{j,d} \omega_j + c_d' \right)$$

is free of rank p^d over R. Since $\mathrm{Frac}(R)$ is algebraically closed and R is integrally closed, this R-algebra admits at least one R-valued point. Hence we can find at least one solution $\underline{\omega}$ of the equation. This concludes the proof of the proposition. $\square$

Consider the exact sequence

$$0 \to \mathcal{G}[p]_{i_n}(\mathcal{O}_{\bar{K}}) \to \mathcal{G}_{i_n}(\mathcal{O}_{\bar{K}}) \xrightarrow{\times p} \mathcal{G}[p^{n-1}]_{pi_n}(\mathcal{O}_{\bar{K}}).$$

Proposition 5.5 implies that the image of the rightmost arrow contains the subgroup

$$\mathcal{G}[p^{n-1}]_{i_{n-1}}(\mathcal{O}_{\bar{K}}) \subseteq \mathcal{G}[p^{n-1}]_{pi_n}(\mathcal{O}_{\bar{K}}),$$

which coincides with $\mathcal{C}_{n-1}(\mathcal{O}_{\bar{K}})$ by induction hypothesis and thus is of order $p^{(n-1)d}$. By Lemma 5.2, the subgroup $\mathcal{G}[p]_{i_n}(\mathcal{O}_{\bar{K}})$ also coincides with $\mathcal{C}_1(\mathcal{O}_{\bar{K}})$ and this is of order p^d. Hence the group $\mathcal{G}_{i_n}(\mathcal{O}_{\bar{K}})$ is of order no less than p^{nd}. Since Lemma 5.4 implies the inclusions

$$\mathcal{G}_{i_n}(\mathcal{O}_{\bar{K}}) \subseteq \mathcal{G}_{i_n'}(\mathcal{O}_{\bar{K}}) \subseteq \mathcal{C}_n(\mathcal{O}_{\bar{K}}),$$

Theorem 1.2 follows by comparing orders. $\square$

To prove Theorem 1.3, we need the following lemma, which is a "lower" variant of [Hattori 2013, Lemma 4.5].

Lemma 5.6. *Let $K/\mathbb{Q}_p$ be an extension of complete discrete valuation fields and i be a positive rational number. Let $\mathfrak{X}$ be an admissible formal scheme over* $\mathrm{Spf}(\mathcal{O}_K)$ *and X be its Raynaud generic fiber. Let $\mathfrak{G}$ be a finite locally free formal group scheme over $\mathfrak{X}$ with Raynaud generic fiber G. Then there exists an admissible open subgroup G_i of G over X such that the open immersion $G_i \to G$ is quasicompact and that for any finite extension L/K and $x \in X(L)$, the fiber $(G_i)_x$ coincides with the lower ramification subgroup* $(\mathfrak{G}_x)_i \times \mathrm{Spec}(L)$ *of the finite flat group scheme* $\mathfrak{G}_x = \mathfrak{G} \times_{\mathfrak{X},x} \mathrm{Spf}(\mathcal{O}_L)$ *over $\mathcal{O}_L$.*

Proof. Let $\mathscr{I}$ be the augmentation ideal sheaf of the formal group scheme $\mathfrak{G}$. Write $i = m/n$ with positive integers m, n and put $\mathscr{J} = p^m\mathcal{O}_{\mathfrak{G}} + \mathscr{I}^n$. Let $\mathfrak{B}$ be the admissible blow-up of $\mathfrak{G}$ along the ideal $\mathscr{J}$ and $\mathfrak{G}_{m,n}$ be the formal open subscheme of $\mathfrak{B}$ where p^m generates the ideal $\mathscr{J}\mathcal{O}_{\mathfrak{B}}$. Since the Raynaud generic fiber of $\mathfrak{G}_{m,n}$ is the admissible open subset of G whose set of $\bar{K}$-valued points is given by

$$\{x \in G(\bar{K}) \mid v_p(\mathscr{I}(x)) \geq i\},$$

it is independent of the choice of m, n, and we write it as G_i. Using the universality of dilatations as in the proof of [Abbes and Mokrane 2004, Proposition 8.2.2], we can show that G_i is an admissible open subgroup of the rigid-analytic group G. For any affinoid open subset $U = \mathrm{Sp}(A)$ of G, put $I = \Gamma(U, \mathscr{I})$. Then the intersection $U \cap G_i$ is the affinoid $\mathrm{Sp}(A\langle I^n/p^m\rangle)$ and thus the open immersion $G_i \to G$ is quasicompact. This concludes the proof of the lemma. □

Proof of Theorem 1.3. Set C_n to be the admissible open subgroup $G_{i'_n}$ of G as in Lemma 5.6 with $i'_n = 1/(p^n(p-1))$. Then, by this lemma and Theorem 1.2, each fiber $(C_n)_x$ coincides with the generic fiber of the level n canonical subgroup of $\mathfrak{G}_x$, and its group of $\bar{K}$-valued points is isomorphic to the group $(\mathbb{Z}/p^n\mathbb{Z})^d$. Moreover, C_n is etale, quasicompact and separated over $X(r_n)$. Thus [Conrad 2006, Theorem A.1.2] implies that C_n is finite over $X(r_n)$, and the theorem follows by a similar argument to the proof of [Hattori 2013, Corollary 1.2]. □

Acknowledgments

This work was supported by JSPS KAKENHI grant no. 23740025.

References

[Abbes and Mokrane 2004] A. Abbes and A. Mokrane, "Sous-groupes canoniques et cycles évanescents p-adiques pour les variétés abéliennes", *Publ. Math. Inst. Hautes Études Sci.* 99 (2004), 117–162. MR 2005f:14090 Zbl 1062.14057

[Berthelot et al. 1982] P. Berthelot, L. Breen, and W. Messing, *Théorie de Dieudonné cristalline, II*, Lecture Notes in Mathematics **930**, Springer, Berlin, 1982. MR 85k:14023 Zbl 0516.14015

[Breuil 2000] C. Breuil, "Groupes p-divisibles, groupes finis et modules filtrés", *Ann. of Math.* (2) **152**:2 (2000), 489–549. MR 2001k:14087 Zbl 1042.14018

[Breuil 2002] C. Breuil, "Integral p-adic Hodge theory", pp. 51–80 in *Algebraic geometry 2000, Azumino* (Hotaka, 2000), edited by S. Usui et al., Adv. Stud. Pure Math. **36**, Math. Soc. Japan, Tokyo, 2002. MR 2004e:11135 Zbl 1046.11085

[Conrad 2006] B. Conrad, "Modular curves and rigid-analytic spaces", *Pure Appl. Math. Q.* **2**:1 (2006), 29–110. MR 2007a:14026 Zbl 1156.14312

[Faltings 1999] G. Faltings, "Integral crystalline cohomology over very ramified valuation rings", *J. Amer. Math. Soc.* **12**:1 (1999), 117–144. MR 99e:14022 Zbl 0914.14009

[Fargues 2011] L. Fargues, "La filtration canonique des points de torsion des groupes p-divisibles", *Ann. Sci. Éc. Norm. Supér.* (4) **44**:6 (2011), 905–961. MR 2919687 Zbl 06026211

[Fontaine 1994] J.-M. Fontaine, "Le corps des périodes p-adiques", pp. 59–111 Astérisque **223**, Société Mathématique de France, Paris, 1994. MR 95k:11086 Zbl 0940.14012

[Gabriel 1965] P. Gabriel, "Étude infinitésimale des schémas en groupe et groupes formels", pp. 1–65+4 in *Schémas en Groupes (Sém. Géométrie Algébrique, Inst. Hautes Études Sci., 1963/64), Fasc. 2b, Exposé 7a*, Inst. Hautes Études Sci., Paris, 1965. MR 35 #4221

[Hattori 2012] S. Hattori, "Ramification correspondence of finite flat group schemes over equal and mixed characteristic local fields", *J. Number Theory* **132**:10 (2012), 2084–2102. MR 2944746 Zbl 1270.14021

[Hattori 2013] S. Hattori, "Canonical subgroups via Breuil–Kisin modules", *Math. Z.* **274**:3-4 (2013), 933–953. MR 3078253 Zbl 06197119

[Hattori 2014] S. Hattori, "Canonical subgroups via Breuil–Kisin modules for $p = 2$", *J. Number Theory* **137** (2014), 142–159. MR 3157783 Zbl 06256935

[Illusie 1985] L. Illusie, "Déformations de groupes de Barsotti–Tate (d'après A. Grothendieck)", pp. 151–198 in *Seminar on arithmetic bundles: The Mordell conjecture* (Paris, 1983/84), Astérisque **127**, Société Mathématique de France, Paris, 1985. MR 801922 Zbl 1182.14050

[de Jong and Messing 1999] A. J. de Jong and W. Messing, "Crystalline Dieudonné theory over excellent schemes", *Bull. Soc. Math. France* **127**:2 (1999), 333–348. MR 2001b:14075 Zbl 0963.14008

[Kim 2012] W. Kim, "The classification of p-divisible groups over 2-adic discrete valuation rings", *Math. Res. Lett.* **19**:1 (2012), 121–141. MR 2923180 Zbl 06194639

[Kisin 2006] M. Kisin, "Crystalline representations and F-crystals", pp. 459–496 in *Algebraic geometry and number theory*, edited by V. Ginzburg, Progr. Math. **253**, Birkhäuser, Boston, MA, 2006. MR 2007j:11163 Zbl 1184.11052

[Kisin 2009] M. Kisin, "Modularity of 2-adic Barsotti–Tate representations", *Invent. Math.* **178**:3 (2009), 587–634. MR 2010k:11089 Zbl 05636297

[Lau 2010] E. Lau, "A relation between Dieudonné displays and crystalline Dieudonné theory", preprint, 2010. arXiv 1006.2720v2

[Liu 2008] T. Liu, "On lattices in semi-stable representations: A proof of a conjecture of Breuil", *Compos. Math.* **144**:1 (2008), 61–88. MR 2009c:14087 Zbl 1133.14020

[Liu 2013] T. Liu, "The correspondence between Barsotti–Tate groups and Kisin modules when $p = 2$", *J. Théor. Nombres Bordeaux* **25**:3 (2013), 661–676.

[Rabinoff 2012] J. Rabinoff, "Higher-level canonical subgroups for p-divisible groups", *J. Inst. Math. Jussieu* **11**:2 (2012), 363–419. MR 2905309 Zbl 06028609

[Tian 2010] Y. Tian, "Canonical subgroups of Barsotti–Tate groups", *Ann. of Math.* (2) **172**:2 (2010), 955–988. MR 2012a:14105 Zbl 1203.14026

[Tian 2012] Y. Tian, "An upper bound on the Abbes–Saito filtration for finite flat group schemes and applications", *Algebra Number Theory* **6**:2 (2012), 231–242. MR 2950153 Zbl 1260.14055

Communicated by Brian Conrad
Received 2012-10-12 Revised 2013-11-18 Accepted 2013-11-19

shin-h@math.kyushu-u.ac.jp *Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.331

Wild models of curves

Dino Lorenzini

Let K be a complete discrete valuation field with ring of integers $\mathcal{O}_K$ and algebraically closed residue field k of characteristic $p > 0$. Let X/K be a smooth proper geometrically connected curve of genus $g > 0$ with $X(K) \neq \varnothing$ if $g = 1$. Assume that X/K does not have good reduction and that it obtains good reduction over a Galois extension L/K of degree p. Let $\mathcal{Y}/\mathcal{O}_L$ be the smooth model of X_L/L. Let $H := \mathrm{Gal}(L/K)$.

In this article, we provide information on the regular model of X/K obtained by desingularizing the wild quotient singularities of the quotient $\mathcal{Y}/H$. The most precise information on the resolution of these quotient singularities is obtained when the special fiber $\mathcal{Y}_k/k$ is ordinary. As a corollary, we are able to produce for each odd prime p an infinite class of wild quotient singularities having pairwise distinct resolution graphs. The information on the regular model of X/K also allows us to gather insight into the p-part of the component group of the Néron model of the Jacobian of X.

1. Introduction

Let K be a complete discrete valuation field with valuation v, ring of integers $\mathcal{O}_K$ and residue field k of characteristic $p > 0$, assumed to be *algebraically closed*. Let X/K be a smooth proper geometrically connected curve of genus $g > 0$ with $X(K) \neq \varnothing$ if $g = 1$.

Assume that X/K does not have good reduction and that it obtains good reduction over a Galois extension L/K. Let $\mathcal{Y}/\mathcal{O}_L$ be the smooth model of X_L/L. Let $H := \mathrm{Gal}(L/K)$, and let $\mathcal{Z}/\mathcal{O}_K$ denote the quotient $\mathcal{Y}/H$. A regular model for X/K can be obtained by resolving the singularities of the scheme $\mathcal{Z}$. Our goal is to obtain information on this regular model when p divides $[L : K]$. Since the presence of wild ramification renders the subject quite challenging, we will restrict our attention in this article to the case where $[L : K] = p$.

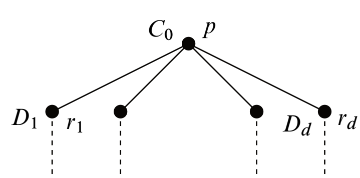
Lorenzini was supported by NSF Grant 0902161.
MSC2010: primary 14G20; secondary 14G17, 14K15, 14J17.
Keywords: model of a curve, ordinary curve, cyclic quotient singularity, wild ramification, arithmetical tree, resolution graph, component group, Néron model.

Beyond our interest in models of curves per se, our motivation for understanding these regular models is twofold. First, since $\mathcal{X}$ is obtained by desingularizing certain quotient singularities, we hope to gain more insight in the general theory of resolutions of wild quotient singularities by producing interesting classes of examples where the singularities can be resolved explicitly. Second, since from a regular model of the curve one can compute much of the Néron model of its Jacobian, we hope to bring new insight into the structure of the rather mysterious p-part of the component group of the Néron model of a general abelian variety from an increased understanding of the special case of Jacobians of curves.

Let us introduce some notation needed to state our theorems. Let σ denote a generator of $H := \mathrm{Gal}(L/K)$. Denote also by σ the automorphism of $\mathcal{Y}_k$ induced by the action of H on $\mathcal{Y}$. The scheme $\mathcal{Z}$ is singular exactly at the images $Q_1, \ldots, Q_d$ of the ramification points $P_1, \ldots, P_d$ of the map $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ (5.2). Consider the regular model $\mathcal{X} \to \mathcal{Z}$ obtained from $\mathcal{Z}$ by a minimal desingularization. Let $\mathcal{X}' \to \mathcal{X}$ denote the regular model of X/K minimal with the property that $\mathcal{X}'_k$ has smooth components and normal crossings. Let f denote the composition $\mathcal{X}' \to \mathcal{Z}$. Let C_0/k denote the strict transform in $\mathcal{X}'$ of the irreducible closed subscheme $\mathcal{Z}_k^{\mathrm{red}}$ of $\mathcal{Z}$. Let $D_1, \ldots, D_d$ denote the irreducible components of $\mathcal{X}'_k$ that meet C_0. Let r_i denote the multiplicity of D_i, $i = 1, \ldots, d$, in $\mathcal{X}'_k$.

Recall that to any connected curve $\bigcup_{\ell=1}^{n} C_\ell$ on a regular model $\mathcal{X}$ we associate a graph G as follows: the vertices are the irreducible components C_ℓ, and in G, the vertices C_i and C_j $(i \neq j)$ are linked by exactly $(C_i \cdot C_j)_{\mathcal{X}}$ edges, where $(C_i \cdot C_j)_{\mathcal{X}}$ denotes the intersection number of C_i and C_j on the regular scheme $\mathcal{X}$. Recall that the *degree* of a vertex v of a graph is the number of edges attached to v. A *node* on a graph is a vertex of degree at least 3. A vertex of degree 1 is a *terminal vertex*. A *chain* is a subgraph of G with vertices $C_0, C_1, \ldots, C_n$, $n \geq 1$, such that C_i is linked to C_{i+1} by exactly one edge in G when $i = 0, \ldots, n-1$ and the degree of C_i is 2 when $i = 1, \ldots, n-1$. If the chain contains a terminal vertex (which can only be C_0 or C_n), the chain is called a *terminal chain*.

Let G denote the graph associated with $\mathcal{X}'_k$. We assume $d \geq 1$. For each $i = 1, \ldots, d$, let G_{Q_i} denote the graph associated with the curve $f^{-1}(Q_i)$. In particular, D_i corresponds to a vertex of G_{Q_i}. We have the following configuration on the graph G (where a positive integer next to a vertex denotes the multiplicity of the corresponding irreducible component in $\mathcal{X}'$):



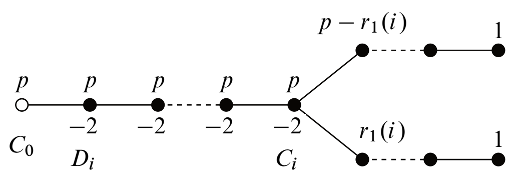
Theorem 5.3. *Let X/K be a curve with potentially good reduction after a wildly ramified extension L/K of degree p, as above. Keep the above notation. Then, for all $i = 1, \dots, d$, the graph G_{Q_i} contains a node of G and p divides r_i.*

In contrast, when H is of prime order $q \neq p$, then it is known that $q > r_i$ and that the graph G_{Q_i} does not contain a node of G. In particular, when L/K is tame and $d \geq 3$, the graph G has only a single node, the component C_0 (see, e.g., [Lorenzini 1990a, Theorem 2.1]).

We propose in 6.1 a combinatorial measure $\gamma_{Q_i} g_{Q_i}$ of the complexity of the graph G_{Q_i}, which we conjecturally relate in 6.2 to the higher ramification data of the morphism $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$. This conjectural relationship expresses the fact that the graph G_{Q_i} is "complicated" only if the higher ramification above Q_i is "large". We prove this conjecture in the ordinary case (Theorem 6.4).

Recall that a smooth proper curve Y/k of genus g is called *ordinary* if its Jacobian J/k is an ordinary abelian variety (that is, $J(k)$ has exactly p^g points of order dividing p). When $\mathcal{Y}_k$ is ordinary, the morphism $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ has the smallest possible ramification data at each Q_i (2.2), and in this case, we can use Theorem 5.3 to describe the graph G_{Q_i} explicitly, as in the following theorem, whose statement is slightly strengthened in the version given in Section 6. In the graph below, a bullet $\bullet$ represents an irreducible component of the desingularization of Q_i. A negative number next to a vertex is the self-intersection of the component. A positive number next to a vertex is the multiplicity of the corresponding component in $\mathcal{X}'_k$.

Theorem (see Theorem 6.8). *Let X/K be a curve with potentially good reduction after a Galois extension L/K of degree p, as above. Assume that $\mathcal{Y}_k$ ordinary. Then, for all $i = 1, \dots, d$, we have $r_i = p$ and G_{Q_i} is a graph with a single node C_i of degree 3:*



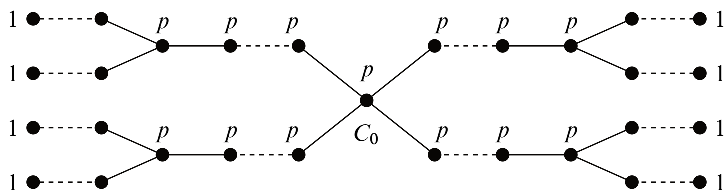
The intersection matrix $N(p, \alpha_i, r_1(i))$ of the resolution of Q_i is uniquely determined as in 4.7 by the two integers α_i and $r_1(i)$ with $1 \leq r_1(i) < p$. The integer α_i denotes the number of vertices of self-intersection -2 (including the node C_i) on the chain in G_{Q_i} connecting the node C_0 to the single node C_i of G_{Q_i}, and the integer α_i is divisible by p.

To further determine the regular model, one would need to determine explicitly the integers α_i and $r_1(i)$. We address this issue in [Lorenzini 2013b]. In all cases where we have been able to compute α_i and $r_1(i)$, we found them to be related to the valuation of the different of L/K. More precisely, let $(s_{L/K} + 1)(p - 1)$

denote the valuation of the different of L/K. In [Lorenzini 2013b, Theorem 1.1], we present some instances where $\alpha_i = ps_{L/K}$ and $r_1(i) \equiv -s_{L/K}^{-1}$ modulo p. We also show in [Lorenzini 2013b, Theorem 4.1] that the singularities Q_i are rational.

Remark 1.1. The same type of intersection matrix, $N(p, \alpha_i, r_1(i))$, also occurs in the resolution of the singularities of the model $\mathcal{Z}$ when X/K has genus $p-1$ and $\mathrm{Jac}(X)/K$ has purely toric reduction after an extension of degree p [Lorenzini 2010, Theorem 2.2].

Remark 1.2. The special fiber of the model $\mathcal{X}/\mathcal{O}_K$ of X/K in Theorem 6.8 has thus a graph with a central vertex to which d branches are attached, of the form described below, where we picture the case $d = 4$.



Fix any $d > 1$. We establish in Theorem 6.8 and Example 6.13 the existence of some field K of residue characteristic $p > 0$ and of some smooth proper curve X/K with a regular model whose special fiber has a graph of the above type. This is clearly a weak existence result, but our understanding of models in the presence of wild ramification is so limited that even this weak existence result does not follow from the general existence results of Viehweg [1977] and Winters [1974].

An immediate but surprising corollary to Theorem 6.8 is as follows.

Corollary (see Corollary 6.10). *Let X/K be a curve of genus $g > 1$ with potentially good reduction after a Galois extension L/K of degree p, as above. Assume that $\mathcal{Y}_k$ is ordinary. Then $X(K) \neq \varnothing$.*

The information on the regular model of X/K obtained in Theorem 6.8, while incomplete to fully describe the special fiber of the model, suffices to compute several invariants of arithmetical interest. For instance, the set of components of multiplicity 1 on the special fiber of the model is determined, and this information is one of the ingredients needed to apply the Chabauty–Coleman method to bound the number of $\mathbb{Q}$-rational points on a curve $X/\mathbb{Q}$ using the reduction at a small prime p, as in [Lorenzini and Tucker 2002, Theorem 1.1]. Let A/K denote the Jacobian of A/K with Néron model $\mathcal{A}/\mathcal{O}_K$ and component group $\Phi_{A/K}$. The information obtained in Theorem 6.8 suffices to compute $\Phi_{A/K}$ and a new canonical subgroup $\Phi^0_{A/K}$ of $\Phi_{A/K}$ that we now define.

1.3. Let A/K be an abelian variety with Néron model $\mathcal{A}/\mathcal{O}_K$. Let L/K be any finite extension, and let $\mathcal{A}'/\mathcal{O}_L$ denote the Néron model of A_L/L. Denote by

$$\eta : \mathcal{A} \times_{\mathcal{O}_K} \mathcal{O}_L \to \mathcal{A}'$$

the canonical map induced by the functoriality property of Néron models. The special fiber $\mathcal{A}_k$ is an extension of a finite group $\Phi_{A/K}$, called the *group of components*, by the *connected component of zero* $\mathcal{A}_k^0$ of $\mathcal{A}_k$:

$$0 \to \mathcal{A}_k^0 \to \mathcal{A}_k \to \Phi_{A/K} \to 0.$$

Assume that A_L/L has semistable reduction, and consider the natural map $\Phi_{A/K} \to \mathcal{A}'_k/\eta(\mathcal{A}_k^0)$. We let

$$\Phi^0_{A/K} := \mathrm{Ker}(\Phi_{A/K} \to \mathcal{A}'_k/\eta(\mathcal{A}_k^0)).$$

The subgroup $\Phi^0_{A/K}$ does not depend on the choice of such an extension L/K and is functorial in A. Our interest in this subgroup stems from the following conjectures.

When A/K has potentially good reduction and, more generally, when the toric rank of $\mathcal{A}_k^0$ is trivial, we conjecture that the order of the group $\Phi_{A/K}$ is bounded by a constant depending only on the dimension g of A/K [Lorenzini 1990b, p. 146]. This statement is true when A/K is a Jacobian [Lorenzini 1990b, Theorem 2.4] and for the prime-to-p part of $\Phi_{A/K}$ [Lorenzini 1993, Theorem 2.15]. Since $[L:K]^2$ kills the group $\Phi_{A/K}$ when the toric rank of $\mathcal{A}$ is trivial [Liu and Lorenzini 2001, Proposition 1.8], we find that, to prove the conjecture that $\Phi_{A/K}$ is bounded by a constant depending only on g, it suffices to prove that the minimal number of generators of $\Phi_{A/K}$ can be bounded by a constant depending on g only. We guess, under the above hypotheses, that $\Phi_{A/K}$ can be generated by $2g$ elements.

Assume now that A/K has potentially good reduction. The p-torsion in $\mathcal{A}'_k$ can always be generated by at most g elements. Thus, the above conjecture is proved if the p-part of the kernel $\Phi^0_{A/K}$ can be generated by a number of elements bounded by a constant depending on g only (possibly $2g$). In the ordinary case, where the p-torsion in $\mathcal{A}'_k$ is minimally generated by g elements, one may wonder if $\Phi^0_{A/K}$ can also be generated by g elements. Our next corollary gives some evidence that this latter question may have a positive answer for all abelian varieties with potentially good ordinary reduction.

Let A/K be the Jacobian of a curve X/K with $X(K) \neq \varnothing$. We denote by $\langle\cdot,\cdot\rangle : \Phi_{A/K} \times \Phi_{A/K} \to \mathbb{Q}/\mathbb{Z}$ Grothendieck's pairing, which is nondegenerate [Bosch and Lorenzini 2002, Theorem 4.6]. We denote by $(\Phi^0_{A/K})^\perp$ the orthogonal of $\Phi^0_{A/K}$ under Grothendieck's pairing.

Corollary (see Corollary 6.12). *Let A/K be the Jacobian of a curve X/K of genus $g > 1$ having potentially good ordinary reduction after a Galois extension L/K of degree p, as above. Then $\Phi_{A/K}$ is a $\mathbb{Z}/p\mathbb{Z}$-vector space of dimension $2d-2$, and $\Phi^0_{A/K}$ is a subspace of dimension $d-1$. Moreover, $\Phi^0_{A/K} = (\Phi^0_{A/K})^\perp$.*

It is natural in view of Corollary 6.12 to wonder whether the same result holds for all principally polarized abelian varieties A/K having potentially good ordinary reduction after a Galois extension L/K of degree p. We may also wonder, for any principally polarized abelian variety A/K with potentially good reduction, whether the order of $\Phi^0_{A/K} \cap (\Phi^0_{A/K})^\perp$ can be bounded by a constant depending only on the p-rank of $\mathcal{A}'_k$. We hope to return to these questions in the future.

1.4. Our explicit computation of a regular model of a curve having potentially good ordinary reduction also has an application to quotient singularities. Our current understanding of wild $\mathbb{Z}/p\mathbb{Z}$-quotient singularities of surfaces is quite limited, and few explicit examples are known (see, e.g., [Artin 1975], [Katsura 1978] for $p = 2$ and [Peskin 1983] for $p = 3$). In contrast to the case of a tame cyclic quotient singularity, where the number of possible resolution graphs is finite once the order of the group is fixed, we show below that, for any fixed odd prime p, there are infinitely many graphs that can occur as the resolution graphs of a wild $\mathbb{Z}/p\mathbb{Z}$-quotient singularity in both mixed characteristic and in the equicharacteristic case. The analogous result when $p = 2$ is discussed in [Lorenzini 2013a, Theorem 4.1].

Corollary 6.14. *Fix any odd prime p. For each integer $m > 0$, there exist a 2-dimensional regular local ring B of equicharacteristic p endowed with an action of $H := \mathbb{Z}/p\mathbb{Z}$ and a 2-dimensional regular local ring B' of mixed characteristic $(0, p)$ endowed with an action of $\mathbb{Z}/p\mathbb{Z}$ such that $\operatorname{Spec} B^H$ and $\operatorname{Spec}(B')^H$ are singular exactly at their closed point and the graphs associated with a minimal resolution of $\operatorname{Spec} B^H$ and $\operatorname{Spec}(B')^H$ have one node and more than m vertices.*

This article is organized as follows. The proof of Theorem 5.3, in Section 5, is of a global nature and includes in particular a study of the natural map $\Phi_{A/K} \to \mathcal{A}'_k/\eta(\mathcal{A}^0_k)$. The proof uses two auxiliary results of independent interest. The first result, Proposition 2.5, is discussed in Section 2 and is a relation between torsion points in a quotient of two Jacobians. This proposition is one place in our arguments where the tame and wild cases can be seen to differ in an explicit way. The second result, Proposition 3.5, is the main result of Section 3 and is a general relation between elements in the component group Φ_M of an arithmetical tree.

Section 4 presents further results of a combinatorial nature on arithmetical trees that are needed in the proof of Theorem 6.8. Section 6 contains the proof of Theorem 6.8 and of its applications.

2. Cyclic morphisms and torsion

Let k be an algebraically closed field of characteristic p. Let $f : D \to C$ be a ramified Galois morphism of smooth connected projective curves over k. Our main result in this section is Proposition 2.5, which will be applied to the case of the quotient morphism $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ in the course of the proof of Theorem 5.3.

2.1. Assume that the Galois group H of f is cyclic of degree q^s with q prime. Let $P_1, \dots, P_d$ in $D(k)$ be the ramification points. Assume that, at each P_i, the morphism is totally ramified, and let $Q_i := f(P_i)$, $i = 1, \dots, d$, be the branch points.

When $q \neq p$, the Riemann–Hurwitz formula is

$$2g(D) - 2 = q^s(2g(C) - 2) + d(q^s - 1). \tag{2.1.1}$$

Moreover, $d \geq 2$. When $g(C) = 0$, this follows immediately from the formula; the general case requires a separate proof.

Assume now that $q = p$. For $P \in D(k)$, let $H_0(P) \supseteq H_1(P) \supseteq \cdots$ denote the sequence of higher ramification groups. If P is a ramification point, then $|H_0(P)| = |H_1(P)| = p^s$. Set

$$\delta(P) := \sum_i (|H_i(P)| - 1).$$

Then the *Riemann–Hurwitz formula* is

$$2g(D) - 2 = p^s(2g(C) - 2) + \sum_{P \in D(k)} \delta(P), \tag{2.1.2}$$

and it may happen that $d = 1$.

2.2. Let $\gamma(D)$ denote the p-rank of D (i.e., the p-rank of $\mathrm{Jac}(D)$). The *Deuring–Shafarevich formula* relates the p-ranks of C and D:

$$\gamma(D) - 1 = p^s(\gamma(C) - 1) + d(p^s - 1). \tag{2.2.1}$$

The curve D is *ordinary* when $\gamma(D) = g(D)$. When D is ordinary, we find, comparing the formulas (2.1.2) and (2.2.1), that $|H_2(P)| = 1$ for all P and that C is also ordinary. Moreover, when $g(D) > 0$, Equation (2.2.1) shows that $p \leq g(D) + 1$.

When a ramification point P of a Galois morphism $f : D \to C$ is such that $H_2(P) = (0)$, we will say that the morphism is *weakly ramified* at P.

2.3. We record here the following well-known fact (see [Hasse 1934, p. 42], or [Singh 1974, Lemma 1.3], when $K = k(x)$). Let K be a field with $\mathrm{char}(K) = p$. Let $(A, \mathcal{M})$ be a discrete valuation ring with field of fractions K, valuation v_K and uniformizer π_K. Assume that the residue field k of A is algebraically closed. Let L/K be a cyclic ramified Galois extension of degree p with Galois group H. Let $(B, \mathcal{N})$ denote the integral closure of A in L. Let $H = H_0 \supseteq H_1 \supseteq \cdots$ denote the sequence of ramification groups. Then $\sum_{i=0}^{\infty}(|H_i| - 1) = (m+1)(p-1)$ for some integer m prime to p.

2.4. Examples of curves with an automorphism of degree p in characteristic p can be given in Artin–Schreier form. Consider the curve $y^p - y = \prod_{i=1}^{d}(x-a_i)^{-n_i}$, where $a_1, \dots, a_s \in k$ are distinct and the n_i are positive integers coprime to p. The automorphism $y \mapsto y+1$ has order p. The genus g of the smooth complete curve defined by the above equation is given by the Riemann–Hurwitz formula $2g-2 = -2p + (p-1)\big(\sum_{i=1}^{d}(n_i+1)\big)$ (see [Subrao 1975, p. 8]).

The following simple proposition exhibits a key difference between the tame and wild cases:

Proposition 2.5. *Let q be a prime. Let $f : D \to C$ be a ramified cyclic morphism of degree q^s between smooth connected projective curves over k. Let $P_1, \dots, P_d$, $d \geq 2$, denote the ramification points, assumed to be totally ramified. For $i \neq j$, the image ω_{ij} of $P_i - P_j$ in $\mathrm{Jac}(D)/f^*(\mathrm{Jac}(C))$ is of finite order q^s. Let T denote the finite subgroup $\mathrm{Jac}(D)/f^*(\mathrm{Jac}(C))$ generated by $\{\omega_{id} \mid i = 1, \dots, d-1\}$.*

(a) *If $q = p$, then T is isomorphic to $(\mathbb{Z}/p^s\mathbb{Z})^{d-1}$ and is generated by the set $\{\omega_{id} \mid i = 1, \dots, d-1\}$.*

(b) *If $q \neq p$, then T is isomorphic to $(\mathbb{Z}/q^s\mathbb{Z})^{d-2}$ and is generated by the set $\{\omega_{id} \mid i = 1, \dots, d-2\}$.*

Proof. Let S denote the subgroup of $\mathrm{Div}^0(D)$ with support on the set $\{P_1, \dots, P_d\}$. It is clear that $\{P_i - P_d \mid i = 1, \dots, d-1\}$ is a $\mathbb{Z}$-basis for S. Let $S \to T$ denote the natural surjective map. This map factors through S/q^sS since $q^s\big(\sum_i b_iP_i\big) = f^*\big(\sum_i b_iQ_i\big)$ with $\sum_i b_iQ_i \in \mathrm{Div}^0(C)$.

Let σ be a generator of $\mathrm{Aut}(D/C)$. Suppose that $\sigma(\mathrm{div}_D(g)) = \mathrm{div}_D(g)$ for some $g \in k(D)^*$. Then $g^\sigma = cg$ for some $c \in k^*$. Since σ has finite order q^s, we find that $c^{q^s} = 1$.

Consider first the case where $q = p$. Then $c = 1$. Thus, $g^\sigma = g$ and $g \in k(C)^*$. Suppose that the divisor $\big(\sum_i b_iP_i\big)$ has trivial image in T. Then it is possible to write $\big(\sum_i b_iP_i\big) = f^*\big(\sum_j R_j\big) + \mathrm{div}_D(h)$ for some $R_j \in C(k)$ and $h \in k(D)^*$. Then we have $\sigma(\mathrm{div}_D(h)) = \mathrm{div}_D(h)$, and we conclude that $h \in k(C)^*$. Therefore, we have an equality of divisors of the form $\big(\sum_i b_iP_i\big) = f^*(E)$ for some $E \in \mathrm{Div}^0(C)$. It follows that $E = \sum_i c_iQ_i$ for some c_i. Hence, the map $S/p^sS \to T$ is an isomorphism, proving Part (a).

Suppose now that $q \neq p$. Fix a primitive q^s-th root ξ of 1. Then $k(D)/k(C)$ is a Kummer extension, generated by the root α of $y^{q^s} - a \in k(C)[y]$ such that $\alpha^\sigma = \xi\alpha$. It is easy to check that, for each $i = 0, \dots, q^s - 1$,

$$\{\beta \in k(D) \mid \beta^\sigma = \xi^i\beta\} = k(C)\alpha^i.$$

The equality $\alpha^\sigma = \xi\alpha$ implies that $\mathrm{div}_D(\alpha)$ can be written as

$$\Big(\sum_{i=1}^{d} a_iP_i\Big) + \sum_j c_j\Big(\sum_{i=0}^{q^s-1}\sigma^i(S_j)\Big)$$

for some integers a_i and c_j, and some $S_j \in D(k) \setminus \{P_1, \ldots, P_d\}$. It follows that q^s divides $\sum_{i=1}^{d} a_i$ since $\deg(\mathrm{div}_D(\alpha)) = 0$. This means that the divisor

$$\sum_j c_j \Big(\sum_{i=0}^{q^s-1} \sigma^i(S_j) \Big) + \Big(\sum_{i=1}^{d} a_i \Big) P_d$$

defines an element in $f^*(\mathrm{Jac}(C))$. Hence, the image ν of $\big(\sum_i a_i P_i\big) - \big(\sum_i a_i\big) P_d$ in T is trivial. We thus have a map

$$\varphi : S \Big/ \Big\langle q^s S, \Big(\sum_i a_i P_i\Big) - \Big(\sum_i a_i\Big) P_d \Big\rangle \to T.$$

Let us note that $\big(\sum_i a_i P_i\big) - \big(\sum_i a_i\big) P_d \notin q^s S$ because, otherwise, the morphism f given by the Kummer equation $y^{q^s} - a$ would not be totally ramified at $P_1, \ldots, P_d$.

Suppose that the divisor $\big(\sum_i b_i P_i\big)$ has trivial image in T. Then it is possible to write $\big(\sum_i b_i P_i\big) = f^*\big(\sum_j R_j\big) + \mathrm{div}_D(h)$ for some $R_j \in C(k)$ and $h \in k(D)^*$. Thus, we have $\sigma(\mathrm{div}_D(h)) = \mathrm{div}_D(h)$, and we conclude that $h^\sigma = \xi^j h$ for some $j \in \{0, \ldots, q^s - 1\}$. Therefore, there exists $b \in k(C)^*$ such that $h = b\alpha^j$. Hence, we have an equality of divisors of the form $\big(\sum_i b_i P_i\big) = f^*(E) + j\big[\big(\sum_i a_i P_i\big) - \big(\sum_i a_i\big) P_d\big]$ for some $E \in \mathrm{Div}^0(C)$. It follows that $E = \sum_i c_i Q_i$ for some c_i. Hence, the map φ is an isomorphism, proving Part (b). □

Corollary 2.6. *Assume that $p \neq 2$. Let D/k be a smooth projective connected hyperelliptic curve of genus g. Denote by τ the hyperelliptic involution. Let σ be an automorphism of order p. Then either σ has a single fixed point, fixed by τ, or it has exactly two fixed points, permuted by τ.*

Proof. The hyperelliptic involution commutes with σ, and hence, it permutes the fixed points $\{P_1, \ldots, P_d\}$. If $d \geq 2$ and two fixed points P_1 and P_2 of σ are fixed by τ, then the divisor class $P_1 - P_2$ is fixed by τ. Proposition 2.5 shows that the class of $P_1 - P_2$ is not trivial and, since $p > 2$, this divisor class is not equal to the class of $-(P_1 - P_2)$. This is a contradiction since τ acts as the $[-1]$-map on $\mathrm{Jac}(D)$. Thus, τ fixes at most one point P_i.

If $d \geq 3$, then we may assume that either $\tau(P_1) = P_2$ and P_3 is fixed or that $\tau(P_1) = P_2$ and $\tau(P_3) = P_4$. In the first case, we find that $\tau(P_1 - P_3) = (P_2 - P_3) = -(P_1 - P_2) + (P_1 - P_3)$. Using the fact that τ acts as the $[-1]$-map on $\mathrm{Jac}(D)$, we find the relation $-(P_1 - P_3) = -(P_1 - P_2) + (P_1 - P_3)$ in $\mathrm{Jac}(D)$. Looking at this relation in T contradicts Proposition 2.5. The other case is similar and is left to the reader. □

Example 2.7. Assume that $p \neq 2$. Consider a smooth hyperelliptic curve C/k given by an affine equation $y^2 = f(x)$, and let D be its Galois cover given by the equation $z^p - z = x$. The automorphism $\sigma : D \to D$ with $\sigma(z) = z + 1$ has one fixed point P with $\delta(P) = 3(p-1)$ when $\deg(f)$ is odd, and it has two fixed points P_1 and P_2 with $\delta(P_1) = \delta(P_2) = 2(p-1)$ when $\deg(f)$ is even.

3. Arithmetical trees

Our main result in this section is Proposition 3.5, which will be needed in the proof of Theorem 5.3. This proposition pertains to arithmetical graphs, and we now recall how one associates such an object to any regular model of a curve.

Let X/K be any smooth, proper, geometrically connected curve of genus g. Let $\mathcal{X}/\mathcal{O}_K$ be a regular model of X/K. Let $\mathcal{X}_k := \sum_{i=1}^{v} r_i C_i$ denote the special fiber of $\mathcal{X}$, where C_i is an irreducible component and r_i is its multiplicity. Let $M := ((C_i \cdot C_j))_{1 \le i, j \le v}$ be the associated *intersection matrix*. Denote by G the associated graph. Let ${}^tR := (r_1, \dots, r_v)$ so that $MR = 0$. We call the triple (G, M, R) an *arithmetical graph* (in [Lorenzini 1989], the additional condition that $\gcd(r_1, \dots, r_v) = 1$ is assumed, and it is $(G, -M, R)$ that is called an arithmetical graph). For the purpose of simplifying the statements of some definitions, we sometimes think of G as a metric space with the natural topology where each edge of G with its two endpoints is homeomorphic to the closed unit interval [0, 1].

Let (G, M, R) be any arithmetical graph on v vertices. Let $M : \mathbb{Z}^v \to \mathbb{Z}^v$ and ${}^tR : \mathbb{Z}^v \to \mathbb{Z}$ be the linear maps associated to the matrices M and R. The *group of components* of (G, M, R) is defined as

$$\Phi_M := \operatorname{Ker}({}^tR)/\operatorname{Im}(M) = (\mathbb{Z}^v/\operatorname{Im}(M))_{\text{tors}}.$$

Motivated by the case of degenerations of curves, we shall denote by $(C, r(C))$ a vertex of G, where $r(C)$ is the coefficient of R corresponding to C. The integer $r(C)$, also denoted simply by r, is called the multiplicity of C. The matrix M is written as $M := ((C_i \cdot C_j))_{1 \le i, j \le v}$, and we write $|C_i \cdot C_i| := |(C_i \cdot C_i)|$.

3.1. Denote by $\langle \cdot , \cdot \rangle : \Phi_M \times \Phi_M \to \mathbb{Q}/\mathbb{Z}$ the perfect pairing $\langle \cdot , \cdot \rangle_M$ attached in [Bosch and Lorenzini 2002, Lemma 1.1] to the symmetric matrix M. Explicit values of this pairing are computed as follows. Let (C, r) and (C', r') be two distinct vertices of G. Define

$$E(C, C') := {}^t\Big(0, \dots, 0, \frac{r'}{\gcd(r, r')}, 0, \dots, 0, \frac{-r}{\gcd(r, r')}, 0, \dots, 0\Big) \in \mathbb{Z}^v,$$

where the first nonzero coefficient of $E(C, C')$ is in the column corresponding to the vertex C and, similarly, the second nonzero coefficient is in the column corresponding to the vertex C'. We say that the pair (C, C') is *uniquely connected* if there exists a path $\mathcal{P}$ in G between C and C' such that, for each edge e on $\mathcal{P}$, the graph $G \setminus \{e\}$ is disconnected. Note that, when a pair (C, C') is uniquely connected, then the path $\mathcal{P}$ is the unique shortest path between C and C'. A graph is a tree if and only if every pair of vertices of G is uniquely connected.

Let (C, r) and (C', r') be a uniquely connected pair with associated path $\mathcal{P}$. While walking on $\mathcal{P} \setminus \{C, C'\}$ from C to C', label each encountered vertex consecutively by $(C_1, r_1), (C_2, r_2), \dots, (C_n, r_n)$. Let G_i denote the connected component of C_i in $G\setminus\{\text{edges of } \mathcal{P}\}$. The graph G_i is reduced to a single vertex if and only if C_i is not a node of G. For convenience, we write $(C, r) = (C_0, r_0)$ and $(C', r') = (C_{n+1}, r_{n+1})$ and define G_0 and G_{n+1} accordingly.

3.2. The following facts are proved in [Bosch and Lorenzini 2002, Proposition 5.1]. Let (G, M, R) be any arithmetical graph. Let C and C' be two vertices such that (C, C') is a *uniquely connected* pair of G. Let γ denote the image of $E(C, C')$ in Φ_M. For (D, s) and (D', s') any two distinct vertices on G, let δ denote the image of $E(D, D')$ in Φ_M. Writing $\mathcal{P}$ for the oriented shortest path from C to C' as above, let C_α denote the vertex of $\mathcal{P}$ closest to D in G, and let C_β denote the vertex of $\mathcal{P}$ closest to D'. In other words, $D \in G_\alpha$ and $D' \in G_\beta$. Assume that $\alpha \le \beta$. (Note that we may have $\alpha = \beta$, and we may have $D = C_\alpha$ or $D' = C_\beta$.) Then if $\alpha < \beta$,

$$\langle \gamma, \delta\rangle = -\operatorname{lcm}(r, r')\operatorname{lcm}(s, s')\Big(\frac{1}{r_\alpha r_{\alpha+1}} + \frac{1}{r_{\alpha+1} r_{\alpha+2}} + \cdots + \frac{1}{r_{\beta-1} r_\beta}\Big) \mod \mathbb{Z}, \tag{3.2.1}$$

and if $C_\alpha = C_\beta$, then $\langle \gamma, \delta\rangle = 0$. Moreover,

$$\langle \gamma, \gamma\rangle = -\operatorname{lcm}(r, r')^2\Big(\frac{1}{r r_1} + \frac{1}{r_1 r_2} + \cdots + \frac{1}{r_n r'}\Big) \mod \mathbb{Z}. \tag{3.2.2}$$

Note that the negative signs in the expressions (3.2.1) and (3.2.2) are missing in [Bosch and Lorenzini 2002, Proposition 5.1]. Thus, all expressions for $\langle \gamma, \delta\rangle$ computed in Section 5 of [Bosch and Lorenzini 2002] using Proposition 5.1 are correct only after having been multiplied by -1. Similar sign mistakes occurred in [Lorenzini 2000]. The proof of [Bosch and Lorenzini 2002, Proposition 5.1] is correct except that its last line produces the opposite of the stated values for $\langle \gamma, \delta\rangle$ since we assume $\alpha \le \beta$.

3.3. Let (C, r) be a vertex of G of degree $d \ge 2$. Let (D_i, r_i), $i = 1, \dots, d$, denote the neighbors of C, that is, the vertices of G linked to C. Let τ_i denote the image of $E(D_i, D_d)$ in Φ_M for $i \in \{1, \dots, d-1\}$. We will use repeatedly the following expressions computed using (3.2.1) and (3.2.2):

$$\langle \tau_i, \tau_i\rangle = -\operatorname{lcm}(r_i, r_d)^2\frac{r_i + r_d}{r_i r_d r} \mod \mathbb{Z}$$

and, when $i \ne j$,

$$\langle \tau_i, \tau_j\rangle = -\operatorname{lcm}(r_i, r_d)\operatorname{lcm}(r_j, r_d)\frac{1}{r_d r} \mod \mathbb{Z}.$$

These formulas allow us to easily show that τ_i may not always be trivial. For example, let p be a prime dividing r. When $p \nmid r_i r_d(r_i + r_d)$, we find that $\langle \tau_i, \tau_i\rangle \ne 0$

and, thus, $\tau_i \neq 0$. Similarly, when for three distinct indices i, j and d we have $p \nmid r_i r_j r_d$, we find that $\langle \tau_i, \tau_j \rangle \neq 0$, showing that both τ_i and τ_j are not trivial.

We claim that r *kills* τ_i. Indeed, we find, using [Lorenzini 2000, Lemma 2.2], that the images in Φ_M of $E(D_i, C)$ and $E(C, D_d)$ have order dividing $\gcd(r_i, r)$ and $\gcd(r, r_d)$, respectively. Consider the following easy relation between vectors in $\mathbb{Z}^v$ [Lorenzini 2000, Remark 3.5]: given any three vertices (A, a), (B, b) and (C, c),

$$bE(A, C) = \frac{c}{\gcd(a, c)} \gcd(a, b) E(A, B) + \frac{a}{\gcd(a, c)} \gcd(b, c) E(B, C). \quad (3.3.1)$$

Using this relation, we find that $r\tau_i = 0$.

Lemma 3.4. *Let (G, M, R) be an arithmetical graph. Consider any two distinct vertices (A, a) and (A', a'), and let $\alpha_{A,A'}$ denote the image of $E(A, A')$ in Φ_M. Then the set $\{\alpha_{AA'} \mid A \neq A'\}$ is a set of generators for Φ_M.*

Proof. Let us note first that the statement is proved for (G, M, R) as soon as it is proved for $(G, M, R/\gcd(r_1, \dots, r_v))$. We will thus assume now that $\gcd(r_1, \dots, r_v) = 1$. Fix a vertex A, and consider the subgroup $(\Phi_M)_A$ of Φ_M generated by $\{\alpha_{AA'} \mid \text{all } A' \neq A\}$. We claim that $a\Phi_M \subseteq (\Phi_M)_A$. Indeed, an element $\phi \in \Phi_M$ is represented by the class of a vector $(f_D \mid D \in G)$ such that $\sum f_D r(D) = 0$. It follows that $a\phi = -\sum \gcd(a, r(D)) f_D \alpha_{AD}$. Since $\gcd(r_1, \dots, r_v) = 1$, ϕ can be expressed in terms of elements of the form $\alpha_{AA'}$. □

The following is a key relation between the τ_i's:

Proposition 3.5. *Let (G, M, R) be an arithmetical tree. Let (C, r) be a vertex of degree $d \geq 2$. Keep the notation introduced in 3.3. Then $\sum_{i=1}^{d-1} \gcd(r_i, r_d)\tau_i = 0$.*

Proof. Consider any two distinct vertices (A, a) and (A', a'), and let α denote the image of $E(A, A')$ in Φ_M. The previous lemma shows that the group Φ_M is generated by such elements α.

Let $\tau := \sum_{i=1}^{d-1} \gcd(r_i, r_d)\tau_i$. We claim that $\langle \tau, \alpha \rangle = 0$ for all such elements α. This claim, proved below, implies immediately that $\tau = 0$. Indeed, recall that, $\langle \cdot, \cdot \rangle$ being perfect, the element τ is trivial if and only if $\langle \tau, \phi \rangle = 0$ for all $\phi \in \Phi_M$.

Let us now prove our claim. Assume first that the path $\mathcal{Q}$ between A and A' contains the vertices D_i and D_d with $i \neq d$. We use (3.2.1) to compute modulo $\mathbb{Z}$ that

$$\begin{aligned}\langle \tau, \alpha \rangle &= \pm \operatorname{lcm}(a, a') \\ &\times \left(\gcd(r_i, r_d) \operatorname{lcm}(r_i, r_d) \left(\frac{1}{r_i r} + \frac{1}{r r_d} \right) + \sum_{j \neq i, d} \gcd(r_j, r_d) \operatorname{lcm}(r_j, r_d) \left(\frac{1}{r r_d} \right) \right),\end{aligned}$$

which simplifies to

$$\langle \tau, \alpha \rangle = \pm \operatorname{lcm}(a, a') \left(\sum_{j=1}^{d} r_j \right) \frac{1}{r}.$$

Since $\sum_{j=1}^{d} r_j = |C \cdot C| r$, we find that $\langle \tau, \alpha \rangle = 0$. When $\mathcal{Q}$ contains D_i and D_j with $i, j \neq d$ and $i \neq j$, we find that modulo $\mathbb{Z}$

$$\begin{aligned}\langle \tau, \alpha \rangle &= \pm \operatorname{lcm}(a, a')\Big(\gcd(r_i, r_d) \operatorname{lcm}(r_i, r_d) \frac{1}{r_i r} - \gcd(r_j, r_d) \operatorname{lcm}(r_j, r_d) \frac{1}{r_j r} \Big) \\ &= \pm \operatorname{lcm}(a, a') \Big(\frac{r_d}{r} - \frac{r_d}{r} \Big) = 0.\end{aligned}$$

It is clear that if the path $\mathcal{Q}$ contains no vertices D_i, or if it contains exactly one vertex D_i and does not contain the vertex C, then $\langle \tau, \alpha \rangle = 0$. It remains to consider the case where the path $\mathcal{Q}$ contains exactly one vertex D_i and the vertex C. Then C is an endpoint of $\mathcal{Q}$, and thus, r divides $\operatorname{lcm}(a, a')$. When $i \neq d$, we find that

$$\langle \tau, \alpha \rangle = \pm \operatorname{lcm}(a, a') \operatorname{lcm}(r_i, r_d) \gcd(r_i, r_d) \frac{1}{r_i r}$$

is 0 modulo $\mathbb{Z}$, and when $i = d$, we find that

$$\langle \tau, \alpha \rangle = \pm \operatorname{lcm}(a, a') \Big(\sum_{i=1}^{d-1} \operatorname{lcm}(r_i, r_d) \gcd(r_i, r_d) \frac{1}{r_i r} \Big)$$

is also 0 modulo $\mathbb{Z}$. □

4. Some combinatorics

Let (G, M, R) be an arithmetical graph. We introduce below a measure $\gamma_D g_D$ of how "complicated" certain subgraphs G_D of G are, and we describe G_D in Proposition 4.3 when $\gamma_D g_D$ is as small as possible. This result is needed in the proof of Theorem 6.8. A geometric motivation for the introduction of the quantity $\gamma_D g_D$ is found in the genus formula (6.1.1).

4.1. Let (G, M, R) be an arithmetical graph. Fix a vertex $(C_0, r(C_0))$ of G. Assume that C_0 is linked to a vertex $(D, r(D))$ by a single edge e and that, when the edge e is removed from G, then D and C_0 are not in the same connected component of the resulting graph. Let G_D *denote the connected component of* $G \setminus \{e\}$ *that contains* D. Consider the minor N_D of M corresponding to the vertices in G_D. Let

$$\gamma_D := \gcd(r(A) \mid A \text{ a vertex of } G_D).$$

Then γ_D *divides* $r(C_0)$. Indeed, γ_D divides the multiplicity of D and of all vertices linked to D except possibly that of C_0. But the relation $MR = 0$ implies then that γ_D divides the multiplicity of C_0. Let R_D denote the vector R restricted to the vertices of G_D. By definition, we find that R_D/γ_D is an integer vector.

Let $\beta(G)$ denote the first Betti number of the graph G. Letting $d_G(A)$ denote the degree of a vertex A in the graph G, we have

$$2\beta(G) - 2 = \sum_{\text{vertices } A \text{ of } G} (d_G(A) - 2).$$

Associated with any arithmetical graph (G, M, R) is the following integer invariant $g_0(G) \geq \beta(G)$ [Lorenzini 1989, Theorem 4.10], defined by the formula

$$2g_0(G) = 2\beta(G) + \sum_{\text{vertices } A \text{ of } G} (r(A) - 1)(d_G(A) - 2). \tag{4.1.1}$$

Let C_0 and D be as above. We now associate to the pair (N_D, R_D) an integer g_D, defined so that the formula below holds:

$$\gamma_D g_D = r(C_0) + r(D) + \sum_{\text{vertices } A \text{ of } G_D} r(A)(d_{G_D}(A) - 2).$$

Since γ_D divides $r(C_0)$, the invariant g_D is indeed an integer. We can rewrite this formula as

$$\begin{aligned} \gamma_D g_D = 2\beta(G_D) &+ (r(C_0) - 1) + (r(D) - 1) \\ &+ \sum_{\text{vertices } A \text{ of } G_D} (r(A) - 1)(d_{G_D}(A) - 2), \end{aligned} \tag{4.1.2}$$

and we find that

$$\begin{aligned} g_D = 2\beta(G_D) &+ \left(\frac{r(C_0)}{\gamma_D} - 1\right) + \left(\frac{r(D)}{\gamma_D} - 1\right) \\ &+ \sum_{\text{vertices } A \text{ of } G_D} \left(\frac{r(A)}{\gamma_D} - 1\right)(d_{G_D}(A) - 2). \end{aligned} \tag{4.1.3}$$

4.2. We will make use below of the following facts. Suppose that, on G, the vertices $D_0, D_1, \ldots, D_n$ are consecutive vertices on a terminal chain and D_n is the terminal vertex on this chain (in other words, D_i is linked by one edge to D_{i+1} for $i = 0, \ldots, n-1$, $d_G(D_i) = 2$ for $i = 1, \ldots, n-1$ and $d_G(D_n) = 1$). *Then* $\gcd(r(D_0), r(D_1)) = r(D_n)$*, and if* $|D_i \cdot D_i| > 1$ *for all* $i = 1, \ldots, n$*, then*

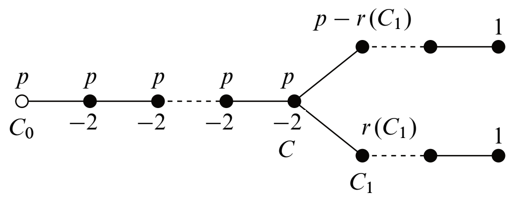
$$r(D_0) > r(D_1) > \cdots > r(D_n).$$

Indeed, the equality $|D_n \cdot D_n| r(D_n) = r(D_{n-1})$ obtained from the relation $MR = 0$ shows that $r(D_n)$ divides $r(D_{n-1})$ and $r(D_n) < r(D_{n-1})$ if $|D_n \cdot D_n| > 1$. Suppose that, for some i, we have $r(D_i) > r(D_{i+1})$. Then $r(D_{i-1}) > r(D_i)$ because $|D_i \cdot D_i| r(D_i) = r(D_{i-1}) + r(D_{i+1})$ and $|D_i \cdot D_i| \geq 2$. The equality $|D_i \cdot D_i| r(D_i) = r(D_{i-1}) + r(D_{i+1})$ implies that $\gcd(r(D_{i-1}), r(D_i)) = \gcd(r(D_i), r(D_{i+1}))$.

Proposition 4.3. *Let (G, M, R) be an arithmetical tree containing a vertex C_0 of prime multiplicity p. Assume that a vertex D linked to C_0 by an edge e has multiplicity divisible by p. Let G_D denote the connected component of $G \setminus \{e\}$ that contains D. Assume in addition that G_D does not contain any vertex A of degree* 1 *or* 2 *in G with $|A \cdot A| = 1$. Then*

$$\gamma_D g_D \geq 2(p-1).$$

If $\gamma_D g_D = 2(p-1)$, then $\gamma_D = 1$ and G_D is a graph of the shape depicted below, containing one node C of G only, of multiplicity p and degree 3 *in G. The two terminal vertices of G that belong to G_D have multiplicity* 1.



Let α denote the number of vertices of G_D on the chain linking C_0 to the node C of G_D (including the node C). Let C_1 and C_1' denote the vertices linked to C on the two terminal chains. Then $1 \leq r(C_1) < p$, and the minor of M corresponding to the vertices of G_D is completely determined by p, α and $r(C_1)$.

The proof of Proposition 4.3 is given in 4.6. We start with a preliminary lemma.

4.4. Let (G, M, R) be an arithmetical tree. For each node $(C, r(C))$ of degree $d(C) \geq 3$ in G, we define an invariant $\mu(C)$ as follows. Let $\rho(C)$ denote the number of terminal chains attached to C, and let $D_1(C), \ldots, D_{\rho(C)}(C)$ be the vertices of G linked to C that belong each to one terminal chain attached to C. Let $r_i(C)$ denote the multiplicity of $D_i(C)$. The multiplicity of the terminal vertex on the chain containing $D_i(C)$ is $\gcd(r(C), r_i(C))$. If no vertex A on the terminal chain has $|A \cdot A| = 1$, then $r_i(C) < r(C)$ (see 4.2). When a chain attached to C is not terminal, we will call it a *connecting chain*. As in [Lorenzini 1989, Theorem 4.7], we let, when $\rho(C) > 0$,

$$\mu(C) := (d(C) - 2)(r(C) - 1) - \sum_{j=1}^{\rho(C)} (\gcd(r(C), r_j(C)) - 1).$$

When $\rho(C) = 0$, we let $\mu(C) := (d(C) - 2)(r(C) - 1)$. It is clear that, if $r(C) = 1$, then $\mu(C) = 0$.

Lemma 4.5. *Assume that the terminal chains attached to C do not contain a vertex A with $|A \cdot A| = 1$. Then $\mu(C) \geq 0$, and $\mu(C) = 0$ if and only if $r(C) = 1$ and $\rho(C) = 0$.*

Proof. It is clear that, if a node C has $\rho(C) = 0$, then $\mu(C) \geq 0$, and $\mu(C) = 0$ only when $r(C) = 1$. Assume now that $\rho(C) > 0$. Our hypothesis implies that $r(C) > \gcd(r(C), r_i(C))$ for each vertex $D_i(C)$, $i = 1, \dots, \rho(C)$. In particular, $r(C) > 1$, and we need to prove that $\mu(C) > 0$. Let

$$s := \gcd(r(C), r_1(C), \dots, r_d(C)).$$

Assume first that $\rho(C) = d(C)$ so that G has a single node. It is proven in [Lorenzini 1989, Proposition 4.1] that, if $\rho(C) = d(C)$ and $s = 1$, then $\mu(C) \geq 0$. When $s > 1$, define

$$\mu_s(C) := (d(C) - 2)\Big(\frac{r(C)}{s} - 1\Big) - \sum_{j=1}^{\rho(C)} \Big(\frac{\gcd(r(C), r_j(C))}{s} - 1\Big).$$

The integer $\mu_s(C)$ is nothing but the μ-invariant of the node on the arithmetical graph obtained from G by dividing all its multiplicities by s. Thus, $\mu_s(C)$ is even [Lorenzini 1989, Definition 3.6] and $\mu_s(C) \geq 0$. Since

$$\mu(C) = -2(s - 1) + s\mu_s(C),$$

we find that $\mu(C) > 0$ if $\mu_s(C) > 0$. *We claim that, under our hypotheses,* $\mu(C) > 0$ *when* $s = 1$. Indeed, our hypotheses imply that $r(C) > \gcd(r(C), r_i(C))$ for each vertex $D_i(C)$, $i = 1, \dots, \rho(C)$. Dropping the reference to C, we can write

$$\begin{aligned} \mu(C) &:= (d - 2)(r - 1) - \sum_{j=1}^{d} (\gcd(r, r_i) - 1) \\ &\geq (d - 2)(r - 1) - d(r/2 - 1) = (d - 4)r/2 + 2. \end{aligned}$$

Thus, $\mu(C) > 0$ if $d \geq 4$. Assume now that $d = 3$. Then $cr = r_1 + r_2 + r_3$ for some c. Let $h_i = \gcd(r, r_i)$, and assume that $h_1 \geq h_2 \geq h_3$. Then $(h_1, h_2, h_3) = (r/2, r/2, r/2)$, $(r/2, r/2, r/3)$, $(r/2, r/3, r/3)$, $(r/2, r/3, r/4)$ cannot occur due to the divisibility $r \mid (r_1 + r_2 + r_3)$. Since the cases $(h_1, h_2, h_3) = (r/3, r/3, r/3)$, $(r/2, r/4, r/4)$, $(r/2, r/3, r/6)$ have $\mu(C) > 0$, we need only consider $(h_1, h_2, h_3) = (r/2, r/3, r/5)$. In this case, $r_1 = r/2$, $r_2 = r/3$ or $2r/3$ and $r_3 = ar/5$ with $a = 1, \dots, 4$. The reader will check that $cr = r_1 + r_2 + r_3$ is impossible to achieve with these values, and our claim is proved.

Let us assume now that $0 < \rho(C) < d(C)$. Then

$$\begin{aligned} \mu(C) &:= (d - 2)(r - 1) - \sum_{j=1}^{\rho} (\gcd(r, r_i) - 1) \\ &\geq (d - 2)(r - 1) - (d - 1)(r/2 - 1) = (d - 3)r/2 + 1 > 0. \end{aligned}$$ □

4.6. *Proof of Proposition 4.3.* We claim that G_D contains a node of G. (This node is also a node of G_D unless it is D itself and $d_G(D) = 3$.) Indeed, the hypotheses that $r(C_0) \leq r(D)$ and $|D \cdot D| > 1$ imply that $d_G(D) > 1$ because the relation $MR = 0$

provides otherwise for the equality $|D\cdot D|r(D)=r(C_0)$, which is a contradiction. Suppose then that D is connected in G_D to D_1. If $d_G(D)=2$, then we find from the relation $|D\cdot D|r(D)=r(C_0)+r(D_1)$ that $r(D)\le r(D_1)$. Repeating this discussion with D and D_1 instead of C_0 and D, we find that the graph G_D has a chain of increasing multiplicities $r(D)\le r(D_1)\le\cdots$, which eventually leads to a node of G_D (and of G).

In G, C_0 and D are adjacent vertices. Consider the connected component $\mathcal{G}$ of $G\setminus\{D\}$ that contains C_0. Two cases can occur: either (a) $\mathcal{G}$ contains a node of G, or (b) $\mathcal{G}$ does not contain a node of G, in which case we will call $\mathcal{G}$ a terminal chain of G. In the latter case, the terminal vertex on this chain has multiplicity $\gcd(r(C_0),r(D))$ (see 4.2), which equals $r(C_0)$ by hypothesis. The definition of $\gamma_D g_D$ in (4.1.2), along with the fact that we assume that G is a tree, allow us to write

$$\gamma_D g_D=(r(C_0)-1)+\sum_{\text{vertices } A \text{ of } G_D}(r(A)-1)(d_G(A)-2).$$

In case (a), C_0 is not on a terminal chain of G so that, by definition of $\mu(C)$ in 4.4, we can write

$$\gamma_D g_D=(r(C_0)-1)+\sum_{\text{nodes } C \text{ of } G \text{ in } G_D}\mu(C) \tag{4.6.1}$$

(where $\mu(C)$ is computed viewing C as a node of G and not of G_D). In case (b) where C_0 is on a terminal chain of G whose terminal vertex has multiplicity $r(C_0)$, we have

$$\gamma_D g_D=2(r(C_0)-1)+\sum_{\text{nodes } C \text{ of } G \text{ in } G_D}\mu(C).$$

We prove below case (a). The arguments to prove (b) are similar and are left to the reader. Case (b) will not be used in the remainder of this article.

Assume that we are in case (a). We can apply Lemma 4.5, and we obtain that each term $\mu(C)$ in the above sum is nonnegative. In view of (4.6.1), since $r(C_0)=p$ by hypothesis, we need to show that $\sum_{\text{nodes } C}\mu(C)\ge p-1$, and we need to describe the graphs for which $\sum_{\text{nodes } C}\mu(C)=(p-1)$.

Denote by C the node of G closest to C_0 in G_D. (This node could be D.) The multiplicity of C is divisible by p since p divides the consecutive multiplicities $r(C_0)$ and $r(D)$ (similar argument as in 4.2). Let np denote the multiplicity of C.

Suppose that C (of degree d in G) has only one connecting chain. If $n=1$, then all terminal multiplicities at C equal 1 and $\mu(C)=(d-2)(p-1)$. The case $d=3$ leads to the case described in the statement of Proposition 4.3 with $\mu(C)=(p-1)$, $\gamma_D g_D=2(p-1)$ and $\gamma_D=1$. When $d>3$, we have $\mu(C)>p-1$, as desired.

When $n > 1$, the inequality

$$\begin{aligned}\mu(C) &\geq (d-2)(np-1) - (d-1)(np/2-1) \\ &= (d-2)np/2 - np/2 + 1\end{aligned}$$

shows that we have $\mu(C) > p-1$ unless $d = 3$. When $n > 1$ and $d = 3$, every vertex on the chain linking C to C_0 has multiplicity divisible by p. Thus, either (i) both terminal multiplicities of C are coprime to p (call them n_1 and n_2), or (ii) both are divisible by p (call them $m_1 p$ and $m_2 p$).

In case (i), $\mu(C) = np - n_1 - n_2 + 1$ with n_1 and n_2 dividing n. It follows that $\mu(C) \geq n(p-2)+1$. Clearly, $\mu(C) > p-1$ unless $p = 2$. Assume that $p = 2$. If $(n_1, n_2) \neq (n, n)$, we find that $\mu(C) = n(p-1)+1 > (p-1)$. The case $(n_1, n_2) = (n, n)$ cannot happen because, in that case, n divides the multiplicity of all the components linked to C, which implies then that $n = 2$. But a node of multiplicity 4 cannot have exactly three vertices of multiplicity 2 attached to it.

In case (ii), $\mu(C) = (n - m_1 - m_2)p + 1$ with m_1 and m_2 dividing n. The equality $(n - m_1 - m_2) = 0$ is not possible. Indeed, it is only possible if $m_1 = m_2 = n/2$. But since $\gcd(m_1, m_2) = 1$, this case can happen only if $n = 2$. But then $|C \cdot C|$ would equal $3/2$, a contradiction. It follows that $\mu(C) = (n - m_1 - m_2)p + 1 > p-1$.

Suppose now that C, of multiplicity np, has at least two connecting chains. If $n > 1$, then

$$\mu(C) \geq (d-2)(np-1) - (d-2)(np/2-1) = (d-2)np/2 > p-1,$$

as desired. If $n = 1$, then $\mu(C) = (d-2)(p-1)$. Thus, $\mu(C) > p-1$ if $d > 3$. Suppose now that $d = 3$. Since G_D is a tree with a node C of degree 3, G_D must have at least three terminal vertices. Thus, there must exist at least one additional node C' on the graph G_D that has a terminal chain. It follows that $\mu(C') \geq 1$ (Lemma 4.5) and, therefore, $\mu(C) + \mu(C') > p-1$, as desired.

4.7. To conclude the proof of Proposition 4.3, we now specify the intersection matrix in the case where $\gamma_D g_D = 2(p-1)$. Let (G, M, R) be as in Proposition 4.3, and assume that the vertex D is such that $\gamma_D g_D = 2(p-1)$. Let N_D denote the matrix M restricted to the vertices of G_D. Let α denote the number of vertices of G_D on the chain linking C_0 to the node C of G_D (including the node C). Each of these vertices except C is of degree 2. The multiplicity of C is p. Since we assume that no vertex of degree 2 has self-intersection -1, we find that the multiplicity of each of these vertices must be p. It follows that each of these vertices except possibly C must have self-intersection -2.

Let C_1 and C_1' denote the vertices linked to C on the two terminal chains. Since they have degree 1 or 2 and cannot have self-intersection -1, we find that $1 \leq r(C_1) < r(C) = p$ and $r(C_1') < r(C)$. Moreover, from $MR = 0$, we find that $p + r(C_1) + r(C'_1) = p|C \cdot C|$. It follows that $|C \cdot C| = 2$, and $r(C'_1) = p - r(C_1)$. We claim that N_D depends only on p, α and $r(C_1)$, and we write it as $N_D = N(p, \alpha, r(C_1))$. Indeed, the pair $(p, r(C_1))$ completely determines all multiplicities and all self-intersections on the terminal chain containing C_1: use $(r, s) = (p, r(C_1))$ in 4.8 below to determine the self-intersections and multiplicities of the terminal chain. Similarly, the pair $(p, r(C'_1))$ completely determines all multiplicities and all self-intersections on the terminal chain containing C'_1. This conclude the proof of Proposition 4.3. The matrix N_D is an intersection matrix also introduced in [Lorenzini 2013a, Example 3.18]. □

4.8. Recall the following standard construction. Given an ordered pair of positive integers $r > s$ with $\gcd(r, s) = 1$, we construct an associated intersection matrix $N = N(r, s)$ with vector $R = R(r, s)$ and $NR = -re_1$ as follows (where e_1 denotes the first standard basis vector of $\mathbb{Z}^n$). Using the division algorithm, we can find positive integers $b_1, \dots, b_m$ and $s_1 = s > s_2 > \cdots > s_m = 1$ such that $r = b_1 s - s_2$, $s_1 = b_2 s_2 - s_3$ and so on until we get $s_{m-1} = b_m s_m$. These equations are best written in matrix form:

$$\begin{pmatrix} -b_1 & 1 & \dots & 0 \\ 1 & -b_2 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & \dots & 1 & -b_m \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ \vdots \\ s_m \end{pmatrix} = \begin{pmatrix} -r \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

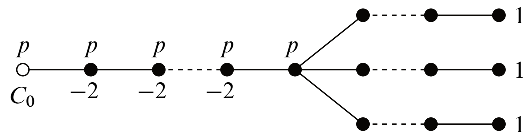
We let $N(r, s)$ denote the above square matrix and $R(r, s)$ be the column matrix on the left of the "equals" sign. It is well-known that $\det(N(r, s)) = \pm r$ (see [Lorenzini 2000, Lemma 2.6]). We recall also for use in Corollary 6.12 that

$$\frac{1}{rs} + \frac{1}{ss_2} + \cdots + \frac{1}{s_{m-1}s_m} = \frac{c}{r}, \tag{4.8.1}$$

where $0 < c < r$ is such that $r \mid cs - 1$ (see [Lorenzini 2000, Lemmas 2.8 and 2.6]).

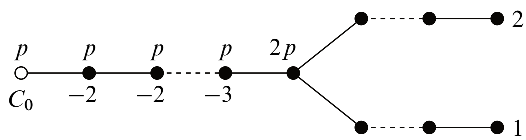
Remark 4.9. In Proposition 4.3, the hypothesis that $\gamma_D g_D = 2(p-1)$ allowed us to completely describe the graph G_D. For a fixed $\gamma_D g_D > 2(p-1)$, the situation is much more complicated and several possible types of graphs G_D may occur. It would follow from our guess in 6.2 that, for applications to models of curves, it suffices to classify the cases where $\gamma_D g_D$ is a multiple of $p-1$. We give below several possible types of graphs G_D with $\gamma_D g_D = 3(p-1)$ when p is odd.

(a) G_D is a graph with one node of G only, of multiplicity p and degree 4 in G. The three terminal vertices of G that belong to G_D have multiplicity 1.



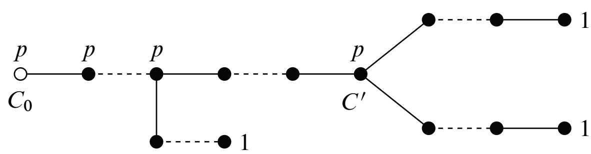
To completely determine the intersection matrix N_D and the vector R_D, one needs to also provide the multiplicities r_1, r_2 and r_3, of the first vertices on each of the three terminal chains, with the conditions $1 \leq r_1, r_2, r_3 < p$ and $r_1 + r_2 + r_3$ divisible by p. Such data can only be provided when p is odd. The self-intersection of the node is then $-(p + r_1 + r_2 + r_3)/p = -2$ or -3.

(b) G_D is a graph with one node of G only, of multiplicity $2p$ and degree 3 in G. The two terminal vertices of G that belong to G_D have multiplicity 1 and 2, respectively.

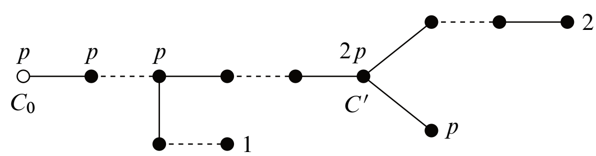


(c) G_D is a graph with 2 nodes C and C' of G. Let C be the node closest to C_0 in G_D. It has multiplicity p and degree 3 in G, and it has a single terminal chain with terminal multiplicity 1. The node C' is connected to C by a connecting chain that contains a vertex of multiplicity *coprime* to p.

(i)



(ii)



We conclude this section with some general remarks concerning the invariant g_D introduced in (4.1.2).

Remark 4.10. Let (G, M, R) be an arithmetical graph. As at the beginning of this section, fix a vertex $(C_0, r(C_0))$ of G. Assume that C_0 is linked to a vertex $(D, r(D))$ by a single edge e and that, when the edge e is removed from G, then D and C_0 are not in the same connected component of the resulting graph. Let G_D denote the connected component of $G \setminus \{e\}$ that contains D. Consider the minor $N = N_D$ of M corresponding to the vertices in G_D. Let n denote the number of vertices of G_D.

(a) *The integer g_D depends only on the matrix N_D and the vertex D on the graph G_D.* To prove this statement, we show that the vector R_D/γ_D is completely determined by N_D and the vertex D. Indeed, let us number the vertices of G_D such that D is the first vertex numbered. Then R_D/γ_D is a vector with positive coefficients such that $N_D(R_D/\gamma_D) = {}^t(-r(C_0)/\gamma_D, 0, \dots, 0)$ (where the superscript t indicates the transpose vector). The existence of such a relation insures that N_D is negative-definite (see [Lorenzini 2013a, §3.3]), and the vector R_D/γ_D is a rational multiple of the first column of the unique matrix N^* such that $NN^* = N^*N = \det(N)\,\mathrm{Id}_n$ [Lorenzini 2013a, Definition 3.4]. The integer $r(C_0)/\gamma_D$ is the order in $\mathbb{Z}^n/\operatorname{Im}(N)$ of the class of the first basis vector e_1 [Lorenzini 2013a, Lemma 3.5].

(b) *The integer g_D is nonnegative*. More precisely,

$$g_D - 2\beta(G_D) \geq \left(\frac{r(C_0)}{\gamma_D}\right) + \gcd\left(\frac{r(D)}{\gamma_D}, \frac{r(C_0)}{\gamma_D}\right) - 2 \geq 0. \tag{4.10.1}$$

To prove the first inequality, complete the pair $(N, R_D/\gamma_D)$ into an arithmetical graph (G', M', R') by adding a chain attached to D, as in [Lorenzini 2013a, §3.15]. Clearly, $\beta(G') = \beta(G_D)$. The graphs G' and G_D differ in only two vertices of degree not equal to 2: the terminal vertex on the new terminal chain on G' has terminal multiplicity $\gcd(r(D)/\gamma_D, r(C_0)/\gamma_D)$, and $d_{G'}(D) = d_{G_D}(D) + 1$. Using (4.1.1) and (4.1.3), it is easy to show that

$$\begin{aligned} &2g_0(G', M', R') - 2\beta(G') \\ &\quad = g_D - 2\beta(G_D) - \left(\frac{r(C_0)}{\gamma_D} - 1\right) - \left(\gcd\left(\frac{r(D)}{\gamma_D}, \frac{r(C_0)}{\gamma_D}\right) - 1\right). \end{aligned} \tag{4.10.2}$$

The integer $g_0(G') - \beta(G')$ is always nonnegative [Lorenzini 1989, Theorem 4.10], and the statement follows.

(c) In analogy with the arithmetic genus of curves on surfaces, we define, given $Z \in \mathbb{Z}^n$, a (possibly negative) integer $p_a(Z)$ as follows. If $Z = C_i$ is a vertex of G_D, we let $p_a(Z) = 0$. We let $p_a(rC_i)$ be defined by the formula $2p_a(rC_i) - 2 = r^2C_i^2 + r(|C_i^2| - 2)$ (where we have abbreviated $C_i \cdot C_i$ by C_i^2). Since $r^2 - r$ is always even, $p_a(rC_i)$ is an integer. In general, when $Z = \sum_{i=1}^n r_iC_i$, we let

$$Z^2 := \sum_{1 \leq i, j \leq n} r_ir_j(C_i \cdot C_j)$$

and set

$$2p_a(Z) - 2 := Z^2 + \sum_{i=1}^n r_i(|C_i^2| - 2).$$

We leave it to the reader to check that

$$g_D = 2p_a(R_D/\gamma_D) - 2 + \frac{r(D)}{\gamma_D}\left(\frac{r(C_0)}{\gamma_D} + 1\right).$$

(d) *The integer g_D is even when either $r(C_0)$ is odd or $r(D)$ is even.* This can be seen from the formula for g_D in (c) or from (4.10.2).

(e) Assume that G_D is a tree. Then the order $|\det(N)|$ of the group $\Phi_N := \mathbb{Z}^n/N(\mathbb{Z}^n)$ can be computed completely in terms of the vector R_D/γ_D and of the graph G_D (see [Lorenzini 2013a, Theorem 3.14]), and we find that

$$|\det(N)| = \frac{r(D)}{\gamma_D}\frac{r(C_0)}{\gamma_D} \prod_{\text{vertices } A \text{ of } G_D} \left(\frac{r(A)}{\gamma_D}\right)^{d_{G_D}(A)-2},$$

where $d_{G_D}(A)$ is the degree of the vertex A in the graph G_D. Recall now the formula (4.1.3):

$$g_D = \left(\frac{r(D)}{\gamma_D} - 1\right) + \left(\frac{r(C_0)}{\gamma_D} - 1\right) + \sum_{\text{vertices } A \text{ of } G_D} \left(\frac{r(A)}{\gamma_D} - 1\right)(d_{G_D}(A) - 2).$$

This last expression is surprisingly similar to the expression for $|\det(N)|$. This motivates the following result.

Let $x > 0$ be any integer, and define the function $\ell(x) := \sum_{q \text{ prime}} \operatorname{ord}_q(x)(q-1)$. *Then*

$$\ell(|\det(N)|) \le g_D. \tag{4.10.3}$$

This result is not used in the remainder of this paper, and we will provide here only a sketch of proof.

Sketch of proof. We complete the pair $(N, R_D/\gamma_D)$ into an arithmetical graph (G', M', R') by adding a chain attached to D, as in [Lorenzini 2013a, §3.15]. The order of the component group $\Phi(M')$ is given in [Lorenzini 1989, Corollary 2.5], and the relation between $\det(N)$ and $|\Phi(M')|$ is discussed in the proof of Theorem 3.14 in [Lorenzini 2013a]. We can then bound $|\Phi(M')|$ in terms of $g_0(G', M', R')$ using [Lorenzini 1989, Corollary 4.8], which states that $\ell(|\Phi(M')|) \le 2g_0(G', M', R')$. The inequality $\ell(|\det(N)|) \le g_D$ follows then from (4.10.2). $\square$

5. The quotient construction

Let K be a complete discrete valuation field with valuation v, ring of integers $\mathcal{O}_K$, uniformizer π_K and residue field k of characteristic $p > 0$, assumed to be algebraically closed. Let X/K be a smooth proper geometrically connected curve of genus $g > 0$. When $g = 1$, assume in addition that $X(K) \neq \varnothing$. Assume that X/K does not have semistable reduction over $\mathcal{O}_K$ and that it achieves good reduction after a cyclic extension L/K of prime degree q.

Let H denote the Galois group of L/K. Let $\mathcal{Y}/\mathcal{O}_L$ be the smooth model of X_L/L. Let σ denote a generator of H. By minimality of the model $\mathcal{Y}$, σ defines an automorphism of $\mathcal{Y}$ also denoted by σ (but note that $\sigma : \mathcal{Y} \to \mathcal{Y}$ is not a morphism of $\mathcal{O}_L$-schemes). We also denote by σ the automorphism of $\mathcal{Y}_k$ induced by the action of σ on $\mathcal{Y}$. Let $\mathcal{Z}/\mathcal{O}_K$ denote the quotient $\mathcal{Y}/H$, and let $\alpha : \mathcal{Y} \to \mathcal{Z}$ denote the quotient map. The scheme $\mathcal{Z}$ is normal. The map α induces a natural map $\mathcal{Y}_k \to \mathcal{Z}_k^{\mathrm{red}}$ that factors as follows:

$$\mathcal{Y}_k \xrightarrow{\rho} \mathcal{Y}_k/\langle\sigma\rangle \to \mathcal{Z}_k^{\mathrm{red}}.$$

5.1. We claim that the first map is Galois of order $|H|$ and that the second map is the normalization map of $\mathcal{Z}_k^{\mathrm{red}}$. Indeed, let $\mathrm{Spec}(B)$ denote a dense open set of $\mathcal{Y}$ invariant under the action of H. Then $\mathrm{Spec}(B^H)$ is a dense open set of $\mathcal{Z}$. Let $A := B^H$. Let $P_B = (\pi_L)$ denote the prime ideal of B corresponding to $\mathcal{Y}_k$, and let $P_A := P_B \cap A$. We have the natural maps

$$B^H/P_A \hookrightarrow (B/P_B)^H \hookrightarrow B/P_B.$$

The extension of discrete valuation rings $(B^H)_{P_A} \to B_{P_B}$ induces an extension of residue fields $(B^H)_{P_A}/P_A(B^H)_{P_A} \to B_{P_B}/P_B B_{P_B}$. We claim that this extension has degree $|H|$. Indeed, our assumption is that the curve X/K does not have good reduction over $\mathcal{O}_K$. If the residue extension is trivial, the normalization of the curve $\mathcal{Z}_k^{\mathrm{red}}$ is isomorphic to $\mathcal{Y}_k$ and, thus, is of genus g. In addition, we find that $P_A B_{P_B} = (P_B B_{P_B})^{|H|}$ so that $\pi_K A_{P_A} = (P_A A_{P_A})$. The special fiber of $\mathcal{Z}$ is then reduced, and the curve X/K has good reduction over $\mathcal{O}_K$, a contradiction. It follows then that $P_A B_{PB} = P_B B_{PB}$ so that $\pi_K A_{P_A} = (P_A A_{P_A})^{|H|}$. Hence, the multiplicity in $\mathcal{Z}$ of the irreducible component $\mathcal{Z}_k^{\mathrm{red}}$ equals $|H|$.

It is easy to check that, for any $x \in (B/P_B)^H$, $|H|x$ and $x^{|H|}$ belong to A/P_A. Thus, when $|H| \neq p$, A/P_A and $(B/P_B)^H$ have the same field of fractions. When $|H| = p$, it could happen that A/P_A and $(B/P_B)^H$ do not have the same field of fractions, and then the extension of fields of fractions is purely inseparable of degree p with $(B/P_B)^H = B/P_B$. It follows that the special fiber of $\mathcal{Z}$ also has genus g. When $g > 1$, this is not possible since the multiplicity of $\mathcal{Z}_k$ is p. When $g = 1$, it could happen that $\mathcal{Z}$ is the minimal model of X/K with a multiple special fiber. This case cannot happen in our situation because we assumed $X(K) \neq \varnothing$: a K-rational point always reduces to a smooth point in the special fiber. Thus, the automorphism $\sigma : \mathcal{Y}_k \to \mathcal{Y}_k$ is not trivial. We find that A/P_A and $(B/P_B)^H$ have the same fields of fractions so that the Dedekind domain $(B/P_B)^H$ is the integral closure of A/P_A.

5.2. Let $P_1, \dots, P_d$ be the ramification points of the map $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$. Let $Q_1, \dots, Q_d$ be their images in $\mathcal{Z}$. The normal scheme $\mathcal{Z}$ is singular exactly at $Q_1, \dots, Q_d$. Indeed, the morphism $\mathcal{Y} \to \mathcal{Z}$ is unramified outside these points. If the point Q_i were regular, the morphism would be flat above Q_i [Altman and Kleiman 1970, Corollary V.3.6] and the branch locus would then be pure of codimension 1 [Altman and Kleiman 1970, Theorem VI.6.8], a contradiction.

Consider the regular model $\mathcal{X} \to \mathcal{Z}$ obtained from $\mathcal{Z}$ by a minimal desingularization. After finitely many blow-ups $\mathcal{X}' \to \mathcal{X}$, we can assume that the model $\mathcal{X}'$ is such that $\mathcal{X}'_k$ has smooth components and normal crossings and is minimal with this property. Let f denote the composition $\mathcal{X}' \to \mathcal{Z}$. Let C_0/k denote the strict transform in $\mathcal{X}'$ of the irreducible closed subscheme $\mathcal{Z}_k^{\text{red}}$ of $\mathcal{Z}$. The curve C_0 has multiplicity $|H|$ in $\mathcal{X}'$. Let $D_1, \ldots, D_d$ denote the irreducible components of $\mathcal{X}'_k$ that meet C_0. Let r_i denote the multiplicity of D_i, $i = 1, \ldots, d$. We assume $d \geq 1$. Our main theorem in this section is this:

Theorem 5.3. *Let X/K be a smooth proper geometrically connected curve of genus $g > 0$ with $X(K) \neq \varnothing$ if $g = 1$. Assume that X/K does not have semistable reduction over $\mathcal{O}_K$ and that it achieves good reduction after a cyclic extension L/K with Galois group H of prime degree p. Keep the above notation, and let Q_i be a singular point of the quotient $\mathcal{Z} := \mathcal{Y}/H$. Let G_{Q_i} denote the graph associated with the curve $f^{-1}(Q_i)$. Let G denote the graph associated with the special fiber $\mathcal{X}'_k$. Then, for all $i = 1, \ldots, d$, the graph G_{Q_i} contains a node of G and p divides r_i.*

Proof. When $d = 1$, the theorem is immediate: the component C_0 of multiplicity p is a terminal vertex of the graph of $\mathcal{X}'$, and thus, $p|C_0 \cdot C_0| = r_1$. Assume that G_{Q_1} does not contain a node of G. Then since $d = 1$, G does not contain a node. Since the resolution is minimal with normal crossings, none of the components of $\mathcal{X}'_k$ can have self-intersection -1 except possibly for C_0. It is clear that the graph G is not reduced to a single vertex since the model $\mathcal{Z}$ is singular. Thus, the graph G has a second terminal vertex C' in addition to C_0. But then, walking on G from C' towards C_0, we find that the multiplicities can only be strictly increasing. This is a contradiction since all multiplicities on G are divisible by p (because two consecutive ones are), and G must contain a node. We assume from now on that $d > 1$.

Let $A := \operatorname{Jac}(X/K)$. Let $\mathcal{A}_K/\mathcal{O}_K$ denote the Néron model of A/K. Let $\mathcal{A}_L/\mathcal{O}_L$ denote the Néron model of A_L/L, and denote by $\eta : \mathcal{A}_K \times_{\mathcal{O}_K} \mathcal{O}_L \to \mathcal{A}_L$ the canonical map induced by the functoriality property of Néron models. The special fiber $(\mathcal{A}_K)_k$ is an extension of a finite group $\Phi_{A/K}$, called the group of components, by the connected component of zero $(\mathcal{A}_K)_k^0$ of $(\mathcal{A}_K)_k$:

$$0 \to (\mathcal{A}_K)_k^0 \to (\mathcal{A}_K)_k \to \Phi_{A/K} \to 0.$$

Assume by contradiction that p is coprime to one of the r_i's. Without loss of generality, we may assume that $p \nmid r_d$. For each $i = 1, \ldots, d$, choose a point $x_i \in D_i$ such that x_i is a regular point of $(\mathcal{X}'_k)^{\text{red}}$. Since K is complete, we can find a closed point R_i of X of degree r_i over K and such that the closure of R_i in $\mathcal{X}'$ meets the special fiber $\mathcal{X}_k$ exactly in x_i (see, e.g., [Gabber et al. 2013, Proposition 8.4(3)]). For each $i = 1, \ldots, d-1$, consider the following divisor of degree 0 on X:

$$S_i := \frac{r_d}{\gcd(r_i, r_d)} R_i - \frac{r_i}{\gcd(r_i, r_d)} R_d.$$

We also denote by S_i its image in $\mathrm{Jac}(X)/K$. We recall below Raynaud's description of the Néron model of a Jacobian in order to be able to describe explicitly the image of S_i under both the reduction map $\mathrm{Jac}(X)(K) \to \Phi_{A/K}$ and the reduction map $\mathrm{Jac}(X)(L) \to (\mathcal{A}_L)_k(k)$. We will be able to contradict the hypothesis that $p \nmid r_d$ by considering the reductions of $\sum_{i=1}^{d-1} \gcd(r_i, r_d) S_i$.

Raynaud [1970] exhibited an explicit separated quotient $Q_K/\mathcal{O}_K$ of the open subfunctor of $\mathrm{Pic}_{\mathcal{X}'/\mathcal{O}_K}$ consisting of line bundles of total degree 0, and he showed that, when the residue field k is algebraically closed, $Q_K/\mathcal{O}_K$ is isomorphic to the Néron model of A/K [Bosch et al. 1990, Theorem 9.5.4(a)]. The canonical map $Q_K(K) \to \Phi_{Q_K}$ is described as follows [Bosch et al. 1990, Lemma 9.5.9, Theorem 9.6.1]. Represent an element of $Q_K(K)$ by a line bundle $\mathcal{L}$ on X of degree 0. Let $\overline{\mathcal{L}}$ denote an extension of $\mathcal{L}$ to $\mathcal{X}'$. Number the irreducible components of $\mathcal{X}'_k$ as $C_1, \dots, C_v$. Consider the map $\bigoplus_i \mathbb{Z}C_i \to \mathrm{Hom}\big(\bigoplus_i \mathbb{Z}C_i, \mathbb{Z}\big)$ that sends C_i to the map δ_{C_i} with $\delta_{C_i}(C_j) := (C_i \cdot C_j)$. The group Φ_M is isomorphic to the torsion subgroup of the cokernel of this map. The group of components Φ_{Q_K} is isomorphic to Φ_M, and under this isomorphism, the image of $\mathcal{L}$ under $Q_K(K) \to \Phi_{Q_K}$ is the map $\delta_{\mathcal{L}}$ with $\delta_{\mathcal{L}}(C_i) := (C_i \cdot \overline{\mathcal{L}})$. It follows immediately from these facts that the image in Φ_{Q_K} of $S_i \in \mathrm{Jac}(X)(K)$ can be identified with the image τ_i of the vector $E(D_i, D_d)$ in Φ_M (notation as in 3.1 and 3.3).

Consider now the reduction map $Q_L(L) \to (Q_L)_k(k)$. The closure of any point in the preimage under $X_L \to X$ of the closed point R_i meets the special fiber of the smooth model $\mathcal{Y}$ of X_L only at the point P_i. The line bundle $\mathcal{L}$ corresponding to the divisor S_i pulls back to a line bundle $\mathcal{L}_L$ on X_L. We find that the reduction of $\mathcal{L}_L \in \mathrm{Jac}(X_L)(L)$ is the point of $\mathrm{Jac}(\mathcal{Y}_k)(k)$ corresponding to the divisor $\mathrm{lcm}(r_i, r_d)(P_i - P_d)$.

We may now find a contradiction to the assertion that $p \nmid r_d$ *when the quotient of* $\mathcal{Y}_k$ *by the action of* H *has genus* 0. As we indicated above, the element $\sum_{i=1}^{d-1} \gcd(r_i, r_d) S_i$ in $\mathrm{Jac}(X)(K)$ reduces to the element $\sum_{i=1}^{d-1} \gcd(r_i, r_d)\tau_i$ in Φ_M. Proposition 3.5 shows that the latter element is zero in Φ_M. Thus, $\sum_{i=1}^{d-1} \gcd(r_i, r_d) S_i$ reduces in the connected component $(Q_K)_k^0$. Our additional hypothesis implies that this connected component is unipotent. This follows from [Bosch et al. 1990, Theorem 9.5.4] if the greatest common divisor of the multiplicities of the components of $\mathcal{X}'_k$ is 1 and from [Liu et al. 2004, Proposition 7.1] in general. It follows that the image of $(Q_K)_k^0$ under the canonical map $\eta : \mathcal{A}_K \times_{\mathcal{O}_K} \mathcal{O}_L \to \mathcal{A}_L$ is trivial.

Consider now the element $\sum_{i=1}^{d-1} \gcd(r_i, r_d) S_i$ in $\mathrm{Jac}(X_L)(L)$. Our discussion above shows that it reduces to the element $r_d\big(\sum_{i=1}^{d-1} r_i(P_i - P_d)\big)$ in $\mathrm{Jac}(\mathcal{Y}_k)(k)$. We have thus proved that $r_d\big(\sum_{i=1}^{d-1} r_i(P_i - P_d)\big) = 0$ in $\mathrm{Jac}(\mathcal{Y}_k)(k)$. Our hypothesis on the quotient of $\mathcal{Y}_k$ by H implies that each $P_i - P_d$ has order p (Proposition 2.5). Since $r_d\big(\sum_{i=1}^{d-1} r_i(P_i - P_d)\big) = 0$ and we assume that p does not divide r_d, we can conclude that $\sum_{i=1}^{d-1} r_i(P_i - P_d) = 0$. Then Proposition 2.5 implies that p divides r_i for all $i = 1, \dots, d-1$. Since $|C_0 \cdot C_0| p = r_1 + \cdots + r_d$, it follows that p divides r_d, which contradicts our assumption.

When the quotient of $\mathcal{Y}_k$ by the action of H has positive genus, the image of $(Q_K)_k^0$ under the canonical map $\eta : \mathcal{A}_K \times_{\mathcal{O}_K} \mathcal{O}_L \to \mathcal{A}_L$ is not trivial, and the following additional considerations must be discussed. Let $\mathrm{Norm}(\mathcal{X}')$ denote the normalization of $\mathcal{X}'$ in the field of fractions of $\mathcal{Y}$. Since $\mathcal{Y}$ is integral over $\mathcal{Z}$, we have a natural map $\mathrm{Norm}(\mathcal{X}') \to \mathcal{Y}$. All components of $\mathcal{X}'$ are rational except possibly the component C_0 [Lorenzini 2013a, Lemma 2.10].

By construction, we have a natural map $\mathrm{Norm}(\mathcal{X}') \to \mathcal{X}' \times_{\mathcal{O}_K} \mathcal{O}_L$. Let $\mathcal{N} \to \mathrm{Norm}(\mathcal{X}')$ denote a resolution of the singularities of $\mathrm{Norm}(\mathcal{X}')$. Consider the commutative diagram of $\mathcal{O}_L$-morphisms

$$\begin{array}{ccccc} \mathcal{N} & \longrightarrow & \mathrm{Norm}(\mathcal{X}') & \longrightarrow & \mathcal{Y} \\ & & \big\downarrow & & \big\downarrow \\ & & \mathcal{X}' \times_{\mathcal{O}_K} \mathcal{O}_L & \longrightarrow & \mathcal{Z} \times_{\mathcal{O}_K} \mathcal{O}_L \end{array}$$

The maps $\mathcal{N} \to \mathrm{Norm}(\mathcal{X}') \to \mathcal{X}' \times_{\mathcal{O}_K} \mathcal{O}_L$ induce maps of the associated Picard functors

$$\mathrm{Pic}_{\mathcal{X}'/\mathcal{O}_K} \times_{\mathcal{O}_K} \mathcal{O}_L \cong \mathrm{Pic}_{\mathcal{X}' \times_{\mathcal{O}_K} \mathcal{O}_L/\mathcal{O}_L} \to \mathrm{Pic}_{\mathrm{Norm}(\mathcal{X}')/\mathcal{O}_L} \to \mathrm{Pic}_{\mathcal{N}/\mathcal{O}_L},$$

whose composition induces the canonical map of Néron models

$$\eta : Q_K \times_{\mathcal{O}_K} \mathcal{O}_L \to Q_L.$$

Considering the special fibers over k, we obtain a commutative diagram

$$\begin{array}{ccc} \mathrm{Pic}^0_{\mathcal{N}_k/k} & \longrightarrow & (Q_L)_k^0 \\ \big\uparrow & & \big\uparrow \\ \mathrm{Pic}^0_{\mathcal{X}'_k/k} & \longrightarrow & (Q_K)_k^0 \end{array}$$

Since we do not have additional information on the special fiber $\mathcal{X}'_k$, we cannot conclude that the bottom horizontal map is an isomorphism. It is however faithfully flat [Raynaud 1970, Corollaire 4.1.2]. Since the special fiber of $\mathcal{Y}$ is reduced, we find that the top horizontal map is an isomorphism [Bosch et al. 1990, Theorem 9.5.4].

Let D denote the irreducible component of $\mathcal{N}_k$ lying above $\mathcal{Y}_k$. The composition $D \hookrightarrow \mathcal{N}_k \to \mathcal{Y}_k$ is an isomorphism. The image of D in $(\mathcal{X}')_k^{\mathrm{red}}$ is the curve C_0, and we will identify the map $D \to C_0$ with the quotient map $\rho : \mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$. Consider the following diagram, whose top right horizontal morphism is an isomorphism:

$$\begin{tikzcd}
\mathrm{Pic}^0_D(k) & \mathrm{Pic}^0_{\mathcal{N}_k}(k) \arrow[l] \arrow[r, "\sim"] & (Q_L)^0_k(k) \\
\mathrm{Pic}^0_{C_0}(k) \arrow[u, "\rho^*"] & \mathrm{Pic}^0_{\mathcal{X}'_k}(k) \arrow[l] \arrow[r] \arrow[u] & (Q_K)^0_k(k) \arrow[u]
\end{tikzcd}$$

We may now conclude the proof of Theorem 5.3 using the same method as in the case where the reduction of $\mathrm{Jac}(X)/K$ is purely unipotent. Consider again the element $\sum_{i=1}^{d-1} \gcd(r_i, r_d)S_i$ in $\mathrm{Jac}(X)(K)$, which reduces to the element $\sum_{i=1}^{d-1} \gcd(r_i, r_d)\tau_i$ in Φ_M. Proposition 3.5 shows that the latter element is zero in Φ_M. Thus, $\sum_{i=1}^{d-1} \gcd(r_i, r_d)S_i$ reduces in the connected component $(Q_K)^0_k$. Consider now the element $\sum_{i=1}^{d-1} \gcd(r_i, r_d)S_i$ in $\mathrm{Jac}(X_L)(L)$. Our discussion above shows that it reduces to the element $r_d\bigl(\sum_{i=1}^{d-1} r_i(P_i - P_d)\bigr)$ in $\mathrm{Jac}(\mathcal{Y}_k)(k)$.

Since $\mathrm{Pic}^0_{\mathcal{X}'_k/k} \to (Q_K)^0_k$ is a faithfully flat morphism and each of the above squares commutes, the element $\sum_{i=1}^{d-1} \gcd(r_i, r_d)S_i$, which reduces to $r_d\bigl(\sum_{i=1}^{d-1} r_i(P_i - P_d)\bigr)$ in $\mathrm{Pic}^0_{\mathcal{Y}_k/k}(k)$, in fact reduces to an element in $\rho^*(\mathrm{Jac}(\mathcal{Y}_k/\langle\sigma\rangle))$. Thus, the image of $r_d\bigl(\sum_{i=1}^{d-1} r_i(P_i - P_d)\bigr)$ in $\mathrm{Jac}(\mathcal{Y}_k)/\rho^*(\mathrm{Jac}(\mathcal{Y}_k/\langle\sigma\rangle))$ is trivial. Each $P_i - P_d$ defines an element of order p in $\mathrm{Jac}(\mathcal{Y}_k)/\rho^*(\mathrm{Jac}(\mathcal{Y}_k/\langle\sigma\rangle))$ (Proposition 2.5). Since $r_d\bigl(\sum_{i=1}^{d-1} r_i(P_i - P_d)\bigr) = 0$, we conclude that $\sum_{i=1}^{d-1} r_i(P_i - P_d) = 0$. Then Proposition 2.5 implies that p divides r_i for all $i = 1, \dots, d-1$, and since $|C_0 \cdot C_0|p = r_1 + \cdots + r_d$, we find that p divides r_d, which contradicts our assumption.

Now that we know that p divides r_i, we see that the multiplicities on the chain of G that leaves C_0 starting with D_i can only be increasing or constant because this chain of vertices of degree 2 contains no vertex of self-intersection -1. If D_i is not a node of G, we continue along this chain and find either a terminal vertex or a node of G. We cannot find a terminal vertex because the multiplicity of a terminal vertex can only be at most the multiplicity of its unique neighbor with equality only if the self-intersection of the terminal vertex is -1. Thus, G_{Q_i} contains a node of G. □

Remark 5.4. Let N_i denote the intersection matrix of the exceptional divisor, with smooth components and normal crossings, of a resolution of the $\mathbb{Z}/p\mathbb{Z}$-quotient singularity Q_i. We recall here some properties of N_i:

(a) It is negative definite (attributed to Du Val in [Lipman 1969, Lemma 14.1]).

(b) The graph $G(N_i)$ associated with N_i is a tree, and all components of the exceptional divisor are rational [Lorenzini 2013a, Theorem 2.8].

(c) Let n_i denote the number of irreducible components in the exceptional divisor. The Smith group $\Phi_{N_i} := \mathbb{Z}^{n_i}/\mathrm{Im}(N_i)$ is killed by p [Lorenzini 2013a, Theorem 2.6].

(d) The fundamental cycle Z of N_i is such that $|Z^2| \le p$ [Lorenzini 2013a, Theorem 2.3, Remark 2.4].

6. The weakly ramified case

We present in this section some applications of Theorem 5.3. Let us recall our notation. Let K be a complete discrete valuation field with valuation v, ring of integers $\mathcal{O}_K$, uniformizer π_K and residue field k of characteristic $p > 0$, assumed to be algebraically closed. Let X/K be a smooth proper geometrically connected curve of genus $g > 0$. When $g = 1$, we assume in addition that $X(K) \neq \varnothing$.

Assume that X/K does not have semistable reduction over $\mathcal{O}_K$ and that it achieves good reduction after a cyclic extension L/K of prime degree p. Let $H = \langle\sigma\rangle$ denote the Galois group of L/K. Let $\mathcal{Y}/\mathcal{O}_L$ be the smooth model of X_L/L. Let $\mathcal{Z}/\mathcal{O}_K$ denote the quotient $\mathcal{Y}/H$ with singular points $Q_1, \dots, Q_d$ and $d \geq 1$. Recall the regular model $f : \mathcal{X}' \to \mathcal{Z}$ introduced in 5.2.

6.1. The resolution of a singularity Q of $\mathcal{Z}$ is a local process and depends only on the local ring $\mathcal{O}_{\mathcal{Z},Q}$. It seems therefore natural to try to relate the "complexity" of the resolution graph to some local invariants of $\mathcal{O}_{\mathcal{Z},Q}$. In this respect, we propose the following.

Consider the Galois morphism $\rho : \mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$. Associated with any point $Q \in \mathcal{Y}_k/\langle\sigma\rangle$ is the following measure of the ramification of ρ over Q:

$$\nu(Q) := \delta(P) = \sum_{j=0}^{\infty}(|H_j(P)| - 1),$$

where P is the preimage of Q in $\mathcal{Y}_k$ and $H_j(P)$ denotes the j-th higher ramification group at P. (For more general morphisms, we would define $\nu(Q) := \sum_{P \in \rho^{-1}(Q)} \delta(P)$.) Recall from 2.2 that the morphism is *weakly ramified* at P if $\delta(P) = 2(p-1)$. Our guess is that $\nu(Q)$ should also be an important measure of how complicated the exceptional divisor of the resolution of Q is. To formulate this guess more precisely, we compare the expressions of the genus g in the Riemann–Hurwitz formula and in the adjunction formula. The Riemann–Hurwitz formula for the morphism ρ can be rephrased as

$$2g = 2g(\mathcal{Y}_k) = 2|H|g(C_0) - 2(|H| - 1) + \sum_{i=1}^{d} \nu(Q_i).$$

Consider now the model $\mathcal{X}'$. By hypothesis, it is minimal with the property that the special fiber has smooth components and normal crossings. Thus, none of the vertices A in the graph $G := G(\mathcal{X}')$ with degree 1 or 2 can have self-intersection -1 (we use here also the fact that only the curve C_0 can have positive genus [Lorenzini 2013a, Lemma 2.10]). Moreover, since the curve X/K has potentially good reduction, the graph $G(\mathcal{X}')$ is a tree [Lorenzini 2013a, Lemma 2.10].

The adjunction formula

$$2g - 2 = \mathcal{X}'_k \cdot \mathcal{X}'_k + \mathcal{X}'_k \cdot \Omega,$$

with Ω a relative canonical divisor of $\mathcal{X}'/\mathcal{O}_K$, can be rewritten as

$$\begin{aligned} 2g &= 2|H|g(C_0) + \sum_{\text{vertex } A \text{ of } G} (r(A)-1)(d_G(A)-2) \\ &= 2|H|g(C_0) - 2(|H|-1) + \sum_{i=1}^{d} \Bigl(|H| - 1 + \sum_{\text{vertex } A \text{ of } G_{Q_i}} (r(A)-1)(d_G(A)-2) \Bigr) \\ &= 2|H|g(C_0) - 2(|H|-1) + \sum_{i=1}^{d} \gamma_{D_i} g_{D_i}, \end{aligned} \tag{6.1.1}$$

where $D_1, \dots, D_d$ are the vertices attached to C_0 in the tree $G(\mathcal{X}')$ and the integers γ_{D_i} and g_{D_i} are defined as in 4.1 and (4.1.2). Since the graph G_{D_i} is nothing but the graph G_{Q_i} of the desingularization of Q_i, we define our measure of the desingularization of Q_i to be $\gamma_{Q_i} g_{Q_i} := \gamma_{D_i} g_{D_i}$ for each $i = 1, \dots, d$. The integer $g_{Q_i} := g_{D_i}$ depends only on the intersection matrix of the desingularization and the marked vertex D_i on its graph. Since $r(C_0) = p$ and is divisible by γ_{Q_i}, we find that $\gamma_{Q_i} = 1$ or p.

6.2. Our guess regarding the resolution $\mathcal{X}' \to \mathcal{Z}$ of the singularities of $\mathcal{Z}$ is that

$$\gamma_{Q_i} g_{Q_i} = \nu(Q_i) \text{ holds for all } i = 1, \dots, d.$$

This equality would have interesting implications. For instance, since $H = \mathbb{Z}/p\mathbb{Z}$, we always have $\nu(Q)$ divisible by $p-1$ so that $p-1$ divides $\gamma_{Q_i} g_{Q_i}$ when $\gamma_{Q_i} g_{Q_i} = \nu(Q_i)$. Since $\gamma_{Q_i} = 1$ or p, we find that

$$p - 1 \text{ divides } g_{Q_i} \text{ when } \gamma_{Q_i} g_{Q_i} = \nu(Q_i).$$

Examples where $g_{Q_i} = 2(p-1)$ and $3(p-1)$ are given in 4.7 and Remark 4.9. It immediately follows from the Riemann–Hurwitz formula and the adjunction formula that:

Lemma 6.3. *With the above notation and hypotheses,*

$$\sum_{i=1}^{d} \nu(Q_i) = \sum_{i=1}^{d} \gamma_{Q_i} g_{Q_i}. \tag{6.3.1}$$

We now prove the equality $\gamma_{Q_i} g_{Q_i} = \nu(Q_i) = 2(p-1)$ for all $i = 1, \dots, d$ in the weakly ramified case, using Theorem 5.3.

Theorem 6.4. *Let X/K be a curve with potentially good reduction after a ramified extension L/K of prime degree p. Keep the above notation. Then for all $i = 1, \dots, d$:*

(a) *We have $\gamma_{Q_i} g_{Q_i} \geq 2(p-1)$ and $\nu(Q_i) \geq 2(p-1)$.*

(b) *If the ramification points of $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ are all weakly ramified (in particular, if $\mathcal{Y}_k$ is ordinary), then $\gamma_{Q_i} g_{Q_i} = \nu(Q_i) = 2(p-1)$.*

Proof. (a) The fact that $\nu(Q_i) \geq 2(p-1)$ follows immediately from the properties of a wildly ramified extension: the higher ramification groups H_0 and H_1 must be nontrivial. To prove that $\gamma_{Q_i} g_{Q_i} \geq 2(p-1)$, we note first that Theorem 5.3 shows that $p \mid r_i$. The inequality follows then from Proposition 4.3.

(b) When the ramification points of $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ are all weakly ramified, we have $\nu(Q_i) = 2(p-1)$ (2.2). It follows from (6.3.1) and from the fact that $\gamma_{Q_i} g_{Q_i} \geq 2(p-1)$ proven in (a) that $\gamma_{Q_i} g_{Q_i} = 2(p-1)$. □

Remark 6.5. Without the use of Theorem 5.3, we could only argue that $\gamma_{Q_i} g_{Q_i} \geq p-1$. Indeed, if $r(C_0)$ does not divide $r(D_i)$, then $\gamma_{D_i} = 1$. Then we can use the fact that $g_{Q_i} \geq r(C_0) - 1$ established in Remark 4.10.

Using the notation γ_{Q_i} introduced in this section, we may now state a corollary to Theorem 5.3.

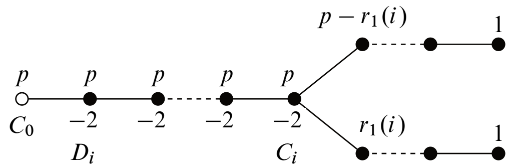
Corollary 6.6. *Let X/K be a curve with potentially good reduction after a wildly ramified Galois extension L/K of degree p as in Theorem 5.3. Let N_i denote the intersection matrix associated with the resolution of Q_i. Assume that $\gamma_{Q_i} = 1$. Then p^2 divides* $\det(N_i)$.

Proof. The graph associated with the matrix N_i is G_{Q_i} with a marked vertex D_i on it. Let R_{D_i} denote the vector of multiplicities of the components of the resolution of Q_i. Then the determinant of N_i can be computed in terms of the coefficients of R_{D_i}/γ_{D_i} (see [Lorenzini 2013a, Theorem 3.14]). In particular, it is known that $(r(C_0)/\gamma_{D_i}) \gcd(r(C_0)/\gamma_{D_i}, r(D_i)/\gamma_{D_i})$ divides $\det(N_i)$. Under our hypotheses, $r(C_0) = p$, p divides $r(D_i)$ (Theorem 5.3) and $\gamma_{D_i} = 1$. □

Remark 6.7. Let X/K be a curve with potentially good reduction after a wildly ramified extension L/K of degree p as in Theorem 5.3. Let N_i denote the intersection matrix associated with the resolution of Q_i. Then p kills the Smith group Φ_{N_i} [Lorenzini 2013a, Theorem 2.6], and thus, $|\det(N_i)|$ is a power of p. It follows from (4.10.3) that $\mathrm{ord}_p(|\det(N_i)|)(p-1) \leq g_{D_i}$.

In the examples of graphs and matrices N_i given in Remark 4.9 with $g_{D_i} = 3(p-1)$, we find that both $|\det(N_i)| = p^2$ and $|\det(N_i)| = p^3$ can occur: the former in (b) and (c)(ii), and the latter in (a) and (c)(i).

Theorem 6.8. *Let X/K be a curve with potentially good reduction after a wildly ramified Galois extension L/K of degree p. Assume that all ramification points of $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ are weakly ramified (this is the case if $\mathcal{Y}_k$ is ordinary). Keep the above notation. Then, for all $i = 1, \dots, d$, we have $r_i = p$, and G_{Q_i} is a graph*[1] *with a single node C_i of degree* 3:



The intersection matrix $N(p, \alpha_i, r_1(i))$ of the resolution of Q_i is uniquely determined as in 4.7 by the two integers α_i and $r_1(i)$ with $1 \le r_1(i) < p$. The integer α_i is the number of vertices of self-intersection -2 (including the node C_i) on the chain in G_{Q_i} connecting the node C_0 to the single node C_i of G_{Q_i}, and this integer α_i is divisible by p.

Proof. Theorem 6.4(b) shows that $\gamma_{Q_i} g_{Q_i} = 2(p-1)$ for all $i = 1, \dots, d$. Proposition 4.3 classifies the graphs with $\gamma_{Q_i} g_{Q_i} = 2(p-1)$, and the statement on the shape of the graph follows.

The Smith group of the intersection matrix $N(p, \alpha_i, r_1(i))$ is computed in [Lorenzini 2013a, §3.19, Lemma 3.21] and is found to be of order p^2 and killed by p if and only if p divides α_i. Theorem 2.6(c) of [Lorenzini 2013a] shows that this Smith group must be killed by p. The divisibility $p \mid \alpha_i$ follows. □

Remark 6.9. It is natural to wonder whether the statements of Theorems 6.4(b) and 6.8 hold for the resolution of Q_i when P_i is a weakly ramified ramification point of $\mathcal{Y}_k \to \mathcal{Y}_k\langle\sigma\rangle$ without also assuming as we do in Theorems 6.4(b) and 6.8 that all ramification points are weakly ramified.

Corollary 6.10. *Let X/K be a curve with potentially good reduction after a wildly ramified Galois extension L/K of degree p as in Theorem 6.8. Suppose that $g > 1$ and that all ramification points of $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ are weakly ramified. Then*:

(a) $X(K) \neq \varnothing$.

(b) *Let A/K denote the Jacobian of X/K. Let $\mathcal{A}/\mathcal{O}_K$ be its Néron model. Then the unipotent part U/k of the connected component of the identity in $\mathcal{A}_k/k$ is a product of additive groups $\mathbb{G}_{a,k}$.*

(c) *The group of components $\Phi_{A,K}$ of the Néron model is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^{2d-2}$.*

[1] A bullet • represents an irreducible component of the desingularization of Q_i. A positive number next to a vertex is the multiplicity of the corresponding component while a negative number next to a vertex is the self-intersection of the component.

Proof. Part (a) is immediate since it follows from Theorem 6.8 that a regular model of X/K contains a component of multiplicity 1. It follows from [Penniston 2000, Theorem 2.4] that p kills U since the maximal multiplicity in the regular model $\mathcal{X}'/\mathcal{O}_K$ is equal to p. That U is now split follows from [Serre 1959, Proposition VII.11.11]. This proves (b).

The order of $\Phi_{A,K}$ can be computed using the intersection matrix of the regular model $\mathcal{X}'$. Since the associated graph is a tree, we find using [Lorenzini 1989, Corollary 2.5] that $|\Phi_{A,K}| = p^{2d-2}$. Part (c) follows since $\Phi_{A,K}$ is killed by $[L:K]$ because A/K has potentially good reduction [Edixhoven et al. 1996]. $\square$

Note that in general the special fiber $\mathcal{A}_k/k$ need not be killed by p even when its subgroup U and quotient $\Phi_{A,K}$ are both killed by p (see [Liu and Lorenzini 2001] for a general discussion of such phenomena).

6.11. Let A/K be the Jacobian of a smooth proper and geometrically connected curve X/K having a K-rational point. For use in our next corollary, we recall below the main result of [Bosch and Lorenzini 2002, Theorem 4.6]. Identify A/K with its dual A'/K via the map $-\varphi_{[\Theta]} : A \to A'$ as in [Bosch and Lorenzini 2002] just before Theorem 4.6. Let $\mathcal{X}/\mathcal{O}_K$ denote a regular model of X/K. Let M be the intersection matrix of $\mathcal{X}_k$. Identify, as recalled in [Bosch and Lorenzini 2002, Theorem 2.3], the component group $\Phi_{A/K}$ with the group of components Φ_M of M (Φ_M is the torsion subgroup of $\mathbb{Z}^v/\operatorname{Im}(M)$). Then Grothendieck's pairing

$$\langle\,\cdot\,,\,\cdot\,\rangle_K : \Phi_{A/K} \times \Phi_{A/K} \to \mathbb{Q}/\mathbb{Z}$$

coincides with the pairing $\langle\,\cdot\,,\,\cdot\,\rangle_M : \Phi_{A/K} \times \Phi_{A/K} \to \mathbb{Q}/\mathbb{Z}$ considered in 3.1. In particular, this pairing is nondegenerate. Recall also the definition of the functorial subgroup $\Phi^0_{A/K}$ of $\Phi_{A/K}$ in 1.3. We denote by $(\Phi^0_{A/K})^\perp$ the orthogonal of $\Phi^0_{A/K}$ under Grothendieck's pairing.

Corollary 6.12. *Let A/K be the Jacobian of a curve X/K of genus $g > 1$ having potentially good reduction after a Galois extension L/K of degree p as in Theorem 6.8. Assume that all ramification points of $\mathcal{Y}_k \to \mathcal{Y}_k/\langle\sigma\rangle$ are weakly ramified. Then $\Phi_{A/K}$ is a $\mathbb{Z}/p\mathbb{Z}$-vector space of dimension $2d-2$, and $\Phi^0_{A/K}$ is a subspace of dimension $d-1$. Moreover, $\Phi^0_{A/K} = (\Phi^0_{A/K})^\perp$.*

Proof. It follows from Corollary 6.10 that $X(K) \neq \varnothing$. We can thus use the results of [Bosch and Lorenzini 2002] recalled above. We produce below explicit generators for the groups $\Phi_{A/K}$ and $\Phi^0_{A/K}$. For each singular point Q_i on the model $\mathcal{Z}/\mathcal{O}_K$, denote by A_i and B_i the terminal components of multiplicity 1 in the exceptional divisor of the resolution of Q_i in $\mathcal{X}'$. Let $\boldsymbol{\alpha}_i$ denote the image in $\Phi_{A/K}$ of the vector $E(A_i, B_i)$, $i = 1, \ldots, d-1$ (notation as in 3.1). Let $\boldsymbol{\beta}_i$ denote the image in $\Phi_{A/K}$ of the vector $E(A_i, A_d)$, $i = 1, \ldots, d-1$. We have seen in Corollary 6.10 that $\Phi_{A/K}$ is a $\mathbb{Z}/p\mathbb{Z}$-vector space of dimension $2(d-1)$.

We claim that

$$\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{d-1}, \boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{d-1}\}$$

is a basis for $\Phi_{A/K}$ *and that* $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{d-1}\}$ *is a basis for* $\Phi^0_{A/K}$. To prove our claim, consider the matrix $V := (\langle \boldsymbol{\alpha}_i, \boldsymbol{\beta}_j \rangle)_{1 \le i,j \le d-1}$ with coefficients in $\mathbb{Q}/\mathbb{Z}$. We can use the computation (4.8.1) to show that V is the diagonal matrix

$$\operatorname{diag}\bigl(c_1/p \ (\mathrm{mod}\ \mathbb{Z}), \dots, c_{d-1}/p \ (\mathrm{mod}\ \mathbb{Z})\bigr),$$

where, for each $i = 1, \dots, d-1$, $0 < c_i < p$ and p divides $c_i r_1(i) - 1$. In particular, $c_i/p \neq 0$ in $\mathbb{Q}/\mathbb{Z}$. It follows that the set $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{d-1}, \boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{d-1}\}$ is linearly independent in $(\mathbb{Z}/p\mathbb{Z})^{2d-2}$. Hence, it is a basis.

It follows from the explicit computations in [Lorenzini 2000, Proposition 3.7(a)], that $\langle \boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j \rangle = 0$ for all $1 \le i, j \le d-1$. Since the pairing $\langle \cdot, \cdot \rangle$ is perfect on $(\mathbb{Z}/p\mathbb{Z})^{2d-2}$, we find that $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{d-1}\}$ generates a maximal isotropic subspace.

It remains to show that $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{d-1}$ belong to $\Phi^0_{A/K}$ and that neither $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{d-1}$ nor any nontrivial linear combination of $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{d-1}$ belong to $\Phi^0_{A/K}$. For this, since K is complete, we can pick for each $i = 1, \dots, d-1$ two K-rational points a_i and b_i of X whose closure in $\mathcal{X}'$ intersects $\mathcal{X}'_k$ in a smooth point of A_i and B_i, respectively (see, e.g., [Bosch et al. 1990, Corollary 9.1.9]). Then $a_i - b_i$ and $a_i - a_d$ are divisors of degree 0 on X, which we identify with K-rational points in the Jacobian A/K of X/K. These rational points reduce in the component group $\Phi_{A/K}$ of the Néron model of A/K to the points $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$, respectively. Since $A(K) \subset A(L)$, we can reduce $a_i - b_i$ in the special fiber of the Néron model $\mathcal{A}'/\mathcal{O}_L$. This special fiber is isomorphic to the Jacobian of the special fiber $\mathcal{Y}_k$ of the smooth model $\mathcal{Y}/\mathcal{O}_L$ of X_L/L. It is clear that, by construction, the reduction of $a_i - b_i$ is trivial so that $\boldsymbol{\alpha}_i \in \Phi^0_{A/K}$ for $i = 1, \dots, d-1$. On the other hand, the reduction of $a_i - a_d$ is the divisor $P_i - P_d$, which is a nontrivial p-torsion point when viewed in the quotient $\mathcal{A}'_k/\eta(\mathcal{A}_k)$. This shows that $\boldsymbol{\beta}_i \notin \Phi^0_{A/K}$ for $i = 1, \dots, d-1$. Moreover, any nontrivial linear combination of the images of the divisors $P_i - P_d$ is not zero in $\mathcal{A}'_k/\eta(\mathcal{A}_k)$ (Proposition 2.5), so no nontrivial linear combination of $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{d-1}$ belongs to $\Phi^0_{A/K}$. □

Example 6.13. Examples of curves having good reduction after an extension of degree p can be obtained as twists as follows. Choose a smooth proper curve C/k having an automorphism σ_k of order p. Over an appropriate ring $\mathcal{O}_K$ with residue field k, there exists a smooth scheme $\mathcal{Y}^0/\mathcal{O}_K$ with an $\mathcal{O}_K$-automorphism σ such that C is k-isomorphic to $\mathcal{Y}^0_k$ and σ restricted to $\mathcal{Y}^0_k$ induces the given automorphism σ_k. It is shown in [Sekiguchi et al. 1989, §IV, Theorem 2.2] that one can take $\mathcal{O}_K$ to be the Witt ring $W(k)(\zeta_p)$ with ζ_p a primitive p-th root of unity. If one wants a lift in equicharacteristic p, one can trivially take $\mathcal{O}_K = k[\![t]\!]$.

Choose any cyclic (ramified) extension L/K of degree p. The twist of $\mathcal{Y}_K^0/K$ by L/K and σ is a curve X/K that achieves good reduction over L. Starting with an ordinary curve C/k produces a curve X/K having potentially good ordinary reduction over L.

Corollary 6.14. *Fix any odd prime p. For each integer $m > 0$, there exist a regular local ring B of equicharacteristic p endowed with an action of $H := \mathbb{Z}/p\mathbb{Z}$ and a regular local ring B' of mixed characteristic $(0, p)$ endowed with an action of $\mathbb{Z}/p\mathbb{Z}$ such that $\operatorname{Spec} B^H$ and $\operatorname{Spec}(B')^H$ are singular exactly at their closed point, and the graphs associated with a minimal resolution of $\operatorname{Spec} B^H$ and $\operatorname{Spec}(B')^H$ have one node and more than m vertices.*

Proof. As we noted in Example 6.13, there exist a field K of either mixed characteristic $(0, p)$ or of equicharacteristic p and a curve X/K without good reduction over K and with good ordinary reduction over a Galois extension L/K of degree p. Let $H := \operatorname{Gal}(L/K)$. Let $\mathcal{Y}/\mathcal{O}_L$ denote the smooth model of X_L/L. Let $\mathcal{Z}/\mathcal{O}_K$ denote the quotient $\mathcal{Y}/H$. Let P denote a ramification point of the morphism $\mathcal{Y}_k \to \mathcal{Y}_k/H$, and let $B := \mathcal{O}_{\mathcal{Y},P}$. Theorem 6.8 shows that the resolution of singularity of $\operatorname{Spec} B^H$ has an intersection matrix of type $N(p, \alpha, r_1)$ for some $\alpha \geq 1$ and $0 < r_1 < p$.

Immediately after the statement of Theorem 6.8 given in the introduction, we briefly alluded to the fact that the integer α is likely to be related to the valuation of the different of L/K. Thus, in principle, by choosing K and L/K appropriately, the above method will produce examples with α as large, as desired. Since at this time we do not know how to prove in general that α is related to the valuation of the different of L/K (except when $p = 2$ and $g = 1$; see [Lorenzini 2013a, Theorem 4.1]), we proceed below with a different argument to prove the existence of resolutions with α as large, as desired.

Consider a quadratic extension K'/K. Since p is odd by hypothesis, the extension K'/K is tame, and one knows how to compute a regular model of $X_{K'}/K'$ from the model $\mathcal{X}/\mathcal{O}_K$ of X/K obtained in Theorem 6.8: simply normalize the base change $\mathcal{X} \times_{\mathcal{O}_K} \mathcal{O}_{K'}$ and resolve its singularities. A singularity on the normalization can only be the preimage of a closed point of $\mathcal{X}_k$ that belongs to two irreducible components of $\mathcal{X}_k$ and such that both components have odd multiplicity. This singular point is resolved by a single smooth rational curve.

Let $L' := LK'$ with $[L' : K'] = p$. The curve $X_{K'}/K'$ achieves good ordinary reduction over L'. The model $\mathcal{Y}'/\mathcal{O}_{L'} := \mathcal{Y} \times_{\mathcal{O}_L} \mathcal{O}_{L'}$ is smooth, and we let P' denote the preimage of P under the natural map $\mathcal{Y}' \to \mathcal{Y}$. Let $B' := \mathcal{O}_{\mathcal{Y}',P'}$. We leave it to the reader to check, using [Halle 2010, Proposition 4.3] and the desingularization of the normalization of $\mathcal{X} \times_{\mathcal{O}_K} \mathcal{O}_{K'}$, that the resolution of the singularity of $\operatorname{Spec}(B')^H$ has an intersection matrix of type $N(p, 2\alpha, r_1')$, where $r_1' := r_1/2$ if r_1 is even and $r_1' := (r_1 + p)/2$ if r_1 is odd.

Since we can make an infinite chain of quadratic extensions $K \subset K' \subset K'' \subset \cdots$ and since the graph associated with $N(p, \beta, r_1)$ has at least β irreducible components, the corollary is proved. $\square$

Remark 6.15. Consider an intersection matrix N, and assume that, for some prime p, it satisfies all the conditions listed in Remark 5.4, conditions that would have to be satisfied if this intersection matrix was associated with the resolution of a $\mathbb{Z}/p\mathbb{Z}$-singularity: its graph $G(N)$ is a tree, $|\det(N)|$ is a power of p, the Smith group Φ_N is killed by p and the fundamental cycle Z has $|Z^2| \leq p$. If $\det(N) = 1$ and $G(N)$ is a tree, then the above conditions are satisfied for every prime at least equal to $|Z^2|$. In particular, when $\det(N) = 1$, the matrix N could potentially be associated with the resolution of a $\mathbb{Z}/p\mathbb{Z}$-singularity for infinitely many primes p.

An interesting consequence of our guess in 6.2 that $\gamma_{Q_i} g_{Q_i} = \nu(Q_i)$ holds for all $i = 1, \ldots, d$ is that a matrix N as above can be associated with the resolution of a $(\mathbb{Z}/p\mathbb{Z})$-quotient singularity $\mathcal{X}' \to \mathcal{Z}$ occurring in models of curves as at the beginning of this section *only for finitely many primes p*. Indeed, the choice of a vertex D on N lets us define the integer g_D associated with N and D. If N is the intersection matrix of the resolution of a singularity Q_i of $\mathcal{Z}$ with the marked vertex D linked to C_0, we noted in 6.2 that $p - 1$ must then divide g_D when the equality $\gamma_{Q_i} g_{Q_i} = \nu(Q_i)$ holds. Since there are only finitely many vertices D, the set of integers g_D is finite, and hence, any prime p larger than the maximum of the integers g_D cannot have the property that $p - 1$ divides some g_D.

Remark 6.16. Let X/K be a curve with potentially good reduction over an extension L/K of degree p as at the beginning of this section. Let Q_i be a singular point of the quotient $\mathcal{Z}$, and consider the graph G_{Q_i} associated with the resolution of Q_i in $\mathcal{X}' \to \mathcal{Z}$. One may wonder whether a node of G in G_{Q_i} could have its multiplicity in $\mathcal{X}'_k$ divisible by p^2. Similar considerations are found in [Lorenzini 2010, Question 1.4].

Acknowledgements

It is my pleasure to thank Qing Liu, Werner Lütkebohmert and Michel Raynaud for helpful comments. I also thank the referee for a careful reading of the article.

References

[Altman and Kleiman 1970] A. Altman and S. Kleiman, *Introduction to Grothendieck duality theory*, Lecture Notes in Mathematics **146**, Springer, Berlin, 1970. MR 43 #224 Zbl 0215.37201

[Artin 1975] M. Artin, "Wildly ramified $Z/2$ actions in dimension two", *Proc. Amer. Math. Soc.* **52** (1975), 60–64. MR 51 #10336 Zbl 0315.14015

[Bosch and Lorenzini 2002] S. Bosch and D. Lorenzini, "Grothendieck's pairing on component groups of Jacobians", *Invent. Math.* **148**:2 (2002), 353–396. MR 2003b:14053 Zbl 1061.14042

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **21**, Springer, Berlin, 1990. MR 91i:14034 Zbl 0705.14001

[Edixhoven et al. 1996] B. Edixhoven, Q. Liu, and D. Lorenzini, "The p-part of the group of components of a Néron model", *J. Algebraic Geom.* **5**:4 (1996), 801–813. MR 98m:14051 Zbl 0898.14007

[Gabber et al. 2013] O. Gabber, Q. Liu, and D. Lorenzini, "The index of an algebraic variety", *Invent. Math.* **192**:3 (2013), 567–626. MR 3049930 Zbl 1268.13009

[Halle 2010] L. H. Halle, "Stable reduction of curves and tame ramification", *Math. Z.* **265**:3 (2010), 529–550. MR 2011e:14030 Zbl 1195.14011

[Hasse 1934] H. Hasse, "Theorie der relativ-zyklischen algebraischen Funktionenkörper, insbesondere bei endlichem Konstantenkörper", *J. Reine Angew. Math.* **172** (1934), 37–54. Zbl 0010.00501

[Katsura 1978] T. Katsura, "On Kummer surfaces in characteristic 2", pp. 525–542 in *Proceedings of the International Symposium on Algebraic Geometry* (Kyoto, 1977), edited by M. Nagata, Kinokuniya Book Store, Tokyo, 1978. MR 82d:14022 Zbl 0411.14010

[Lipman 1969] J. Lipman, "Rational singularities, with applications to algebraic surfaces and unique factorization", *Inst. Hautes Études Sci. Publ. Math.* **36** (1969), 195–279. MR 43 #1986 Zbl 0181.48903

[Liu and Lorenzini 2001] Q. Liu and D. Lorenzini, "Special fibers of Néron models and wild ramification", *J. Reine Angew. Math.* **532** (2001), 179–222. MR 2001m:11094 Zbl 0970.14025

[Liu et al. 2004] Q. Liu, D. Lorenzini, and M. Raynaud, "Néron models, Lie algebras, and reduction of curves of genus one", *Invent. Math.* **157**:3 (2004), 455–518. MR 2005m:14039 Zbl 1060.14037

[Lorenzini 1989] D. Lorenzini, "Arithmetical graphs", *Math. Ann.* **285**:3 (1989), 481–501. MR 91b: 14026 Zbl 0662.14008

[Lorenzini 1990a] D. Lorenzini, "Dual graphs of degenerating curves", *Math. Ann.* **287**:1 (1990), 135–150. MR 91j:14018 Zbl 0668.14019

[Lorenzini 1990b] D. Lorenzini, "Groups of components of Néron models of Jacobians", *Compositio Math.* **73**:2 (1990), 145–160. MR 92d:14019 Zbl 0737.14008

[Lorenzini 1993] D. Lorenzini, "On the group of components of a Néron model", *J. Reine Angew. Math.* **445** (1993), 109–160. MR 94k:11065 Zbl 0781.14029

[Lorenzini 2000] D. Lorenzini, "Reduction of points in the group of components of the Néron model of a Jacobian", *J. Reine Angew. Math.* **527** (2000), 117–150. MR 2001i:11073 Zbl 0983.14022

[Lorenzini 2010] D. Lorenzini, "Models of curves and wild ramification", *Pure Appl. Math. Q.* **6**:1 (2010), 41–82. MR 2012a:11068 Zbl 1200.14052

[Lorenzini 2013a] D. Lorenzini, "Wild quotient singularities of surfaces", *Math. Z.* **275**:1–2 (2013), 211–232. MR 3101805 Zbl 06224338

[Lorenzini 2013b] D. Lorenzini, "Wild quotients of products of curves", preprint, 2013, available at http://www.math.uga.edu/~lorenz/Paper3.pdf.

[Lorenzini and Tucker 2002] D. Lorenzini and T. J. Tucker, "Thue equations and the method of Chabauty–Coleman", *Invent. Math.* **148**:1 (2002), 47–77. MR 2003d:11088 Zbl 1048.11023

[Penniston 2000] D. Penniston, "Unipotent groups and curves of genus two", *Math. Ann.* **317**:1 (2000), 57–78. MR 2001b:14050 Zbl 1005.14010

[Peskin 1983] B. R. Peskin, "Quotient-singularities and wild p-cyclic actions", *J. Algebra* **81**:1 (1983), 72–99. MR 84k:13018 Zbl 0517.14017

[Raynaud 1970] M. Raynaud, "Spécialisation du foncteur de Picard", *Inst. Hautes Études Sci. Publ. Math.* **38** (1970), 27–76. MR 44 #227 Zbl 0207.51602

[Sekiguchi et al. 1989] T. Sekiguchi, F. Oort, and N. Suwa, "On the deformation of Artin–Schreier to Kummer", *Ann. Sci. École Norm. Sup.* (4) **22**:3 (1989), 345–375. MR 91g:14041 Zbl 0714.14024

[Serre 1959] J.-P. Serre, *Groupes algébriques et corps de classes*, Publications de l'institut de mathématique de l'université de Nancago, VII. Hermann, Paris, 1959. MR 21 #1973 Zbl 0097.35604

[Singh 1974] B. Singh, "On the group of automorphisms of function field of genus at least two", *J. Pure Appl. Algebra* **4** (1974), 205–229. MR 50 #13047 Zbl 0284.12007

[Subrao 1975] D. Subrao, "The p-rank of Artin–Schreier curves", *Manuscripta Math.* **16**:2 (1975), 169–193. MR 51 #12868 Zbl 0321.14017

[Viehweg 1977] E. Viehweg, "Invarianten der degenerierten Fasern in lokalen Familien von Kurven", *J. Reine Angew. Math.* **293/294** (1977), 284–308. MR 58 #16655 Zbl 0349.14017

[Winters 1974] G. B. Winters, "On the existence of certain families of curves", *Amer. J. Math.* **96** (1974), 215–228. MR 50 #9874 Zbl 0334.14004

Communicated by Brian Conrad
Received 2013-01-03 Revised 2013-06-06 Accepted 2013-07-16

lorenzin@uga.edu *Department of Mathematics, University of Georgia, Athens, GA 30602, United States*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.369

Geometry of Wachspress surfaces

Corey Irving and Hal Schenck

Let P_d be a convex polygon with d vertices. The associated Wachspress surface W_d is a fundamental object in approximation theory, defined as the image of the rational map

$$\mathbb{P}^2 \xrightarrow{w_d} \mathbb{P}^{d-1},$$

determined by the Wachspress barycentric coordinates for P_d. We show w_d is a regular map on a blowup X_d of $\mathbb{P}^2$ and, if $d > 4$, is given by a very ample divisor on X_d so has a smooth image W_d. We determine generators for the ideal of W_d and prove that, in graded lex order, the initial ideal of I_{W_d} is given by a Stanley–Reisner ideal. As a consequence, we show that the associated surface is arithmetically Cohen–Macaulay and of Castelnuovo–Mumford regularity 2 and determine all the graded Betti numbers of I_{W_d}.

1. Introduction

Introduced by Möbius [1827], barycentric coordinates for triangles appear in a host of applications. Recent work in approximation theory has shown that it is also useful to define barycentric coordinates for a convex polygon P_d with $d \geq 4$ vertices (a d-gon). The idea is as follows. To deform a planar shape, first place the shape inside a control polygon. Then move the vertices of the control polygon, and use barycentric coordinates to extend this motion to the entire shape.

For a d-gon with $d \geq 4$, barycentric coordinates were defined by Wachspress [1975] in his work on finite elements; these coordinates are rational functions depending on the vertices $\nu(P_d)$ of P_d. Warren [2003] shows that Wachspress' coordinates are the unique rational barycentric coordinates of minimal degree. The Wachspress coordinates define a rational map w_d on $\mathbb{P}^2$, whose value at a point $p \in P_d$ is the d-tuple of barycentric coordinates of p. The closure of the image of w_d is the Wachspress surface W_d, first defined and studied by Garcia-Puente and Sottile [2010] in their work on linear precision.

Irving is supported by Texas Advanced Research Program 010366-0054-2007. Schenck is supported by NSF 1068754 and NSA H98230-11-1-0170.
MSC2010: primary 13D02; secondary 52C35, 14J26, 14C20.
Keywords: barycentric coordinates, Wachspress variety, rational surface.

In Definition 1.3, we fix linear forms ℓ_i that are positive inside P_d and vanish on an edge. Let $A = \ell_1 \cdots \ell_d$, Z be the $\binom{d}{2}$ singular points of $\mathbb{V}(A)$, and $Y = Z \setminus \nu(P_d)$. We call Y the *external vertices* of P_d and show that w_d has basepoints only at Y. Let X_d be the blowup of $\mathbb{P}^2$ at Y. In Section 2, we prove that W_d is the image of X_d, embedded by a certain divisor D_{d-2} on X_d. The global sections of D_{d-2} have a simple interpretation in terms of the edges $\mathbb{V}(\ell_i)$ of P_d: we prove that

$$H^0(\mathbb{O}_{X_d}(D_{d-2})) \quad \text{has basis} \quad \{\ell_3 \cdots \ell_d, \ell_1\ell_4 \cdots \ell_d, \ldots, \ell_2 \cdots \ell_{d-1}\}.$$

We show that D_{d-2} is very ample if $d > 4$; hence, $W_d \subseteq \mathbb{P}^{d-1}$ is a smooth surface.

1A. ***Statement of main results.*** For a d-gon P_d with $d \geq 4$:

(1) We give explicit generators for $I_{W_d} \subseteq S = \mathbb{K}[x_1, \ldots, x_d]$.

(2) We determine $\mathrm{in}_\prec(I_{W_d})$, where $\prec$ is graded lex order.

(3) We prove $\mathrm{in}_\prec(I_{W_d})$ is the Stanley–Reisner ideal of a graph Γ.

(4) We prove that S/I_{W_d} is Cohen–Macaulay, and $\mathrm{reg}(S/I_{W_d}) = 2$.

(5) We determine the graded Betti numbers of S/I_{W_d}.

In Section 1B, we give some quick background on geometric modeling, and in Section 1C, we do the same for algebraic geometry (in particular, we define all the terms above). Our strategy runs as follows. In Section 2, we study I_{W_d} by blowing up $\mathbb{P}^2$ at the external vertices. Define a divisor

$$D_{d-2} = (d-2)E_0 - \sum_{p \in Y} E_p$$

on X_d, where E_0 is the pullback of a line and E_p is the exceptional fiber over p. We show that D_{d-2} is very ample and that I_{W_d} is the ideal of the image of

$$X_d \to \mathbb{P}(H^0(D_{d-2})).$$

Riemann–Roch then yields the Hilbert polynomial of S/I_{W_d}.

In Sections 3 and 4, we find distinguished sets of quadrics and cubics vanishing on W_d and use them to generate a subideal $I(d) \subseteq I_{W_d}$. In Section 5, we tie everything together, showing that, in graded lex order, $I_\Gamma(d) \subseteq \mathrm{in}_\prec I(d)$, where $I_\Gamma(d)$ is the Stanley–Reisner ideal of a certain graph. Using results on flat deformations and an analysis of associated primes, we prove

$$I_\Gamma(d) = \mathrm{in}_\prec(I(d)).$$

The description in terms of the Stanley–Reisner ring yields the Hilbert series for $S/I_\Gamma(d)$. We prove that $S/I_\Gamma(d)$ is Cohen–Macaulay and has Castelnuovo–Mumford regularity 2, and it follows from uppersemicontinuity that the same is true for $S/I(d)$. The differentials on the quadratic generators of $I_\Gamma(d)$ turn out to be easy to describe, and combining this with the regularity bound and knowledge of the Hilbert series yields the graded Betti numbers for $\mathrm{in}_{\prec}(I(d))$.

Finally, we show that $I(d)$ has no linear syzygies on its quadratic generators, which allows us to prune the resolution of $\mathrm{in}_{\prec}(I(d))$ to obtain the graded Betti numbers of $I(d)$. Comparing Hilbert polynomials shows that up to saturation

$$S/I(d) = S/I_{W_d}.$$

Since I_{W_d} is prime, it is saturated, and a short-exact-sequence argument shows that $S/I(d)$ is also saturated, concluding the proof.

1B. ***Geometric modeling background.*** Let P_d be a d-gon with vertices $v_1, \dots, v_d$ and indices taken modulo d.

Definition 1.1. Functions $\{\beta_i : P_d \to \mathbb{R} \mid 1 \le i \le d\}$ are *barycentric coordinates* if, for all $p \in P_d$,

$$\beta_i(p) \ge 0, \qquad p = \sum_{i=1}^{d} \beta_i(p) v_i, \qquad \sum_{i=1}^{d} \beta_i(p) = 1.$$

Wachspress coordinates have a geometric description in terms of areas of subtriangles of the polygon. Let $A(a, b, c)$ denote the area of the triangle with vertices a, b, and c. For $1 \le j \le d$, set $\alpha_j := A(v_{j-1}, v_j, v_{j+1})$ and $A_j := A(p, v_j, v_{j+1})$.

Definition 1.2. For $1 \le i \le d$, the functions

$$\beta_i = \frac{b_i}{\sum_{j=1}^{d} b_j}, \quad \text{where } b_i = \alpha_i \prod_{j \ne i-1, i} A_j$$

are Wachspress barycentric coordinates for the d-gon P_d; see Figure 1.

We embed P_d in the plane $z = 1 \subseteq \mathbb{R}^3$ and form the cone with $\mathbf{0} \in \mathbb{R}^3$. Explicitly, to each vertex $v_i \in \nu(P_d)$, we associate the ray $\boldsymbol{v}_i := (v_i, 1) \in \mathbb{R}^3$. Let $\boldsymbol{P}_d$ denote the cone generated by the rays $\boldsymbol{v}_i$, and $\nu(\boldsymbol{P}_d) := \{\boldsymbol{v}_i \mid v_i \in \nu(P_d)\}$. The cone over

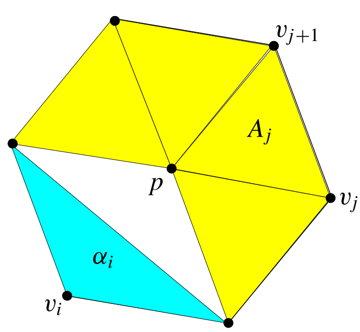


Figure 1. Wachspress coordinates for a polygon.

the edge $[v_i, v_{i+1}]$ corresponds to a facet of $\boldsymbol{P}_d$ with normal vector $\boldsymbol{n}_i := \boldsymbol{v}_i \times \boldsymbol{v}_{i+1}$. We redefine α_j and A_j to be the determinants $|\boldsymbol{v}_{j-1}\boldsymbol{v}_j\boldsymbol{v}_{j+1}|$ and $|\boldsymbol{v}_j\boldsymbol{v}_{j+1}\boldsymbol{p}|$, where $\boldsymbol{p} = (x, y, z)$. This scales the b_i by a factor of 2 so leaves the β_i unchanged, save for homogenizing the A_j with respect to z, and allows us to define Wachspress coordinates for nonconvex polygons, although Property 1 of barycentric coordinates fails when P_d is nonconvex.

Definition 1.3. $$\ell_j := A_j = \boldsymbol{n}_j \cdot \boldsymbol{p} = |\boldsymbol{v}_j\boldsymbol{v}_{j+1}\boldsymbol{p}|.$$

The ℓ_j are homogeneous linear forms in (x, y, z) and vanish on the cone over the edge $[v_j, v_{j+1}]$. We use Theorem 1.6 below, but Warren's proof does not require convexity. Our results hold over an arbitrary field $\mathbb{K}$ as long as no three of the lines $\mathbb{V}(\ell_i) \subseteq \mathbb{P}^2$ meet at a point. For the first condition of Definition 1.1 to make sense, $\mathbb{K}$ should be an ordered field.

Definition 1.4. The *dual cone* to $\boldsymbol{P}_d$ is the cone spanned by the normals $\boldsymbol{n}_1, \dots, \boldsymbol{n}_d$ and is denoted $\boldsymbol{P}_d^*$.

Triangulating P_d yields a triangulation of $\boldsymbol{P}_d$, and the volume of the parallelepiped S spanned by vertices $\{\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k, \boldsymbol{0}\}$ is $a_S = |\boldsymbol{v}_i\boldsymbol{v}_j\boldsymbol{v}_k|$.

Definition 1.5. Let C be a cone defined by a polygon P_d and $T(C)$ a triangulation of C obtained from a triangulation of P_d as above. The *adjoint* of C is

$$\mathcal{A}_{T(C)}(\boldsymbol{p}) = \sum_{S \in T(C)} a_S \prod_{\boldsymbol{v} \in \nu(\boldsymbol{P}_d) \setminus \nu(S)} (\boldsymbol{v} \cdot \boldsymbol{p}) \in \mathbb{K}[x, y, z]_{d-3}.$$

Theorem 1.6 [Warren 1996]. *$\mathcal{A}_{T(C)}(\boldsymbol{p})$ is independent of the triangulation $T(C)$.*

1C. *Algebraic geometry background.* Next, we review some background in algebraic geometry, referring to [Eisenbud 1995; Hartshorne 1977; Schenck 2003] for more detail. Homogenizing the numerators of Wachspress coordinates yields our main object of study:

Definition 1.7. The Wachspress map defined by a polygon P_d is the rational map $\mathbb{P}^2 \overset{w_d}{\dashrightarrow} \mathbb{P}^{d-1}$ given on the open set $U_{z\neq 0} \subseteq \mathbb{P}^2$ by $(x, y) \mapsto (b_1(x, y), \dots, b_d(x, y))$. The Wachspress variety W_d is the closure of the image of w_d.

The polynomial ring $S = \mathbb{K}[x_1, \dots, x_d]$ is a graded ring: it has a direct-sum decomposition into homogeneous pieces. A finitely generated graded S-module N admits a similar decomposition; if $s \in S_p$ and $n \in N_q$, then $s \cdot n \in N_{p+q}$. In particular, each N_q is a $(S_0 = \mathbb{K})$-vector space.

Definition 1.8. For a finitely generated graded S-module N, the Hilbert series $\mathrm{HS}(N, t) = \sum \dim_{\mathbb{K}} N_q t^q$.

Definition 1.9. A free resolution for an S-module N is an exact sequence

$$\mathbb{F} : \cdots \to F_i \xrightarrow{d_i} F_{i-1} \to \cdots \to F_0 \to N \to 0,$$

where the F_i are free S-modules.

If N is graded, then the F_i are also graded, so letting $S(-m)$ denote a rank-1 free module generated in degree m, we may write $F_i = \bigoplus_j S(-j)^{a_{i,j}}$. By the Hilbert syzygy theorem [Eisenbud 1995], a finitely generated, graded S-module N has a free resolution of length at most d with all the F_i of finite rank.

Definition 1.10. For a finitely generated graded S-module N, a free resolution is minimal if, for each i, $\mathrm{Im}(d_i) \subseteq \mathfrak{m}F_{i-1}$, where $\mathfrak{m} = \langle x_1, \ldots, x_d \rangle$. The graded Betti numbers of N are the $a_{i,j}$ that appear in a minimal free resolution, and the Castelnuovo–Mumford regularity of N is $\max_{i,j}\{a_{i,j} - i\}$.

While the differentials d_i that appear in a minimal free resolution of N are not unique, the ranks and degrees of the free modules that appear are unique. The graded Betti numbers are displayed in a *Betti table*. Reading this table right and down, starting at (0, 0), the entry $b_{ij} := a_{i,i+j}$, and the regularity of N is the index of the bottommost nonzero row in the Betti table for N.

Example 1.11. In Examples 2.9 and 3.11 of [Garcia-Puente and Sottile 2010], it is shown that I_{W_6} is generated by three quadrics and one cubic. The variety $\mathbb{V}(\ell_1 \cdots \ell_6)$ of the edges of P_6 has $\binom{6}{2} = 15$ singular points, of which six are vertices of P_6, and S/I_{W_6} has Betti table

total	1	4	6	3
0	1	–	–	–
1	–	3	–	–
2	–	1	6	3

For example, $b_{1,2} = a_{1,3} = 1$ reflects that I_{W_6} has a cubic generator, and S/I_{W_6} has regularity 2. The Hilbert series can be read off the Betti table:

$$\mathrm{HS}(S/I_{W_6}, t) = \frac{1 - 3t^2 - t^3 + 6t^4 - 3t^5}{(1-t)^6} = \frac{1 + 3t + 3t^2}{(1-t)^3}.$$

Theorem 5.11 gives a complete description of the Betti table of S/I_{W_d}.

2. $H^0(D_{d-2})$ and the Wachspress surface

2A. ***Background on blowups of*** $\mathbb{P}^2$**.** Fix points $p_1, \ldots, p_k \in \mathbb{P}^2$, and let

$$X \xrightarrow{\pi} \mathbb{P}^2 \tag{1}$$

be the blowup of $\mathbb{P}^2$ at these points. Then $\mathrm{Pic}(X)$ is generated by the exceptional curves E_i over the points p_i and the proper transform E_0 of a line in $\mathbb{P}^2$. A classical geometric problem asks for a relationship between numerical properties of a divisor $D_m = mE_0 - \sum a_i E_i$ on X and the geometry of

$$X \xrightarrow{\phi} \mathbb{P}(H^0(D_m)^\vee).$$

First, we discuss some basics. Let m and a_i be nonnegative, let I_{p_i} denote the ideal of a point p_i, and define

$$J = \bigcap_{i=1}^{k} I_{p_i}^{a_i} \subseteq \mathbb{K}[x, y, z] = R. \tag{2}$$

Then $H^0(D_m)$ is isomorphic to the m-th graded piece J_m of J (see [Harbourne 2002]). Davis and Geramita [1988] show that, if $\gamma(J)$ denotes the smallest degree t such that J_t defines J scheme theoretically, then D_m is very ample if $m > \gamma(J)$, and if $m = \gamma(J)$, then D_m is very ample if and only if J does not contain m collinear points, counted with multiplicity. Note that $\gamma(J) \leq \operatorname{reg}(J)$.

2B. ***Wachspress surfaces.*** For a polygon P_d, fix defining linear forms ℓ_i as in Definition 1.3 and let $A := \ell_1 \cdots \ell_d$; the edges of P_d are defined by the $\mathbb{V}(\ell_i)$. Let Z denote the $\binom{d}{2}$ singular points of $\mathbb{V}(A)$ and $Y = Z \setminus \nu(P_d)$. Finally, X_d will be the blowup of $\mathbb{P}^2$ at Y. We study the divisor

$$D_{d-2} = (d-2)E_0 - \sum_{p \in Y} E_p$$

on X_d. First, we present some preliminaries.

Definition 2.1. Let L be the ideal in $R = \mathbb{K}[x, y, z]$ given by

$$L = \langle \ell_3 \cdots \ell_d, \ell_1\ell_4 \cdots \ell_d, \ldots, \ell_2 \cdots \ell_{d-1} \rangle = \langle A/\ell_1\ell_2, A/\ell_2\ell_3, \ldots, A/\ell_d\ell_1 \rangle,$$

where $A = \prod_{i=1}^{d} \ell_i$.

For any variety V, we use I_V to denote the ideal of polynomials vanishing on V.

Lemma 2.2. *The ideals L and I_Y are equal up to saturation at $\langle x, y, z \rangle$.*

Proof. Being equal up to saturation at $\langle x, y, z \rangle$ means that the localizations at any associated prime except $\langle x, y, z \rangle$ are equal. The ideal I_p of a point p is a prime ideal. Recall that the localization of a ring T at a prime ideal $\mathfrak{p}$ is a new ring $T_{\mathfrak{p}}$ whose elements are of the form f/g with $f, g \in T$ and $g \notin \mathfrak{p}$. Localize R at I_p, where $p \in Y$. Then in R_{I_p}, ℓ_i is a unit if $p \notin \mathbb{V}(\ell_i)$. Without loss of generality, suppose forms ℓ_1 and ℓ_2 vanish on p (note that all points of Y are intersections of exactly two lines) and the remaining forms do not. Thus, $L_{I_p} = \langle \ell_1, \ell_2 \rangle = (I_Y)_{I_p}$. □

The ideal L is not saturated.

Lemma 2.3. *I_Y is generated by one form F of degree $d-3$ and $d-3$ forms of degree $d-2$. Hence, a basis for L_{d-2} consists of $F \cdot x$, $F \cdot y$, $F \cdot z$, and the $d-3$ forms.*

Proof. First, note that I_Y cannot contain any form of degree $d-4$ since Y contains d sets of $d-3$ collinear points. So the smallest degree of a minimal generator for I_Y is $d-3$. Since Y consists of $\binom{d-1}{2}-1$ distinct points and the space of forms of degree $d-3$ has dimension $\binom{d-1}{2}$, there is at least one form F of degree $d-3$ in I_Y. We claim that it is unique. To see this, first note that no ℓ_i can divide F: by symmetry, if one ℓ_i divides F, they all must, which is impossible for degree reasons. Now suppose G is a second form of degree $d-3$ in I_Y. Let $p \in \nu(P_d)$ and $\mathbb{V}(\ell_i)$ be a line corresponding to an edge containing p. $F(p)$ must be nonzero since if not $\mathbb{V}(F)$ would contain $d-2$ collinear points of $\mathbb{V}(\ell_i)$, forcing $\mathbb{V}(F)$ to contain $\mathbb{V}(\ell_i)$, a contradiction. This also holds for G. But in this case, $F(p)G - G(p)F$ is a polynomial of degree $d-3$ vanishing at $d-2$ collinear points, again a contradiction. So F is unique (up to scaling), which shows that the Hilbert function satisfies

$$\mathrm{HF}(R/L, d-3) = |Y|,$$

so $\mathrm{HF}(R/L, t) = |Y|$ for all $t \geq d-3$ (see [Schenck 2003]). As the polynomials $A/\ell_i\ell_{i+1}$ are linearly independent and there are the correct number, L_{d-2} must be the degree-$(d-2)$ component of I_Y. $\square$

Theorem 2.4. *The minimal free resolution of R/L is*

$$0 \to R(-d) \xrightarrow{d_3} R(-d+1)^d \xrightarrow{d_2} R(-d+2)^d \xrightarrow{\left[\frac{A}{\ell_1\ell_2}\ \frac{A}{\ell_2\ell_3}\ \cdots\ \frac{A}{\ell_d\ell_1}\right]} R \to R/L \to 0,$$

where $d_2 = \begin{bmatrix} \ell_1 & 0 & \cdots & \cdots & 0 & 0 & m_1 \\ -\ell_3 & \ell_2 & 0 & \cdots & \vdots & \vdots & m_2 \\ 0 & -\ell_4 & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & 0 & \ddots & \ddots & \ell_{d-2} & 0 & \vdots \\ \vdots & \vdots & \ddots & \ddots & -\ell_d & \ell_{d-1} & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & -\ell_1 & m_d \end{bmatrix}$

and the m_i are linear forms.

Proof. By Lemma 2.3, the generators of I_Y are known. Since I_Y is saturated, the Hilbert–Burch theorem implies that the free resolution of R/I_Y has the form

$$0 \to R(-d+1)^{d-3} \to R(-d+3) \oplus R(-d+2)^{d-3} \to R \to R/I_Y \to 0.$$

Writing I_Y as $\langle f_1, \dots, f_{d-3}, F\rangle$ and L as $\langle f_1, \dots, f_{d-3}, xF, yF, zF\rangle$, the task is to understand the syzygies on L given the description above of the syzygies on I_Y. From the Hilbert–Burch resolution, any minimal syzygy on I_Y is of the form

$$\sum g_i f_i + qF = 0,$$

where g_i are linear and q is a quadric (or zero). Since

$$qF = g_1 xF + g_2 yF + g_3 zF \quad \text{with } g_i \text{ linear,}$$

all $d-3$ syzygies on I_Y lift to give linear syzygies on L. Furthermore, we obtain three linear syzygies on $\{xF, yF, zF\}$ from the three Koszul syzygies on $\{x, y, z\}$. It is clear from the construction that these d linear syzygies are linearly independent. Since $\mathrm{HF}(R/L, d-1) = |Y|$, this means we have determined all the linear first syzygies. Furthermore, the three Koszul first syzygies on $\{xF, yF, zF\}$ generate a linear second syzygy, so the complex given above is a subcomplex of the minimal free resolution. A check shows that the Buchsbaum–Eisenbud criterion [1973] holds, so the complex above is actually exact and hence a free resolution. The differential d_2 above involves the canonical generators $A/\ell_i\ell_{i+1}$ rather than a set involving $\{xF, yF, zF\}$. Since the $d-1$ linear syzygies appearing in the first $d-1$ columns of d_2 are linearly independent, they agree up to a change of basis; the last column of d_2 is a vector of linear forms determined by the change of basis. □

Theorem 2.5.

(i) $H^0(D_{d-2}) \simeq \mathrm{Span}_{\mathbb{K}}\{A/\ell_1\ell_2, A/\ell_2\ell_3, \dots\}$.

(ii) $H^1(D_{d-2}) = 0 = H^2(D_{d-2})$.

Proof. The remark following Equation (2) shows that $H^0(D_{d-2}) \simeq L_{d-2}$. Since $K = -3E_0 + \sum_{p\in Y} E_p$ (see [Hartshorne 1977]), by Serre duality,

$$H^2(D_{d-2}) \simeq H^0\biggl((-d-1)E_0 + \sum_{p\in Y} E_p\biggr),$$

which is clearly zero. Using that X_d is rational, it follows from Riemann–Roch that

$$h^0(D_{d-2}) - h^1(D_{d-2}) = \frac{D_{d-2}^2 - D_{d-2}\cdot K}{2} + 1.$$

The intersection pairing on X_d is given by $E_i^2 = 1$ if $i = 0$ and -1 if $i \neq 0$, and

$$E_i \cdot E_j = 0 \quad \text{if } i \neq j.$$

Thus,

$$D_{d-2}^2 = (d-2)^2 - |Y| \quad \text{and} \quad -D_{d-2}K = 3(d-2) - |Y|, \tag{3}$$

yielding

$$h^0(D_{d-2}) - h^1(D_{d-2}) = \frac{d^2 - d - 2 - 2|Y|}{2} + 1 = d. \tag{4}$$

Thus, $h^0(D_{d-2}) - h^1(D_{d-2}) = d$. Now apply the remark following Equation (2). □

Corollary 2.6. *If $d > 4$, D_{d-2} is very ample, so the image of X_d in $\mathbb{P}^{d-1}$ is smooth.*

Proof. By Theorem 2.4, the ideal L is $d-2$ regular. Furthermore, the set Y contains d sets of $d-3$ collinear points but no set of $d-2$ collinear points if $d > 4$. The result follows from the Davis–Geramita criterion. $\square$

Theorem 2.7. *$W_4 \simeq \mathbb{P}^1 \times \mathbb{P}^1$, and $X_4 \to W_4$ is an isomorphism away from the (-1) curve $E_0 - E_1 - E_2$, which is contracted to a smooth point.*

Proof. The surface X_4 is $\mathbb{P}^2$ blown up at two points, which is toric, and isomorphic to $\mathbb{P}^1 \times \mathbb{P}^1$ blown up at a point. By Proposition 6.12 of [Cox et al. 2011], D_2 is basepoint free. Since $D_2^2 = 2$, W_4 is an irreducible quadric surface in $\mathbb{P}^3$. As $D_2 \cdot (E_0 - E_1 - E_2) = 0$, the result follows. $\square$

Replacing D_{d-2} with $t\,D_{d-2}$, a computation as in Equations (3) and (4) and Serre vanishing shows that the Hilbert polynomial $\mathrm{HP}(S/I_{W_d}, t)$ is equal to

$$\frac{((d-2)^2-|Y|)t^2+(3(d-2)-|Y|)t}{2}+1=\frac{d^2-5d+8}{4}t^2-\frac{d^2-9d+12}{4}t+1. \quad (5)$$

3. The Wachspress quadrics

In this section, we construct a set of quadrics that vanish on W_d. These quadrics are polynomials that are expressed as a scalar product with a fixed vector τ. The vector τ defines a linear projection $\mathbb{P}^{d-1} \dashrightarrow \mathbb{P}^2$, also denoted by τ, given by

$$\boldsymbol{x} \mapsto \sum_{i=1}^{d} x_i \boldsymbol{v}_i,$$

where $\boldsymbol{x} = [x_1 : \cdots : x_d] \in \mathbb{P}^{d-1}$. By the second property of barycentric coordinates, the composition $\tau \circ w_d : \mathbb{P}^2 \dashrightarrow \mathbb{P}^2$ is the identity map on $\mathbb{P}^2$. Since $\boldsymbol{v}_i \in \mathbb{K}^3$, the vector τ is a triple of linear forms $(\tau_1, \tau_2, \tau_3) \in S^3$. The linear subspace $\mathscr{C}$ of $\mathbb{P}^{d-1}$ where the projection is undefined is the *center of projection*, and $I_{\mathscr{C}} = \langle \tau_1, \tau_2, \tau_3 \rangle$.

3A. *Diagonal monomials.* A *diagonal monomial* is a monomial $x_i x_j \in S_2$ such that $j \notin \{i-1, i, i+1\}$. We write $\mathfrak{D}$ for the subspace of S_2 spanned by the diagonal monomials; identifying x_i with the vertex v_i, a diagonal monomial is a diagonal in P_d; see Figure 2.

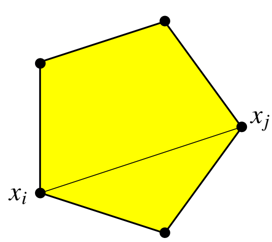


Figure 2. A diagonal monomial.

Lemma 3.1. *Any quadric that vanishes on W_d is a linear combination of elements of $\mathcal{D}$.*

Proof. Let Q be a polynomial in $(I_{W_d})_2$. Then $Q(w_d) = Q(b_1, \dots, b_d) = 0$. On the edge $[v_k, v_{k+1}]$, all the b_i vanish except b_k and b_{k+1}. Thus, on this edge, the expression $Q(w_d) = 0$ is

$$c_1 b_k^2 + c_2 b_k b_{k+1} + c_3 b_{k+1}^2 = 0 \tag{6}$$

for some constants c_1, c_2, and c_3 in $\mathbb{K}$. Recall that $b_i(v_j) = 0$ if $i \neq j$ and $b_i(v_i) \neq 0$ for each i. Evaluating (6) at v_k and v_{k+1}, we conclude $c_1 = c_3 = 0$. At an interior point of edge $[v_k, v_{k+1}]$, neither b_k nor b_{k+1} vanishes. This implies that $c_2 = 0$. A similar calculation on each edge shows that all coefficients of nondiagonal terms in Q are zero. □

3B. *The map to $(I_{\mathscr{C}})_2$.* We define a surjective map onto $(I_{\mathscr{C}})_2$ and use the map to calculate the dimension of the vector space of polynomials in $(I_{\mathscr{C}})_2$ that are supported on diagonal monomials. Let S_1^3 denote the space of triples of linear forms on $\mathbb{P}^{d-1}$. Define the map $\Psi : S_1^3 \to (I_{\mathscr{C}})_2$ by $F \mapsto F \cdot \tau$, where $\cdot$ is the scalar product.

Lemma 3.2. *The kernel of Ψ is three-dimensional.*

Proof. Since $I_{\mathscr{C}}$ is a complete intersection, the kernel is generated by the three Koszul syzygies on the τ_i. □

Next we determine conditions on F so that $\Psi(F) \in \mathcal{D}$. If $\boldsymbol{u}_i \in \mathbb{K}^3$ for $i = 1, \dots, d$, then

$$F = \sum_{i=1}^{d} x_i \boldsymbol{u}_i$$

is an element of S_1^3. Viewing the projection τ as an element of S_1^3, we have

$$\Psi(F) = F \cdot \tau = \left(\sum_{i=1}^{d} x_i \boldsymbol{u}_i \right) \cdot \left(\sum_{i=1}^{d} x_i \boldsymbol{v}_i \right) = \sum_{i,j=1}^{d} (\boldsymbol{u}_i \cdot \boldsymbol{v}_j + \boldsymbol{u}_j \cdot \boldsymbol{v}_i) x_i x_j. \tag{7}$$

If $\Psi(F) \in \mathcal{D}$, then the coefficients of nondiagonal monomials must vanish:

$$\boldsymbol{u}_i \cdot \boldsymbol{v}_i = 0 \quad \text{and} \quad \boldsymbol{u}_i \cdot \boldsymbol{v}_{i+1} + \boldsymbol{u}_{i+1} \cdot \boldsymbol{v}_i = 0 \quad \text{for all } i. \tag{8}$$

Lemma 3.3. *The dimension of the vector space $\mathcal{D} \cap (I_{\mathscr{C}})_2$ is $d - 3$.*

Proof. We show the conditions in (8) give $2d$ independent conditions on the $3d$-dimensional vector space S_1^3, and the solution space is $\Psi^{-1}(\mathcal{D} \cap (I_{\mathscr{C}})_2)$; thus,

$\dim(\Psi^{-1}(\mathscr{D} \cap (I_{\mathscr{C}})_2)) = d$. The conditions are represented by the matrix equation

$$\begin{pmatrix} \boldsymbol{v}_1 \cdot \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{v}_d \cdot \boldsymbol{u}_d \\ \boldsymbol{v}_1 \cdot \boldsymbol{u}_2 + \boldsymbol{v}_2 \cdot \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{v}_d \cdot \boldsymbol{u}_1 + \boldsymbol{v}_1 \cdot \boldsymbol{u}_d \end{pmatrix} = \overbrace{\begin{pmatrix} \boldsymbol{v}_1^T & 0 & \cdots & 0 \\ 0 & \boldsymbol{v}_2^T & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \boldsymbol{v}_d^T \\ \boldsymbol{v}_2^T & \boldsymbol{v}_1^T & & 0 \\ 0 & & \ddots & \\ \boldsymbol{v}_d^T & & & \boldsymbol{v}_1^T \end{pmatrix}}^{M} \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \boldsymbol{u}_d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix},$$

where the $\boldsymbol{v}_i$ and $\boldsymbol{u}_i$ are column vectors and the superscript T indicates transpose. The matrix M in the middle is a $2d \times 3d$ matrix, and the proof will be complete if the rows are shown to be independent. Denote the rows of M by $r_1, \ldots, r_d, r_{d+1}, \ldots, r_{2d}$, and let $c_1 r_1 + \cdots + c_d r_d + c_{d+1} r_{d+1} + \cdots + c_{2d} r_{2d}$ be a dependence relation among them. The first three columns of M give the dependence relation $c_1 \boldsymbol{v}_1 + c_{d+1} \boldsymbol{v}_2 + c_{2d} \boldsymbol{v}_d = 0$. Since $\boldsymbol{v}_d$, $\boldsymbol{v}_1$, and $\boldsymbol{v}_2$ define adjacent rays of a polyhedral cone, they must be independent, so c_1, c_{d+1}, and c_{2d} must be zero. Repeating the process at each triple $\boldsymbol{v}_{i-1}$, $\boldsymbol{v}_i$, and $\boldsymbol{v}_{i+1}$ shows the rest of the c_i's vanish. Since the restriction $\Psi : \Psi^{-1}(\mathscr{D} \cap (I_{\mathscr{C}})_2) \to \mathscr{D} \cap (I_{\mathscr{C}})_2$ remains surjective, we find $\dim(\mathscr{D} \cap (I_{\mathscr{C}})_2) = \dim(\Psi^{-1}(\mathscr{D} \cap (I_{\mathscr{C}})_2)) - \dim(\ker(\Psi)) = d - 3$. □

3C. *Wachspress quadrics.* We now compute the dimension and a generating set for $(I_{W_d})_2$.

Definition 3.4. Let $\gamma(i)$ denote the set $\{1, \ldots, d\} \setminus \{i-1, i\}$, $\gamma(i, j) = \gamma(i) \cap \gamma(j)$, and $\gamma(i, j, k) = \gamma(i) \cap \gamma(j) \cap \gamma(k)$.

The image of a diagonal monomial $x_i x_j$ under the pullback map $w_d^* : S \to R$ is

$$b_i b_j = \alpha_i \alpha_j \prod_{k \in \gamma(i)} \ell_k \prod_{m \in \gamma(j)} \ell_m = \alpha_i \alpha_j \prod_{k=1}^{d} \ell_k \prod_{m \in \gamma(i,j)} \ell_m,$$

and each diagonal monomial has a common factor $A = \prod_{k=1}^{d} \ell_k$. To find the quadratic relations among Wachspress coordinates, it suffices to find linear relations among products $\prod_{m \in \gamma(i,j)} \ell_m \in R_{d-4}$ for diagonal pairs i and j. Define the map $\phi : \mathscr{D} \to R_{d-4}$ by $x_i x_j \mapsto b_i b_j / A$, and extend by linearity; this is w_d^* restricted to $\mathscr{D}$ and divided by A. By Lemma 3.1, it follows that $(I_{W_d})_2 = \ker(\phi) \subseteq \mathscr{D}$.

Lemma 3.5. *The dimension of* $(I_{W_d})_2$ *is* $d - 3$.

Proof. We will show $\phi : \mathscr{D} \to R_{d-4}$ is surjective with $\dim(\ker \phi) = d - 3$. To see this, note that there are $d - 3$ diagonal monomials that have x_1 as a factor. We show

	$x_{2,4}$	$\cdots$	$x_{2,d}$	$x_{3,5}$	$\cdots$	$x_{3,d}$	$\cdots$	$x_{d-3,d-1}$	$x_{d-3,d}$	$x_{d-2,d}$
$p_{1,3}$	$*$									
$\vdots$		$\ddots$								
$p_{1,d-1}$			$*$							
$p_{2,4}$	$*$			$*$						
$\vdots$					$\ddots$					
$p_{2,d-1}$			$*$			$*$				
$\vdots$							$\ddots$			
$p_{(d-4)(d-2)}$								$*$		
$p_{(d-4)(d-1)}$								$*$	$*$	
$p_{(d-3)(d-1)}$								$*$	$*$	$*$

Table 1. Values of images of diagonal monomials at external vertices.

that the images of the remaining

$$d(d-3)/2-(d-3)=(d-3)(d-2)/2=\dim(R_{d-4})$$

diagonal monomials are independent. Let $p_{s,t}=\ell_s\cap\ell_t$ and $x_{p,q}=x_px_q$. In Table 1, a star, $*$, represents a nonzero number and a blank space is zero. The $(i,\, j)$ entry in the table represents the value of the image of the diagonal monomial in column j at the external vertex in row i. The external vertices not lying on ℓ_d are arranged down the rows with their indices in lexicographic order.

Since Table 1 is lower triangular, the images are independent. We have found $\dim(R_{d-4})$ independent images, and hence, ϕ is surjective. This is a map from a vector space of dimension $d(d-3)/2$ to one of dimension $(d-2)(d-3)/2$. The map is surjective, so the kernel has dimension $d(d-3)/2-(d-2)(d-3)/2=d-3$. □

There is a generating set for $(I_{W_d})_2$ where each generator is a scalar product with the vector τ. The other vectors in these scalar products are

$$\Lambda_k=\frac{x_{k+1}}{\alpha_{k+1}}\boldsymbol{n}_{k+1}-\frac{x_k}{\alpha_k}\boldsymbol{n}_{k-1}\in S_1^3.$$

Lemma 3.6. *The vectors $\{\Lambda_1,\dots,\Lambda_d\}$ form a basis for the space $\Psi^{-1}(\mathfrak{D}\cap(I_{\mathscr{C}})_2)$.*

Proof. Suppose that $\sum_{k=1}^{d}c_k\Lambda_k=0$ is a linear dependence relation among the Λ_k. The coefficient of a variable x_k is

$$\frac{1}{\alpha_k}(c_{k-1}\boldsymbol{n}_k-c_k\boldsymbol{n}_{k-1}).$$

By the dependence relation, this must be zero, which implies that $\boldsymbol{n}_{k-1}$ and $\boldsymbol{n}_k$ are scalar multiples. This is impossible since they are normal vectors of adjacent facets of a polyhedral cone. Hence, $c_{k-1} = c_k = 0$ for all k, which shows that the Λ_k are independent.

In the proof of Lemma 3.3, we showed that $\dim(\Psi^{-1}(\mathcal{D} \cap (I_{\mathscr{C}})_2)) = d$, and we have just shown $\dim(\langle \Lambda_k \mid k = 1, \ldots, d\rangle) = d$. To prove the result, it suffices to show $\langle \Lambda_k \mid k = 1, \ldots, d\rangle \subseteq \Psi^{-1}(\mathcal{D} \cap (I_{\mathscr{C}})_2)$. The conditions of (8) are required for $\Lambda_k \in S_1^3$ to lie in $\Psi^{-1}(\mathcal{D} \cap (I_{\mathscr{C}})_2)$. We show these conditions are satisfied for each Λ_k.

Let $\boldsymbol{u}_i = 0$ if $i \neq k, k+1$, $\boldsymbol{u}_k = -\boldsymbol{n}_{k-1}/\alpha_k$, and $\boldsymbol{u}_{k+1} = \boldsymbol{n}_{k+1}/\alpha_{k+1}$ for each fixed k. Then

$$\Lambda_k = \frac{x_{k+1}}{\alpha_{k+1}}\boldsymbol{n}_{k+1} - \frac{x_k}{\alpha_k}\boldsymbol{n}_{k-1} = \sum_{i=1}^{d} \boldsymbol{u}_i x_i.$$

Since $\boldsymbol{n}_{k-1} \cdot \boldsymbol{v}_k = 0$, $\boldsymbol{n}_{k+1} \cdot \boldsymbol{v}_{k+1} = 0$, and $\boldsymbol{u}_i = 0$ for $i \neq k, k+1$, we have that $\boldsymbol{u}_i \cdot \boldsymbol{v}_i = 0$ for each $i = 1, \ldots d$. The expression $\boldsymbol{u}_i \cdot \boldsymbol{v}_{i+1} + \boldsymbol{u}_{i+1} \cdot \boldsymbol{v}_i$ is zero for all $i \neq k-1, k, k+1$ simply because $\boldsymbol{u}_i = 0$ for $i \neq k, k+1$. We have

$$\begin{aligned}\boldsymbol{u}_k \cdot \boldsymbol{v}_{k+1} + \boldsymbol{u}_{k+1} \cdot \boldsymbol{v}_k &= -\frac{\boldsymbol{n}_{k-1}}{\alpha_k} \cdot \boldsymbol{v}_{k+1} + \frac{\boldsymbol{n}_{k+1}}{\alpha_{k+1}} \cdot \boldsymbol{v}_k \\ &= -\frac{\boldsymbol{v}_{k-1} \times \boldsymbol{v}_k \cdot \boldsymbol{v}_{k+1}}{\alpha_k} + \frac{\boldsymbol{v}_{k+1} \times \boldsymbol{v}_{k+2} \cdot \boldsymbol{v}_k}{\alpha_{k+1}} \\ &= -\frac{|\boldsymbol{v}_{k-1}\boldsymbol{v}_k\boldsymbol{v}_{k+1}|}{\alpha_k} + \frac{|\boldsymbol{v}_{k+1}\boldsymbol{v}_{k+2}\boldsymbol{v}_k|}{\alpha_{k+1}} = 0\end{aligned}$$

as $\alpha_j = |\boldsymbol{v}_{j-1}\boldsymbol{v}_j\boldsymbol{v}_{j+1}|$. It is easy to show that the expression $\boldsymbol{u}_i \cdot \boldsymbol{v}_{i+1} + \boldsymbol{u}_{i+1} \cdot \boldsymbol{v}_i$ is zero for $i = k \pm 1$. Thus, the $\boldsymbol{u}_i$ satisfy the conditions in (8), so $\Lambda_k \in \Psi^{-1}(\mathcal{D} \cap (I_{\mathscr{C}})_2)$. □

Theorem 3.7 (Wachspress quadrics). *The Wachspress quadrics $(I_{W_d})_2$ are those elements of S_2 that are diagonally supported and vanish on $\mathscr{C}$. The quadrics $Q_k = \Lambda_k \cdot \tau$ for $k = 1, \ldots, d$ span $(I_{W_d})_2$.*

Proof. Let $\boldsymbol{p}$ be the vector (x, y, z). By definition of Wachspress coordinates,

$$\tau(w_d(\boldsymbol{p})) = \sum_{i=1}^{d} b_i(\boldsymbol{p})\boldsymbol{v}_i = \boldsymbol{p} \sum_{i=1}^{d} b_i(\boldsymbol{p}).$$

We have

$$\begin{aligned}\Lambda_k(w_d(\boldsymbol{p})) &= \frac{b_{k+1}(\boldsymbol{p})}{\alpha_{k+1}}\boldsymbol{n}_{k+1} - \frac{b_k(\boldsymbol{p})}{\alpha_k}\boldsymbol{n}_{k-1} \\ &= \Bigg(\prod_{j\neq k,k+1} \ell_j\Bigg)\boldsymbol{n}_{k+1} - \Bigg(\prod_{j\neq k-1,k} \ell_j\Bigg)\boldsymbol{n}_{k-1} \\ &= \Bigg(\prod_{j\neq k-1,k,k+1} \ell_j\Bigg)(\ell_{k-1}\boldsymbol{n}_{k+1} - \ell_{k+1}\boldsymbol{n}_{k-1}) \\ &= H[\boldsymbol{n}_{k+1}(\boldsymbol{n}_{k-1} \cdot \boldsymbol{p}) - \boldsymbol{n}_{k-1}(\boldsymbol{n}_{k+1} \cdot \boldsymbol{p})],\end{aligned}$$

where $H = \prod_{j\neq k-1,k,k+1} \ell_j$. Set $\overline{H} := H\sum_{i=1}^{d} b_i(\boldsymbol{p})$. Then we have

$$\begin{aligned} Q_k(w_d(\boldsymbol{p})) &= \tau(w_d(\boldsymbol{p}))\cdot\Lambda_k(w_d(\boldsymbol{p})) \\ &= \overline{H}\boldsymbol{p}\cdot[\boldsymbol{n}_{k+1}(\boldsymbol{n}_{k-1}\cdot\boldsymbol{p}) - \boldsymbol{n}_{k-1}(\boldsymbol{n}_{k+1}\cdot\boldsymbol{p})] \\ &= \overline{H}[(\boldsymbol{n}_{k+1}\cdot\boldsymbol{p})(\boldsymbol{n}_{k-1}\cdot\boldsymbol{p}) - (\boldsymbol{n}_{k-1}\cdot\boldsymbol{p})(\boldsymbol{n}_{k+1}\cdot\boldsymbol{p})] = 0. \end{aligned}$$

We have just shown that $Q_k \in (I_{W_d})_2$. By Lemma 3.6, $\Psi^{-1}(\mathfrak{D}\cap(I_{\mathscr{C}})_2)$ is spanned by the Λ_k. Observe that $\langle Q_1,\dots,Q_d\rangle = \Psi(\langle\Lambda_k\rangle) = \mathfrak{D}\cap(I_{\mathscr{C}})_2$. Thus, $\dim(\langle Q_1,\dots,Q_d\rangle) = d-3$, and by Lemma 3.5, $\dim((I_{W_d})_2) = d-3$. Therefore, since $\langle Q_1,\dots,Q_d\rangle \subseteq (I_{W_d})_2$, we have $\langle Q_1,\dots,Q_d\rangle = (I_{W_d})_2 = \mathfrak{D}\cap(I_{\mathscr{C}})_2$. □

Corollary 3.8. *The quadrics $\{\Lambda_2\cdot\tau,\dots,\Lambda_{d-2}\cdot\tau\}$ are a basis for the quadrics in I_{W_d}, and in graded lex order, $\{x_1x_3,\dots,x_1x_{d-1}\}$ is a basis for $\mathrm{in}_\prec(I_{W_d})_2$.*

Proof. Expanding the expression for $\Lambda_i\cdot\tau$ yields

$$\Lambda_i\cdot\tau = x_1x_{i+1}\left(\frac{\boldsymbol{v}_1\cdot\boldsymbol{n}_{i+1}}{\alpha_{i+1}}\right) - x_1x_i\left(\frac{\boldsymbol{v}_1\cdot\boldsymbol{n}_{i-1}}{\alpha_i}\right) + \zeta_i,$$

where $\zeta_i \in \mathbb{K}[x_2,\dots,x_d]$. Since $\boldsymbol{n}_i = \boldsymbol{v}_i\times\boldsymbol{v}_{i+1}$,

$$\Lambda_2\cdot\tau = x_1x_3\left(\frac{\boldsymbol{v}_1\cdot\boldsymbol{n}_3}{\alpha_3}\right) + \zeta_2.$$

Since no three of the lines $\mathbb{V}(l_i)$ are concurrent, $\boldsymbol{v}_i\cdot\boldsymbol{n}_j$ is nonzero unless $j\in\{i,i+1\}$, so we may use the lead term of $\Lambda_2\cdot\tau$ to reduce $\Lambda_3\cdot\tau$ to $x_1x_4 + f(x_2,\dots,x_d)$. Repeating the process proves that

$$\{x_1x_3,\dots,x_1x_{d-1}\} \subseteq \mathrm{in}_\prec(I_{W_d})_2.$$

By Lemma 3.5, $(I_{W_d})_2$ has dimension $d-3$, which concludes the proof. □

Corollary 3.9. *There are no linear first syzygies on $(I_{W_d})_2$.*

Proof. By Corollary 3.8, we may assume that a basis for $(I_{W_d})_2$ has the form

$$\begin{gathered} x_1x_3 + \zeta_3(x_2,\dots,x_d), \\ x_1x_4 + \zeta_4(x_2,\dots,x_d), \\ x_1x_5 + \zeta_5(x_2,\dots,x_d), \\ \vdots \\ x_1x_{d-1} + \zeta_{d-1}(x_2,\dots,x_d). \end{gathered}$$

Since the ζ_i do not involve x_1, this implies that any linear first syzygy on $(I_{W_d})_2$ must be a linear combination of the Koszul syzygies on $\{x_3,\dots,x_{d-1}\}$. Now change the term order to graded lex with $x_i > x_{i+1} > \dots > x_d > x_1 > x_2 > \dots > x_{i-1}$. In this order, arguing as in the proof of Corollary 3.8 shows that we may assume a basis for $(I_{W_d})_2$ has the form

$$
\begin{gathered}
x_i x_{i+2} + \zeta_{i+2}(x_1, \dots, \widehat{x_i}, \dots, x_d), \\
x_i x_{i+3} + \zeta_{i+3}(x_1, \dots, \widehat{x_i}, \dots, x_d), \\
x_i x_{i+4} + \zeta_{i+4}(x_1, \dots, \widehat{x_i}, \dots, x_d), \\
\vdots \\
x_i x_{i-2} + \zeta_{i-2}(x_1, \dots, \widehat{x_i}, \dots, x_d).
\end{gathered}
$$

Hence, any linear first syzygy on $(I_{W_d})_2$ must be a combination of Koszul syzygies on $x_{i+2}, x_{i+3}, \dots, x_{i-2}$. Iterating this process for the term orders above shows there can be no linear first syzygies on $(I_{W_d})_2$. $\square$

3D. *Decomposition of* $\mathbb{V}(\langle (I_{W_d})_2 \rangle)$. We now prove that $\mathbb{V}(\langle (I_{W_d})_2 \rangle) = \mathscr{C} \cup W_d$. The results in Sections 4 and 5 are independent of this fact.

Lemma 3.10. *For any i, j, and k, we have*

$$|\boldsymbol{n}_i \boldsymbol{n}_j \boldsymbol{n}_k| = |\boldsymbol{v}_j \boldsymbol{v}_k \boldsymbol{v}_{k+1}| \cdot |\boldsymbol{v}_i \boldsymbol{v}_{i+1} \boldsymbol{v}_{j+1}| - |\boldsymbol{v}_{j+1} \boldsymbol{v}_k \boldsymbol{v}_{k+1}| \cdot |\boldsymbol{v}_i \boldsymbol{v}_{i+1} \boldsymbol{v}_j|.$$

Proof. Apply the formulas $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}(\boldsymbol{a} \cdot \boldsymbol{b})$ and $|\boldsymbol{a}\boldsymbol{b}\boldsymbol{c}| = \boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c}$:

$$
\begin{aligned}
|\boldsymbol{n}_i \boldsymbol{n}_j \boldsymbol{n}_k| = \boldsymbol{n}_i \times \boldsymbol{n}_j \cdot \boldsymbol{n}_k &= (\boldsymbol{n}_i \times (\boldsymbol{v}_j \times \boldsymbol{v}_{j+1})) \cdot \boldsymbol{n}_k \\
&= [\boldsymbol{v}_j (\boldsymbol{n}_i \cdot \boldsymbol{v}_{j+1}) - \boldsymbol{v}_{j+1} (\boldsymbol{n}_i \cdot \boldsymbol{v}_j)] \cdot \boldsymbol{n}_k \\
&= (\boldsymbol{v}_j \cdot \boldsymbol{n}_k)(\boldsymbol{n}_i \cdot \boldsymbol{v}_{j+1}) - (\boldsymbol{v}_{j+1} \cdot \boldsymbol{n}_k)(\boldsymbol{n}_i \cdot \boldsymbol{v}_j) \\
&= |\boldsymbol{v}_j \boldsymbol{v}_k \boldsymbol{v}_{k+1}| \cdot |\boldsymbol{v}_i \boldsymbol{v}_{i+1} \boldsymbol{v}_{j+1}| - |\boldsymbol{v}_{j+1} \boldsymbol{v}_k \boldsymbol{v}_{k+1}| \cdot |\boldsymbol{v}_i \boldsymbol{v}_{i+1} \boldsymbol{v}_j|. \qquad \square
\end{aligned}
$$

Corollary 3.11. *We have $|\boldsymbol{n}_i \boldsymbol{n}_j \boldsymbol{n}_{j+1}| = \alpha_{j+1} |\boldsymbol{v}_i \boldsymbol{v}_{i+1} \boldsymbol{v}_{j+1}|$.*

Proof. This follows from Lemma 3.10 and the definition of α_{j+1}. $\square$

Corollary 3.12. *We have $|\boldsymbol{n}_{i-1} \boldsymbol{n}_i \boldsymbol{n}_{i+1}| = \alpha_i \alpha_{i+1}$.*

Proof. This follows from Lemma 3.10 and the definition of α_i and α_{i+1}. $\square$

Lemma 3.13. *Let $\boldsymbol{x} = [x_1 : \cdots : x_d] \in \mathbb{V}(\langle (I_{W_d})_2 \rangle) \setminus \mathscr{C}$. If $\tau(\boldsymbol{x})$ is a base point $p_{ij} = \boldsymbol{n}_i \times \boldsymbol{n}_j$, then $\boldsymbol{x}$ lies on the exceptional line $\hat{p}_{ij}$ over p_{ij}.*

Proof. Since indices are cyclic, we assume that $i = 1$. Thus, $\tau(\boldsymbol{x}) = p_{1,j} = \boldsymbol{n}_1 \times \boldsymbol{n}_j$ for some $j \notin \{d, 1, 2\}$. The relation $Q_1(\boldsymbol{x}) = \Lambda_1 \cdot \tau(\boldsymbol{x}) = \Lambda_1 \cdot (\boldsymbol{n}_1 \times \boldsymbol{n}_j) = 0$ yields

$$L(1) := x_2 \boldsymbol{n}_2 \cdot p_{1,j} - x_1 \boldsymbol{n}_d \cdot p_{1,j} = 0. \tag{9}$$

The relation $Q_j(\boldsymbol{x}) = 0$ implies

$$L(j) := x_{j+1} |\boldsymbol{n}_{j+1} \boldsymbol{n}_1 \boldsymbol{n}_j| - x_j |\boldsymbol{n}_2 \boldsymbol{n}_1 \boldsymbol{n}_j| = 0. \tag{10}$$

Also,

$$Q_2(\boldsymbol{x}) = (x_3\boldsymbol{n}_3 - x_2\boldsymbol{n}_1)\cdot\boldsymbol{n}_1\times\boldsymbol{n}_j = x_3|\boldsymbol{n}_3\boldsymbol{n}_1\boldsymbol{n}_j| = 0,$$

implying $x_3 = 0$ since $|\boldsymbol{n}_3\boldsymbol{n}_1\boldsymbol{n}_j| \neq 0$ if $j \neq 3$. Assume $x_k = 0$ for $3 \le k < j-1$. Note that

$$Q_k(\boldsymbol{x}) = (x_{k+1}\boldsymbol{n}_{k+1} - x_k\boldsymbol{n}_{k-1})\cdot\boldsymbol{n}_1\times\boldsymbol{n}_j = x_{k+1}|\boldsymbol{n}_{k+1}\boldsymbol{n}_1\boldsymbol{n}_j| = 0;$$

hence, $x_{k+1} = 0$ since $|\boldsymbol{n}_{k+1}\boldsymbol{n}_1\boldsymbol{n}_j| \neq 0$ and by induction $x_k = 0$ for $3 \le k \le j-1$. An analogous argument shows that $x_k = 0$ for $j+2 \le k \le d$. Hence, $\boldsymbol{x}$ lies on the line $\mathbb{V}(L(1), L(j), x_k \mid k \notin \{1, 2, j, j+1\})$, which is the exceptional line $\hat{p}_{1,j}$. $\square$

Theorem 3.14. *The subset $\mathbb{V}(\langle (I_{W_d})_2\rangle)\setminus\mathscr{C}$ is contained in W_d. It follows that the variety $\mathbb{V}(\langle (I_{W_d})_2\rangle)$ has irreducible decomposition $W_d \cup \mathscr{C}$.*

Proof. Let $\boldsymbol{x} = [x_1 : \cdots : x_d] \in \mathbb{V}(\langle (I_{W_d})_2\rangle)\setminus\mathscr{C}$. The Wachspress quadrics give the relations

$$x_{r+1}\boldsymbol{n}_{r+1}\cdot\tau = x_r\boldsymbol{n}_{r-1}\cdot\tau \tag{11}$$

for each $r = 1, \ldots, d$. By Theorem 1.6, the adjoint is independent of triangulation,

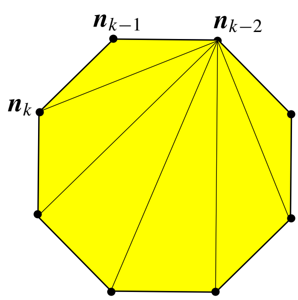


Figure 3. Triangulation used for adjoint.

so we use $\mathscr{A}$ to denote the adjoint, specifying the triangulation if necessary. We now show, for each $k \in \{1, \ldots, d\}$, $b_k(\tau(\boldsymbol{x})) = \mathscr{A}(\tau(\boldsymbol{x}))x_k$, where the triangulation above is used for the adjoint $\mathscr{A}$. It follows from the uniqueness of Wachspress coordinates that the denominator $\sum_{i=1}^{d} b_i$ of β_i is the adjoint of $\boldsymbol{P}_d^*$, so it follows that

$$w_d(\tau(\boldsymbol{x})) = \mathscr{A}(\tau(\boldsymbol{x}))\boldsymbol{x}. \tag{12}$$

Provided $\mathscr{A}(\tau(\boldsymbol{x})) \neq 0$, the result follows since $w_d(\tau(\boldsymbol{x})) \in \mathbb{P}^{d-1}$ is a nonzero scalar multiple of $\boldsymbol{x}$; hence, $\boldsymbol{x}$ is in the image of the Wachspress map and thus lies on W_d. If $\boldsymbol{x} \in \mathbb{V}(\langle (I_{W_d})_2\rangle)\setminus\mathscr{C}$ and $\mathscr{A}(\tau(\boldsymbol{x})) = 0$, then by (12) $w_d(\tau(\boldsymbol{x})) = 0$, and hence, $\tau(\boldsymbol{x})$ is a basepoint of w_d. Thus, $\tau(\boldsymbol{x}) = \boldsymbol{n}_i \times \boldsymbol{n}_j$ for some diagonal pair (i, j). By Lemma 3.13,

$\boldsymbol{x}$ lies on an exceptional line and hence lies on W_d. To prove the claim, note that since all indices are cyclic it suffices to assume $k = 3$. Let $|\boldsymbol{n}_i\boldsymbol{n}_j\boldsymbol{n}_k| = |\boldsymbol{n}_{ijk}|$ and

$$\boldsymbol{n}_{i_1,\dots,i_m} \cdot \tau := \prod_{j=1}^{m} (\boldsymbol{n}_{i_j} \cdot \tau).$$

This is the product of m linear forms in S, and with this notation,

$$b_3(\tau) = \boldsymbol{n}_{1,4,5,\dots,d} \cdot \tau.$$

For each $r \in \{3, \dots, d\}$, define

$$\sigma_r := (\boldsymbol{n}_{4,\dots,r} \cdot \tau)\boldsymbol{n}_1 \cdot \left[\sum_{i=3}^{r} \boldsymbol{v}_i (\boldsymbol{n}_{r+1,\dots,d} \cdot \tau) x_i + \sum_{i=r+1}^{d} \boldsymbol{v}_i (\boldsymbol{n}_{r-1,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau) x_r \right],$$

where we set $\boldsymbol{n}_{i,\dots,j} \cdot \tau = 1$ if $j < i$. We show $x_3\mathcal{A}(\tau(\boldsymbol{x})) = \sigma_3 = \sigma_d = b_3(\tau(\boldsymbol{x}))$. First, we show $\sigma_3 = x_3\mathcal{A}(\tau)$: to see this, note that

$$x_3\mathcal{A}(\tau) = |\boldsymbol{n}_{123}|(\boldsymbol{n}_{4,\dots,d} \cdot \tau)x_3 + \sum_{i=4}^{d} |\boldsymbol{n}_{1,i-1,i}|(\boldsymbol{n}_{2,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau)x_3, \quad (13)$$

where we express the adjoint $\mathcal{A}$ using the triangulation in Figure 3. Applying the scalar triple product to $|\boldsymbol{n}_{123}|$ and $|\boldsymbol{n}_{1,i-1,i}|$ in the expression (13) yields

$$\boldsymbol{n}_1 \cdot (\boldsymbol{n}_2 \times \boldsymbol{n}_3)(\boldsymbol{n}_{4,\dots,d} \cdot \tau)x_3 + \sum_{i=4}^{d} \boldsymbol{n}_1 \cdot (\boldsymbol{n}_{i-1} \times \boldsymbol{n}_i)(\boldsymbol{n}_{2,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau)x_3. \quad (14)$$

Factoring an $\boldsymbol{n}_1$ and noting that $\boldsymbol{n}_i \times \boldsymbol{n}_{i+1} = \boldsymbol{v}_{i+1}$, (14) becomes

$$\boldsymbol{n}_1 \cdot \left[\boldsymbol{v}_3(\boldsymbol{n}_{4,\dots,d} \cdot \tau)x_3 + \sum_{i=4}^{d} \boldsymbol{v}_i(\boldsymbol{n}_{2,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau)x_3 \right] = \sigma_3.$$

Now we show $\sigma_d = b_3(\tau)$. Since $\boldsymbol{n}_{d+1,\dots,d} \cdot \tau = 1$,

$$\sigma_d = (\boldsymbol{n}_{4,\dots,d} \cdot \tau)\boldsymbol{n}_1 \cdot \left(\sum_{i+3}^{d} \boldsymbol{v}_i (\boldsymbol{n}_{d+1,\dots,d} \cdot \tau)x_i \right) = (\boldsymbol{n}_{4,\dots,d} \cdot \tau)\boldsymbol{n}_1 \cdot \left(\sum_{i+3}^{d} \boldsymbol{v}_i x_i \right). \quad (15)$$

Observing that $\boldsymbol{n}_1 \cdot \sum_{i=1}^{2} x_i \boldsymbol{v}_i = 0$, we see that (15) is

$$(\boldsymbol{n}_{4,\dots,d} \cdot \tau)(\boldsymbol{n}_1 \cdot \tau) = \boldsymbol{n}_{1,4,\dots,d} \cdot \tau = b_3(\tau).$$

We now claim that for $r \in \{3, \dots, d-1\}$ we have $\sigma_r = \sigma_{r+1}$. Indeed,

$$\begin{aligned}\sigma_r &= (\boldsymbol{n}_{4,\dots,r} \cdot \tau)\boldsymbol{n}_1 \cdot \Bigg[\sum_{i=3}^{r} \boldsymbol{v}_i (\boldsymbol{n}_{r+1,\dots,d} \cdot \tau) x_i \\ &\qquad\qquad + \sum_{i=r+1}^{d} \boldsymbol{v}_i (\boldsymbol{n}_{r,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau)(\boldsymbol{n}_{r-1} \cdot \tau) x_r \Bigg] \\ &= (\boldsymbol{n}_{4,\dots,r} \cdot \tau)\boldsymbol{n}_1 \cdot \Bigg[\sum_{i=3}^{r} \boldsymbol{v}_i (\boldsymbol{n}_{r+1,\dots,d} \cdot \tau) x_i \\ &\qquad\qquad + \sum_{i=r+1}^{d} \boldsymbol{v}_i (\boldsymbol{n}_{r,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau)(\boldsymbol{n}_{r+1} \cdot \tau) x_{r+1} \Bigg],\end{aligned}$$

where we have applied (11) to the last term. Factoring out $\boldsymbol{n}_{r+1} \cdot \tau$ yields

$$(\boldsymbol{n}_{4,\dots,r+1} \cdot \tau)\boldsymbol{n}_1 \cdot \Bigg[\sum_{i=3}^{r} \boldsymbol{v}_i (\boldsymbol{n}_{r+2,\dots,d} \cdot \tau) x_i + \sum_{i=r+1}^{d} \boldsymbol{v}_i (\boldsymbol{n}_{r,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau) x_{r+1} \Bigg].$$

Lastly, since the expressions in both summations agree at the index $i = r+1$, we can shift the indices of summation,

$$(\boldsymbol{n}_{4,\dots,r+1} \cdot \tau)\boldsymbol{n}_1 \cdot \Bigg[\sum_{i=3}^{r+1} \boldsymbol{v}_i (\boldsymbol{n}_{r+2,\dots,d} \cdot \tau) x_i + \sum_{i=r+2}^{d} \boldsymbol{v}_i (\boldsymbol{n}_{r,\dots,i-2} \cdot \tau)(\boldsymbol{n}_{i+1,\dots,d} \cdot \tau) x_{r+1} \Bigg],$$

which is precisely σ_{r+1}, proving the claim. The claim shows that $\sigma_3 = \sigma_d$; hence, (12) holds, and so $\boldsymbol{x}$ lies in W_d if $\mathscr{A}(\tau(\boldsymbol{x})) \neq 0$. □

4. The Wachspress cubics

Theorem 3.14 shows that the Wachspress quadrics do not suffice to cut out the Wachspress variety W_d. We now construct cubics, the *Wachspress cubics*, that lie in I_{W_d} and do not arise from the Wachspress quadrics. These cubics are determinants of 3×3 matrices of linear forms. The key to showing that they are in I_{W_d} is to write them as a difference of adjoints $\mathscr{A}_{T_1(C)} - \mathscr{A}_{T_2(C)}$, where $T_1(C)$ and $T_2(C)$ are two different triangulations of a subcone C of the dual cone $\boldsymbol{P}_d^*$. By Theorem 1.6, the difference is zero, so the cubic is in I_{W_d}.

4A. *Construction of Wachspress cubics.*

As in Lemma 3.6, let

$$\Lambda_r = \frac{x_{r+1}}{\alpha_{r+1}} \boldsymbol{n}_{r+1} - \frac{x_r}{\alpha_r} \boldsymbol{n}_{r-1}.$$

Theorem 4.1. *If $i \neq j \neq k \neq i$, then $w_{i,j,k} := |\Lambda_i, \Lambda_j, \Lambda_k| \in I_{W_d}$.*

Proof. We break the proof into two parts. First, suppose no pair of (i, j, k) corresponds to an edge of P_d. We call such an (i, j, k) a T-triple. A direct calculation shows that, if (i, j, k) is a T-triple, then evaluating the monomial $x_i x_j x_k$ at Wachspress coordinates yields

$$x_i x_j x_k(w_d) = b_i b_j b_k = A^2 \prod_{m \in \gamma(i,j,k)} \ell_m, \tag{16}$$

where $\gamma(i, j, k)$ is as in Definition 3.4. Since there are no T-triples if $d < 6$, we may assume $d \geq 6$. Changing variables by replacing x_i with x_i/α_i, we may ignore the constants α_i. Using the definition of the Λ's, observe that

$$\begin{aligned} w_{i,j,k} = {} & |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k+1}| x_{i+1}x_{j+1}x_{k+1} - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}| x_{i+1}x_{j+1}x_k \\ & - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k+1}| x_{i+1}x_j x_{k+1} + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}| x_{i+1}x_j x_k \\ & - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k+1}| x_i x_{j+1}x_{k+1} + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}| x_i x_{j+1}x_k \\ & + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k+1}| x_i x_j x_{k+1} - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}| x_i x_j x_k. \end{aligned} \tag{17}$$

There are several situations to consider, depending on various possibilities for interactions among the indices. Interactions may occur if $i+1 = j-1$ or $j+1 = k-1$ or $k+1 = i-1$, so there are four cases:

1. All three hold. 2. Two hold. 3. One holds. 4. None hold.

Case 1. The indices (i, j, k) satisfy Case 1 if and only if $d = 6$. For $d = 6$, there are only two T-triples: $(1, 3, 5)$ and $(2, 4, 6)$. We show that $w_{1,3,5}$ vanishes on Wachspress coordinates; the case of $w_{2,4,6}$ is similar. All but two of the determinants in Equation (17) vanish, leaving

$$w_{1,3,5} = |\Lambda_1, \Lambda_3, \Lambda_5| = |\boldsymbol{n}_2\boldsymbol{n}_4\boldsymbol{n}_6| x_2x_4x_6 - |\boldsymbol{n}_6\boldsymbol{n}_2\boldsymbol{n}_4| x_1x_3x_5. \tag{18}$$

Notice that the coefficients are equal, and we conclude by showing that

$$x_1x_3x_5 - x_2x_4x_6$$

vanishes on Wachspress coordinates. The monomials $x_1x_3x_5$ and $x_2x_4x_6$ evaluated at Wachspress coordinates are $b_1b_3b_5$ and $b_2b_4b_6$, respectively. Both of these are equal to A^2, so $x_1x_3x_5 - x_2x_4x_6$ vanishes on Wachspress coordinates.

Case 2. We can assume without loss of generality $i + 1 \neq j - 1,\ j + 1 = k - 1$, and $k + 1 = i - 1$. Four coefficients vanish in (17), yielding

$$\begin{aligned} w_{i,j,k} = {} & |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{i-1}| x_{i+1}x_{j+1}x_{i-1} \\ & - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{i-1}| x_{i+1}x_j x_{i-1} \\ & + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{j+1}| x_{i+1}x_j x_{i-2} \\ & - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{j+1}| x_i x_j x_{i-2}. \end{aligned}$$

Evaluating this at Wachspress coordinates yields

$$\begin{aligned}
w_{i,j,k} \circ w_d &= |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{i-1}| \prod_{m\in\gamma(i+1,j+1,i-1)} \ell_m + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{i-1}| \prod_{m\in\gamma(i+1,j,i-1)} \ell_m \\
&\quad - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{j+1}| \prod_{m\in\gamma(i+1,j,i-1)} \ell_m - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{j+1}| \prod_{m\in\gamma(i,j,i-1)} \ell_m \\
&= A^2\Bigg(\prod_{m\in\gamma(i-1,i+1,j+1,j)} \ell_m\Bigg)\big(|\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{i-1}|\ell_{j-1} - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{i-1}|\ell_{j+1} \\
&\qquad + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{j+1}|\ell_{i-1} - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{j+1}|\ell_{i+1}\big) \\
&= A^2\Bigg(\prod_{m\in\gamma(i-1,i+1,j+1,j)} \ell_m\Bigg)\big[(|\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{i-1}|\ell_{j-1} + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{j-1}|\ell_{i+1}) \\
&\qquad - (|\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{i-1}|\ell_{j+1} + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{j-1}|\ell_{i-1})\big],
\end{aligned}$$

where

$$A = \prod_{i=1}^{d} \ell_i.$$

The last factor is the difference of two adjoints with respect to the triangulations of the quadrilateral in Figure 4. The vanishing can be seen directly: write $\boldsymbol{n}_1, \dots, \boldsymbol{n}_4$ for $\boldsymbol{n}_{i-1}$, $\boldsymbol{n}_{i+1}$, $\boldsymbol{n}_{j-1}$, and $\boldsymbol{n}_{j+1}$. Then the last factor is

$$|\boldsymbol{n}_2\boldsymbol{n}_3\boldsymbol{n}_4|\ell_1 - |\boldsymbol{n}_1\boldsymbol{n}_3\boldsymbol{n}_4|\ell_2 + |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_4|\ell_3 - |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_3|\ell_4.$$

Applying $\dfrac{d}{dx}$ to this shows the x coefficient is

$$|\boldsymbol{n}_2\boldsymbol{n}_3\boldsymbol{n}_4|\boldsymbol{n}_{11} - |\boldsymbol{n}_1\boldsymbol{n}_3\boldsymbol{n}_4|\boldsymbol{n}_{21} + |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_4|\boldsymbol{n}_{31} - |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_3|\boldsymbol{n}_{41}.$$

This is the determinant of the matrix of the $\boldsymbol{n}_i$ with a repeat row for the x coordinates $\boldsymbol{n}_{i1}$, so it vanishes. Reason similarly for the y and z coefficients.

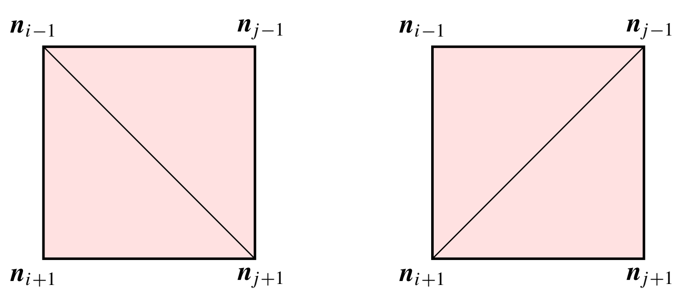


Figure 4. Case 2 triangulation.

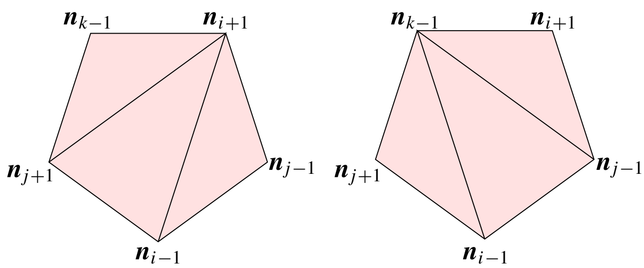


Figure 5. Case 3 triangulation.

Case 3. Assume without loss of generality $i+1 \neq j-1$, $j+1 \neq k-1$, and $k+1 = i-1$. In this case, two coefficients vanish in (17), and after evaluating at Wachspress coordinates, we obtain

$$\begin{aligned}
&w_{i,j,k} \circ w_d \\
&\quad = |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{i-1}| \prod_{m\in\gamma(i+1,j+1,k+1)} \ell_m - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i+1,j+1,k)} \ell_m \\
&\qquad - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{i-1}| \prod_{m\in\gamma(i+1,j,k+1)} \ell_m + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i+1,j,k)} \ell_m \\
&\qquad + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i,j+1,k)} \ell_m - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i,j,k)} \ell_m \\
&\quad = A^2 \Bigg(\prod_{\substack{m\in\gamma(i,j,k,\\ i+1,j+1,k+1)}} \ell_m \Bigg) \big(|\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{i-1}|\ell_{j-1}\ell_{k-1} - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}|\ell_{i-1}\ell_{j-1} \\
&\qquad - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{i-1}|\ell_{j+1}\ell_{k-1} + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}|\ell_{j+1}\ell_{i-1} \\
&\qquad + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}|\ell_{i+1}\ell_{j-1} - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}|\ell_{i+1}\ell_{j+1} \big).
\end{aligned}$$

The last factor is the difference of adjoints with respect to the triangulations of the pentagon in Figure 5.

Case 4. In this case, evaluation at Wachspress coordinates yields

$$\begin{aligned}
w_{i,j,k} \circ w_d &= |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k+1}| \prod_{m\in\gamma(i+1,j+1,k+1)} \ell_m - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i+1,j+1,k)} \ell_m \\
&\quad - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k+1}| \prod_{m\in\gamma(i+1,j,k+1)} \ell_m + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i+1,j,k)} \ell_m \\
&\quad - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k+1}| \prod_{m\in\gamma(i,j+1,k+1)} \ell_m + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i,j+1,k)} \ell_m \\
&\quad + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k+1}| \prod_{m\in\gamma(i,j,k+1)} \ell_m - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}| \prod_{m\in\gamma(i,j,k)} \ell_m
\end{aligned}$$

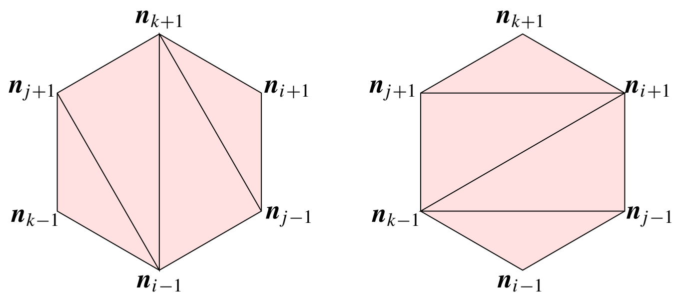


Figure 6. Case 4 triangulation.

$$
\begin{aligned}
&= A^2\Big(\prod_{\substack{m\in\gamma(i,j,k,\\ i+1,j+1,k+1)}} \ell_m\Big)\\
&\qquad \times \big(|\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k+1}|\ell_{i-1}\ell_{j-1}\ell_{k-1} - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}|\ell_{i-1}\ell_{j-1}\ell_{k+1}\\
&\qquad\quad - |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k+1}|\ell_{i-1}\ell_{j+1}\ell_{k-1} + |\boldsymbol{n}_{i+1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}|\ell_{i-1}\ell_{j+1}\ell_{k+1}\\
&\qquad\quad - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k+1}|\ell_{i+1}\ell_{j-1}\ell_{k-1} + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j+1}\boldsymbol{n}_{k-1}|\ell_{j+1}\ell_{i-1}\ell_{k+1}\\
&\qquad\quad + |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k+1}|\ell_{i+1}\ell_{j+1}\ell_{k-1} - |\boldsymbol{n}_{i-1}\boldsymbol{n}_{j-1}\boldsymbol{n}_{k-1}|\ell_{i+1}\ell_{j+1}\ell_{k+1}\big).
\end{aligned}
$$

The last factor is the difference of adjoints expressed using the triangulations of the hexagon in Figure 6. This completes the analysis when (i, j, k) is a T-triple.

Next, we consider the situation when (i, j, k) contains a pair of consecutive indices. Suppose first that there are exactly two consecutive vertices; without loss of generality, we assume the indices are $(2, 3, i)$ with $i > 4$. We have

$$
\begin{aligned}
w_{2,3,i} := |\Lambda_2\Lambda_3\Lambda_i| = |\boldsymbol{n}_2\boldsymbol{n}_4\boldsymbol{n}_{i+1}|x_3x_4x_{i+1} &- |\boldsymbol{n}_3\boldsymbol{n}_4\boldsymbol{n}_{i-1}|x_3x_4x_i\\
- |\boldsymbol{n}_3\boldsymbol{n}_2\boldsymbol{n}_{i+1}|x_3x_3x_{i+1} &+ |\boldsymbol{n}_3\boldsymbol{n}_2\boldsymbol{n}_{i-1}|x_3x_3x_i\\
- |\boldsymbol{n}_1\boldsymbol{n}_4\boldsymbol{n}_{i+1}|x_2x_4x_{i+1} &+ |\boldsymbol{n}_1\boldsymbol{n}_4\boldsymbol{n}_{i-1}|x_2x_4x_i\\
+ |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_{i+1}|x_2x_3x_{i+1} &- |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_{i-1}|x_2x_3x_i.
\end{aligned}
$$

We show that $w_{2,3,i} \circ w_d$ is a multiple of the difference between two expressions of the adjoint polynomial of a polygon with respect to two different triangulations. After evaluation at w_d, each monomial has a common factor of $A \prod_{j\neq 2,3} \ell_j$. Thus, we can express

$$
\frac{w_{2,3,i}(w_d)}{A \prod_{j\neq 2,3} \ell_j}
$$

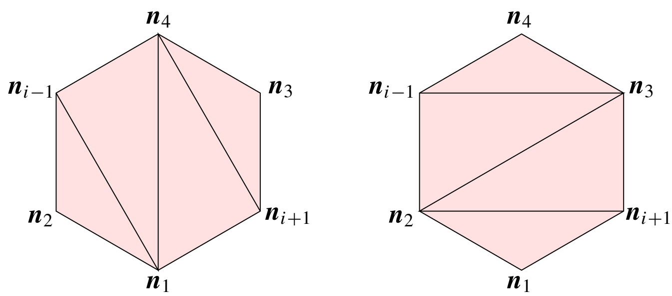


Figure 7. Triangulations for the non-T-triples.

as

$$\begin{aligned}
\frac{w_{2,3,i}(w_d)}{A\prod_{j\neq 2,3}\ell_j} = |\boldsymbol{n}_2\boldsymbol{n}_4\boldsymbol{n}_{i+1}| \prod_{j\neq 3,4,i+1} \ell_j - |\boldsymbol{n}_3\boldsymbol{n}_4\boldsymbol{n}_{i-1}| \prod_{j\neq 3,4,i-1} \ell_j \\
- |\boldsymbol{n}_3\boldsymbol{n}_2\boldsymbol{n}_{i+1}| \prod_{j\neq 2,3,i+1} \ell_j + |\boldsymbol{n}_3\boldsymbol{n}_2\boldsymbol{n}_{i-1}| \prod_{j\neq 2,3,i-1} \ell_j \\
- |\boldsymbol{n}_1\boldsymbol{n}_4\boldsymbol{n}_{i+1}| \prod_{j\neq 1,4,i+1} \ell_j + |\boldsymbol{n}_1\boldsymbol{n}_4\boldsymbol{n}_{i-1}| \prod_{j\neq 1,4,i-1} \ell_j \\
+ |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_{i+1}| \prod_{j\neq 1,2,i+1} \ell_j - |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_{i-1}| \prod_{j\neq 1,2,i-1} \ell_j \\
= \Bigg(\prod_{j\in\gamma(2,4,i,i+1)} \ell_j\Bigg)\big(|\boldsymbol{n}_2\boldsymbol{n}_4\boldsymbol{n}_{i+1}|\ell_1\ell_3\ell_{i-1} - |\boldsymbol{n}_3\boldsymbol{n}_4\boldsymbol{n}_{i-1}|\ell_1\ell_2\ell_{i+1} \\
- |\boldsymbol{n}_3\boldsymbol{n}_2\boldsymbol{n}_{i+1}|\ell_1\ell_4\ell_{i-1} + |\boldsymbol{n}_3\boldsymbol{n}_2\boldsymbol{n}_{i-1}|\ell_1\ell_4\ell_{i+1} \\
- |\boldsymbol{n}_1\boldsymbol{n}_4\boldsymbol{n}_{i+1}|\ell_2\ell_3\ell_{i-1} + |\boldsymbol{n}_1\boldsymbol{n}_4\boldsymbol{n}_{i-1}|\ell_2\ell_3\ell_{i+1} \\
+ |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_{i+1}|\ell_3\ell_4\ell_{i-1} - |\boldsymbol{n}_1\boldsymbol{n}_2\boldsymbol{n}_{i-1}|\ell_3\ell_4\ell_{i+1}\big).
\end{aligned}$$

The factor in parentheses is the difference of the adjoints computed with respect to the triangulations of the polygon in Figure 7.

Finally, for the case where the three vertices are consecutive, assume without loss of generality the triple is $(2, 3, 4)$, and proceed as above. In this case, the triangulations that arise are those that appear in Figure 5. □

Definition 4.2. $I(d)$ is the ideal generated by the Wachspress quadrics appearing in Corollary 3.8 and the Wachspress cubics appearing in Theorem 4.1.

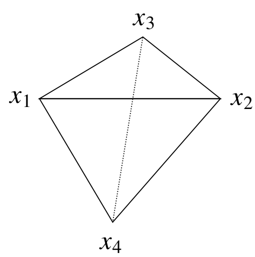
5. Gröbner basis, Stanley–Reisner ring, and free resolution

In this section, we determine the initial ideal of $I(d)$ in graded lex order and prove $I(d) = I_{W_d}$. First, we present some preliminaries.

5A. ***Simplicial complexes and combinatorial commutative algebra.*** An abstract n-simplex is a set consisting of all subsets of an $(n+1)$-element ground set. Typically a simplex is viewed as a geometric object; for example, a 2-simplex on the set $\{a, b, c\}$ can be visualized as a triangle with the subset $\{a, b, c\}$ corresponding to the whole triangle, $\{a, b\}$ an edge, and $\{a\}$ a vertex. For this reason, elements of the ground set are called the vertices.

Definition 5.1 [Ziegler 1995]. A simplicial complex Δ on a vertex set V is a collection of subsets σ of V such that, if $\sigma \in \Delta$ and $\tau \subset \sigma$, then $\tau \in \Delta$. If $|\sigma| = i + 1$, then σ is called an i-face. Let $f_i(\Delta)$ denote the number of i-faces of Δ, and define $\dim(\Delta) = \max\{i \mid f_i(\Delta) \neq 0\}$. If $\dim(\Delta) = n - 1$, we define $f_\Delta(t) = \sum_{i=0}^{n} f_{i-1}t^{n-i}$. The ordered list of coefficients of $f_\Delta(t)$ is the f-vector of Δ, and the coefficients of $h_\Delta(t) := f_\Delta(t - 1)$ are the h-vector of Δ.

Example 5.2. Consider the 1-skeleton of a tetrahedron with vertices x_1, x_2, x_3, x_4, as in the figure.



The corresponding simplicial complex Δ consists of all vertices and edges, so $\Delta = \{\varnothing, \{x_i\}, \{x_i, x_j\} \mid 1 \leq i \leq 4 \text{ and } i < j \leq 4\}$. Thus, $f(\Delta) = (1, 4, 6)$ and $h(\Delta) = (1, 2, 3)$; the empty face gives $f_{-1}(\Delta) = 1$.

A simplicial complex Δ can be used to define a commutative ring, known as the Stanley–Reisner ring. This construction allows us to use tools of commutative algebra to prove results about the topology or combinatorics of Δ.

Definition 5.3. Let Δ be a simplicial complex on vertices $\{x_1, \dots, x_n\}$. The Stanley–Reisner ideal I_Δ is

$$I_\Delta = \langle x_{i_1} \cdots x_{i_j} \mid \{x_{i_1}, \dots, x_{i_j}\} \text{ is not a face of } \Delta \rangle \subseteq \mathbb{K}[x_1, \dots, x_n],$$

and the Stanley–Reisner ring is $\mathbb{K}[x_1, \dots, x_k]/I_\Delta$.

In Example 5.2, since Δ has no 2-faces,

$$I_\Delta = \langle x_1x_2x_3, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4 \rangle = \bigcap_{1 \leq i < j \leq 4} \langle x_i, x_j \rangle.$$

Definition 5.4. A prime ideal P is associated to a graded S-module N if P is the annihilator of some $n \in N$, and $\operatorname{Ass}(N)$ is the set of all associated primes of N.

Definition 5.5. Let $\mathrm{codim}(N) = \min\{\mathrm{codim}(P) \mid P \in \mathrm{Ass}(N)\}$ for a finitely generated graded S-module N. The projective dimension $\mathrm{pdim}(N)$ is the length of a minimal free resolution of N; N is Cohen–Macaulay if $\mathrm{codim}(N) = \mathrm{pdim}(N)$. S/I is arithmetically Cohen–Macaulay if it is Cohen–Macaulay as an S-module.

5B. ***Application to Wachspress surfaces.***

Definition 5.6. Define $I_\Gamma(d) \subseteq \mathbb{K}[x_1, \dots, x_d]$ as

$$I_\Gamma(d) = \langle x_1x_3, \dots, x_1x_{d-1}\rangle + K_{2,d-1},$$

where $K_{2,d-1}$ consists of all square-free cubic monomials in $x_2, \dots, x_{d-1}$.

Theorem 5.7. *The quotient $S/I_\Gamma(d)$ is arithmetically Cohen–Macaulay, of Castelnuovo–Mumford regularity two, and has Hilbert series*

$$\mathrm{HS}(S/I_\Gamma(d), t) = \frac{1 + (d-3)t + \binom{d-3}{2}t^2}{(1-t)^3}.$$

Proof. The ideal $I_\Gamma(d)$ is the Stanley–Reisner ideal of a one-dimensional simplicial complex Γ consisting of a complete graph on vertices $\{x_2, \dots, x_{d-1}\}$ with a single additional edge $\overline{x_1x_2}$ attached. All connected graphs are shellable, so since shellable implies Cohen–Macaulay (see [Miller and Sturmfels 2005]), $S/I_\Gamma(d)$ is Cohen–Macaulay. Since $I_\Gamma(d)$ contains no terms involving x_d, if $S' = \mathbb{K}[x_1, \dots, x_{d-1}]$, then

$$S/I_\Gamma(d) \simeq S'/I_\Gamma(d) \otimes \mathbb{K}[x_d].$$

The Hilbert series of a Stanley–Reisner ring has numerator equal to the h-vector of the associated simplicial complex (see [Schenck 2003]), which in this case is a graph on $d-1$ vertices with $\binom{d-2}{2} + 1$ edges. Converting $f(\Gamma) = (1, d-1, \binom{d-2}{2} + 1)$ to $h(\Gamma)$ yields the Hilbert series of $S'/I_\Gamma(d)$. The Hilbert series of a graph has denominator $(1-t)^2$, and tensoring with $\mathbb{K}[x_d]$ contributes a factor of $1/(1-t)$, yielding the result. □

Theorem 5.8. *In graded lex order*, $\mathrm{in}_\prec I(d) = I_\Gamma(d)$.

Proof. First, note that

$$I_\Gamma(d) \subseteq \mathrm{in}_\prec I(d),$$

which follows from Corollary 3.8 and Theorem 4.1, combined with the observation that, in graded lex order, $\mathrm{in}(|\Lambda_i\Lambda_j\Lambda_k|) = x_ix_jx_k$ if $i < j < k$ as long as $k \neq d$. Since $I(d) \subseteq I_{W_d}$, there is a surjection

$$S/I(d) \twoheadrightarrow S/I_{W_d};$$

hence, $\mathrm{HP}(S/I(d), t) \geq \mathrm{HP}(S/I_{W_d}, t)$. Since

$$\mathrm{HP}(S/I(d), t) = \mathrm{HP}(S/\mathrm{in}_\prec I(d), t)$$

and

$$I_\Gamma(d) \subseteq \mathrm{in}_\prec I(d),$$

we have

$$\mathrm{HP}(S/I_\Gamma(d), t) \geq \mathrm{HP}(S/\mathrm{in}_\prec I(d), t) = \mathrm{HP}(S/I(d), t) \geq \mathrm{HP}(S/I_{W_d}, t).$$

The Hilbert polynomial $\mathrm{HP}(S/I_{W_d}, t)$ is given by Equation (5). The Hilbert series of $S/I_\Gamma(d)$ is given by Theorem 5.7, from which we can extract the Hilbert polynomial:

$$\mathrm{HP}(S/I_\Gamma(d), t) = \binom{d-3}{2}\binom{t}{2} + (d-3)\binom{t+1}{2} + \binom{t+2}{2},$$

and a check shows this agrees with Equation (5). Since $I_\Gamma(d) \subseteq \mathrm{in}_\prec I(d)$, equality of the Hilbert polynomials implies that in high degree (i.e., up to saturation)

$$I_\Gamma(d) = \mathrm{in}_\prec I(d) \quad \text{and} \quad I(d) = I_{W_d}.$$

Consider the short exact sequence

$$0 \to \mathrm{in}_\prec I(d)/I_\Gamma(d) \to S/I_\Gamma(d) \to S/\mathrm{in}_\prec I(d) \to 0.$$

By Lemma 3.6 of [Eisenbud 1995],

$$\mathrm{Ass}(\mathrm{in}_\prec I(d)/I_\Gamma(d)) \subseteq \mathrm{Ass}(S/I_\Gamma(d)). \tag{19}$$

Since $\mathrm{HP}(S/I_\Gamma(d), t) = \mathrm{HP}(S/\mathrm{in}_\prec I(d), t)$, the module $\mathrm{in}_\prec I(d)/I_\Gamma(d)$ must vanish in high degree so is supported at $\mathfrak{m}$, which is of codimension d. But $I_\Gamma(d)$ is a radical ideal supported in codimension $d-3$, so it follows from Equation (19) that $\mathrm{in}_\prec I(d)/I_\Gamma(d)$ must vanish. □

Corollary 5.9. *The ideal $I(d)$ is the ideal of the image of*

$$X_d \to \mathbb{P}(H^0(D_{d-2})).$$

In particular, $I(d) = I_{W_d}$, and $S/I(d)$ is arithmetically Cohen–Macaulay.

Proof. By the results of Sections 2 and 3, $I(d) \subseteq I_{W_d}$, and the proof of Theorem 5.8 showed that they are equal up to saturation. Hence, $I_{W_d}/I(d)$ is supported at $\mathfrak{m}$. Consider the short exact sequence

$$0 \to I_{W_d}/I(d) \to S/I(d) \to S/(I_{W_d}) \to 0.$$

Since $S/I_\Gamma(d) = S/\mathrm{in}_\prec I(d)$ is arithmetically Cohen–Macaulay of codimension $d-3$, by uppersemicontinuity [Herzog 2005], so is $S/I(d)$, so $I_{W_d}/I(d) = 0$. □

Corollary 5.10. *The quotient S/I_{W_d} has regularity* 2.

Proof. Since $S/I(d)$ is Cohen–Macaulay, reducing modulo a linear regular sequence of length 3 yields an Artinian ring with the same regularity, which is equal to the socle degree [Eisenbud 2005]. By Theorems 5.7 and 5.8, this is 2, so the regularity of S/I_{W_d} is 2. □

Theorem 5.11. *The nonzero graded Betti numbers of the minimal free resolution of $S/I(d)$ are given by $b_{12} = d-3$ and for $i \geq 3$ by*

$$b_{i-2,i} = \binom{d-3}{i} - (d-3)\binom{d-3}{i-1} + \binom{d-3}{2}\binom{d-3}{i-2}.$$

Proof. By Corollary 5.10, there are only two rows in the Betti table of $S/I(d)$. By Corollary 3.9, the top row is empty, save for the quadratic generators at the first step. Thus, the entire Betti diagram may be obtained from the Hilbert series, which is given in Theorem 5.7, and the result follows. □

We are at work on generalizing the results here to higher dimensions.

Acknowledgments

Computations were performed using Macaulay2, by Grayson and Stillman, available at http://www.math.uiuc.edu/Macaulay2/. A script is available at http://www.math.uiuc.edu/~schenck/Wpress.m2/. Our collaboration began during a SIAM conference on applied algebraic geometry, and we thank the organizers of that conference, Dan Bates and Frank Sottile. We thank Frank Sottile and Greg Smith for useful comments. Irving also thanks his thesis advisor, Frank Sottile, for introducing him to Wachspress varieties and providing guidance during the dissertation process.

References

[Buchsbaum and Eisenbud 1973] D. A. Buchsbaum and D. Eisenbud, "What makes a complex exact?", *J. Algebra* **25** (1973), 259–268. MR 47 #3369 Zbl 0264.13007

[Cox et al. 2011] D. A. Cox, J. B. Little, and H. K. Schenck, *Toric varieties*, Graduate Studies in Mathematics **124**, American Mathematical Society, Providence, RI, 2011. MR 2012g:14094 Zbl 1223.14001

[Davis and Geramita 1988] E. D. Davis and A. V. Geramita, "Birational morphisms to $\mathbf{P}^2$: an ideal-theoretic perspective", *Math. Ann.* **279**:3 (1988), 435–448. MR 89a:14014 Zbl 0657.14003

[Eisenbud 1995] D. Eisenbud, *Commutative algebra, with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, New York, 1995. MR 97a:13001 Zbl 0819.13001

[Eisenbud 2005] D. Eisenbud, *The geometry of syzygies: a second course in commutative algebra and algebraic geometry*, Graduate Texts in Mathematics **229**, Springer, New York, 2005. MR 2005h:13021 Zbl 1066.14001

[Garcia-Puente and Sottile 2010] L. D. Garcia-Puente and F. Sottile, "Linear precision for parametric patches", *Adv. Comput. Math.* **33**:2 (2010), 191–214. MR 2011f:65030 Zbl 1193.65018

[Harbourne 2002] B. Harbourne, "Problems and progress: a survey on fat points in $\mathbf{P}^2$", pp. 85–132 in *Zero-dimensional schemes and applications* (Naples, 2000), edited by A. V. Geramita, Queen's Papers in Pure and Appl. Math. **123**, Queen's Univ., Kingston, ON, 2002. MR 2003f:13032 Zbl 1052.14052

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Herzog 2005] J. Herzog, "Finite free resolutions", pp. 118–144 in *Computational commutative and non-commutative algebraic geometry*, edited by S. Cojocaru et al., NATO Sci. Ser. III Comput. Syst. Sci. **196**, IOS, Amsterdam, 2005. MR 2007a:13014 Zbl 1105.13012

[Miller and Sturmfels 2005] E. Miller and B. Sturmfels, *Combinatorial commutative algebra*, Graduate Texts in Mathematics **227**, Springer, New York, 2005. MR 2006d:13001 Zbl 1066.13001

[Möbius 1827] A. F. Möbius, *Der barycentrische Calcul: ein neues Hülfsmittel zur analytischen Behandlung der Geometrie*, Johann Ambrosius Barth, Leipzig, 1827.

[Schenck 2003] H. Schenck, *Computational algebraic geometry*, London Mathematical Society Student Texts **58**, Cambridge University Press, 2003. MR 2004k:13001 Zbl 1046.14034

[Wachspress 1975] E. L. Wachspress, *A rational finite element basis*, Mathematics in Science and Engineering **114**, Academic Press, New York, 1975. MR 54 #14403 Zbl 0322.65001

[Warren 1996] J. Warren, "Barycentric coordinates for convex polytopes", *Adv. Comput. Math.* **6**:2 (1996), 97–108. MR 97m:52026 Zbl 0873.52013

[Warren 2003] J. Warren, "On the uniqueness of barycentric coordinates", pp. 93–99 in *Topics in algebraic geometry and geometric modeling*, edited by R. Goldman and R. Krasauskas, Contemp. Math. **334**, Amer. Math. Soc., Providence, RI, 2003. MR 2005b:52022 Zbl 1043.52009

[Ziegler 1995] G. M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics **152**, Springer, New York, 1995. MR 96a:52011 Zbl 0823.52002

Communicated by Craig Huneke
Received 2013-01-06 Revised 2013-04-07 Accepted 2013-05-27

cfirving@scu.edu *Department of Mathematics and Computer Science, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053, United States*

schenck@math.uiuc.edu *Department of Mathematics, University of Illinois at Urbana–Champaign, 1409 West Green Street, Urbana, IL 61801, United States*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.397

Groups with exactly one irreducible character of degree divisible by p

Daniel Goldstein, Robert M. Guralnick, Mark L. Lewis, Alexander Moretó, Gabriel Navarro and Pham Huu Tiep

Let p be a prime. We characterize those finite groups which have precisely one irreducible character of degree divisible by p.

Minimal situations constitute a classical theme in group theory. Not only do they arise naturally, but they also provide valuable hints in searching for general patterns. In this paper, we are concerned with character degrees. One of the key results on character degrees is the Itô–Michler theorem, which asserts that a prime p does not divide the degree of any complex irreducible character of a finite group G if and only if G has a normal, abelian Sylow p-subgroup. In [Isaacs et al. 2009], Isaacs together with the fourth, fifth, and sixth authors of this paper studied the finite groups that have only one character degree divisible by p. They proved, among other things, that the Sylow p-subgroups of those groups were metabelian. This suggested that the derived length of the Sylow p-subgroups might be related with the number of different character degrees divisible by p. However, nothing could be said in [Isaacs et al. 2009] on how large p-Sylow normalizers were inside G. (As a trivial example, the dihedral group of order $2n$ for n odd has a unique character degree divisible by 2, and a self-normalizing Sylow 2-subgroup of order 2.)

In this paper, we go further and completely classify the finite groups with exactly one irreducible character of degree divisible by p. Our focus now therefore is not only on the set of character degrees but also on the multiplicity of the number of irreducible characters of each degree. In Section 1, we define the terms semi-extraspecial, ultraspecial, and doubly transitive Frobenius groups of Dickson type.

We would like to thank the referee for the helpful suggestions. Guralnick and Tiep gratefully acknowledge the support of the NSF (grants DMS-1001962, DMS-0901241, and DMS-1201374). Guralnick was also partially supported by the Simons Foundation fellowship 224965. The research of Moretó and Navarro is partially supported by the Spanish Ministerio de Educación y Ciencia proyecto MTM2010-15296, and Prometeo/Generalitat Valenciana.
MSC2010: 20C15.
Keywords: finite groups, characters, character degrees.

Main theorem. *Let p be a prime, and let G be a finite group. Then G has exactly one irreducible character of degree divisible by p if and only if one of the following statements holds*:

(i) $p = 2$ *and G is an extraspecial 2-group.*

(ii) $p = 2$ *and* $G = \mathsf{S}_4$.

(iii) $p = 3$ *and G is the semidirect product of* $\mathrm{SL}_2(3)$ *acting on its natural module* $(\mathbb{F}_3)^2$.

(iv) *G is a doubly transitive Frobenius group whose Frobenius complement has a nontrivial cyclic normal Sylow p-subgroup.*

(v) *p is odd and $G = H \ltimes K$, where $K = F(G)$ is an ultraspecial q-group for some prime $q \neq p$, H has a normal cyclic Sylow p-subgroup P, P acts trivially on K' and G/K' is a doubly transitive Frobenius group of Dickson type.*

(vi) *$G = HP$, where P is a normal semi-extraspecial Sylow p-subgroup and H is a group of order $|P'| - 1$ so that HP' is a doubly transitive Frobenius group.*

(vii) *Either G is* $\mathrm{PSL}_2(q)$ *or* $\mathrm{SL}_2(q)$ *or there exists a minimal normal elementary abelian p-subgroup V of order q^2 in G so that $G/V = \mathrm{SL}_2(q)$ and V can be viewed as a 2-dimensional irreducible module of G/V over* $\mathrm{End}_{G/V}(V) \cong \mathbb{F}_{q^2}$, *where $q = p^a \geq 4$ is a power of p.*

(viii) $p = 3$ *and* $G = \mathsf{S}_5$.

(ix) $p = 3$ *and* $G = M_{11}$.

Inspecting the groups listed in the main theorem, we see that either these groups have normal Sylow p-subgroups or their Sylow normalizers are maximal subgroups. Thus, as a corollary of the main theorem, we obtain:

Corollary. *Suppose that G is a finite group with exactly one irreducible character of degree divisible by p. Let $P \in \mathrm{Syl}_p(G)$. If P is not normal in G, then $\mathbf{N}_G(P)$ is maximal in G.*

This corollary suggests that, perhaps, the number of irreducible characters of G of degree divisible by p is bounded by the length of any saturated chain of subgroups between $\mathbf{N}_G(P)$ and G.

We now mention a connection of this problem with block theory and Brauer's height zero conjecture. We suppose that G is a group and that G has only one irreducible character whose degree is divisible by p. If G has more than one block, then there must exist at least one block where all the characters have height zero and this block will have maximal defect; in particular, take any of the blocks not containing the character whose degree is divisible by p. Since such a block has maximal defect, its defect group will be a Sylow p-subgroup of G, and then Brauer's height zero conjecture, if true, would imply that the Sylow p-subgroup must be abelian. Thus, in light of Brauer's height zero conjecture, we would expect that either our group would have abelian Sylow p-subgroups or that the principal block is the only p-block.

By the Gluck–Wolf theorem (Theorem 12.10 of [Manz and Wolf 1993]), we know that Brauer's height zero conjecture is true for p-solvable groups. Thus, it is possible that the Gluck–Wolf theorem might give a different approach to proving our result for p-solvable groups (but probably not shorter). In particular, if G does not have a single p-block, then we know that G must have abelian Sylow p-subgroups. If G does have a single p-block, then $\mathbf{O}_{p'}(G) = 1$, and so $\mathbf{O}_p(G) > 1$. In particular, if $\mathbf{O}_p(G)$ is not abelian, then all the nonlinear irreducible characters in $\mathrm{Irr}(\mathbf{O}_p(G))$ must be conjugate in G. On the other hand, by Theorem A of [Isaacs et al. 2009], we know that a p-solvable group G having only one irreducible character degree divisible by p and $\mathbf{O}_p(G)$ nonabelian must have that $\mathbf{O}_p(G)$ is a Sylow p-subgroup. Thus, our theorem could be viewed as classifying those groups G having a normal Sylow p-subgroup P where all the nonlinear irreducible characters of P are G-conjugate.

This suggests that it might be worth studying the following problem which looks to us to be difficult: classify the pairs (G, N) with N normal in G such that all the characters of N with degree divisible by p are conjugate in G. A closely related problem would be to classify the groups G where all the irreducible characters of G with degree divisible by p are Galois conjugate. While we hesitate to predict what such a classification would look like, we observe that an extraspecial p-group for any prime p will be an example. We expect that the p-solvable examples that are not nilpotent will involve Frobenius groups and a careful analysis of the conjugacy of elements of order p as in our examples. For the non-p-solvable groups, one would begin by looking at the simple groups having the property that all irreducible characters whose degrees are divisible by p are conjugate (under outer automorphisms or under Galois automorphisms). Using the classification of finite simple groups, one can show that the only nonabelian simple groups with this property are $\mathrm{PSL}_2(q)$, J_1, and M_{11} (see Corollary 7.5 of [Isaacs et al. 2009]). In fact, [Isaacs et al. 2009] studied a somewhat more general condition. The stronger condition that all nonlinear irreducible characters of the same degree are Galois conjugate was treated in the recent paper [Dolfi et al. 2013].

Our paper is structured as follows: in Section 1, we classify the p-solvable groups with exactly one irreducible character of degree divisible by p. The classification is first proved under the additional hypothesis that the group is solvable. We then produce examples of solvable groups that satisfy conclusion (vi) of the classification. The classification is then proved under the hypothesis that the group is p-solvable, but not solvable. This case is split depending on whether the Sylow p-subgroup is abelian or nonabelian. We also find some additional constraints when the group is not solvable, but it is p-solvable and has a nonabelian Sylow p-subgroup. We construct p-solvable groups that are not solvable and satisfy conclusion (vi) of the classification. Finally, in Section 2 we take care of the classification of the non-p-solvable groups.

1. p-solvable groups

In this section we state the classification for p-solvable groups. In Section 2 of [Dolfi et al. 2009], the reader can find definitions of the relevant terms below. Recall that P is a semi-extraspecial p-group if P/N is an extraspecial p-group for every subgroup N in $\mathbf{Z}(P)$ with $|\mathbf{Z}(P) : N| = p$. It is known that if P is a semi-extraspecial p-group, then $P' = \mathbf{Z}(P) = \Phi(P)$ and $\mathbf{Z}(P)$ is elementary abelian. Also, $|P'|^2 \leq |P : P'|$. For details, see [Fernández-Alcober and Moretó 2001], for instance. A group P is ultraspecial if P is a semi-extraspecial p-group that satisfies $|P'|^2 = |P : P'|$. In this section, we prove the following result, which is the p-solvable portion of the main theorem.

Theorem 1.1. *Fix a prime number p, and let G be a finite p-solvable group. Then G has exactly one irreducible character of degree divisible by p if and only if one of the following holds*:

(i) $p = 2$ *and G is an extraspecial* 2*-group.*

(ii) $p = 2$ *and* $G = \mathsf{S}_4$.

(iii) $p = 3$ *and G is the semidirect product of* $\mathrm{SL}_2(3)$ *acting on its natural module.*

(iv) *G is a doubly transitive Frobenius group whose Frobenius complement has a nontrivial cyclic normal Sylow p-subgroup.*

(v) *p is odd and $G = H \ltimes K$ where $K = F(G)$ is an ultraspecial q-group for some prime $q \neq p$, H has a normal cyclic Sylow p-subgroup P, P acts trivially on K' and G/K' is a doubly transitive Frobenius group of Dickson type.*

(vi) $G = HP$, *where P is a normal semi-extraspecial Sylow p-subgroup and H is a group of order $|P'| - 1$ so that HP' is a doubly transitive Frobenius group.*

Clearly, the groups in conclusions (i), (ii), and (iii) exist and are solvable. Doubly transitive Frobenius groups have been studied in a number of different places. As mentioned in [Dolfi et al. 2009], the doubly transitive Frobenius groups are in bijection with the finite near-rings. Most near-rings are obtained by Galois twists of finite fields and these are said to be of Dickson type. There are also seven near-rings that are said to be of exceptional type. The doubly transitive Frobenius groups of Dickson type are solvable. Four of the doubly transitive Frobenius groups of exceptional type are solvable and the other three are nonsolvable. As mentioned in [Dolfi et al. 2009], one of the solvable and two of the nonsolvable doubly transitive groups of exceptional type have a nontrivial cyclic normal Sylow p-subgroup. Thus, there are two nonsolvable groups that satisfy conclusion (iv).

In conclusion (v), since G/K' is doubly transitive of Dickson type and K is a q-group for some prime, we see that G must be solvable. We will see that the groups in (v) have exactly one conjugacy class whose size is divisible by p. In Section 5 of [Dolfi et al. 2009], they construct groups that satisfy conclusion (v), and they give conditions on when such groups can be constructed.

Groups that satisfy conclusion (vi) can be found on page 383 of [Gagola 1983]. Using this construction, one can find an example with $|P'| = p^a$ for every prime p and integer $a \geq 1$. We will use a variation on this construction to find groups that satisfy conclusion (vi) where HP' is any doubly transitive Frobenius group of Dickson type. We will also present a variation on this construction to produce an example where P is not ultraspecial.

We will also show that we can find groups that satisfy conclusion (vi) where HP' is any of the exceptional doubly transitive Frobenius groups. We will also find additional restrictions on the groups arising in this case. (See Theorem 1.20, where we essentially classify such groups. It is perhaps remarkable how large the exponent of $|P/P'|$ needs to be in any of these groups.)

We claim that it is easy to see that if G is one of the groups in (i)–(vi), then it has exactly one irreducible character of degree divisible by p. Thus, we will work to prove that if G is p-solvable and has exactly one irreducible character of degree divisible by p, then G is one of the groups in (i)–(vi).

Preliminaries. In this section, we present several results from other sources.

Lemma 1.2 [Lewis 2001, Lemma 1]. *Suppose a solvable group G acts faithfully on a group V, and let p be a prime divisor of $|G|$. Assume for each nonidentity element $v \in V$ that $\mathbf{C}_G(v)$ contains a unique Sylow p-subgroup of G. Then G is a subgroup of the semilinear group on V or $p = 3$, $|V| = 9$ and G is one of the groups* $\mathrm{SL}_2(3)$ *or* $\mathrm{GL}_2(3)$.

This next result was proved by Noritzsch. We write $\mathrm{cd}(G) = \{\chi(1) \mid \chi \in \mathrm{Irr}(G)\}$ for the set of irreducible character degrees of G.

Lemma 1.3 [Noritzsch 1995, Lemma 1.10]. *Let $V < N < G$ be normal subgroups of a finite group G such that G/N and N/V are cyclic of order a and b, respectively. Moreover, let V be elementary abelian and suppose that both G/V and N are Frobenius groups with kernel N/V and V, respectively. Then*

$$\mathrm{cd}(G) \cup \{ab\} = \{1, a\} \cup \{ib \mid i \text{ divides } a\}.$$

For nilpotent groups, the result is immediate. (This is essentially proved in [Seitz 1968], but we have decided to include our own short proof.)

Lemma 1.4. *Let G be a nilpotent group, and let p be a prime. Then G has exactly one irreducible character of degree divisible by p if and only if G is an extraspecial 2-group.*

Proof. Obviously G must be a nonabelian p-group. Let $Z = \mathbf{Z}(G) \cap G'$, and note that $Z > 1$. If α and β are distinct nontrivial characters of Z, then any irreducible constituent of α^G or β^G has degree a nontrivial power of p and clearly α^G and β^G have no common constituents. Thus, Z has only one nontrivial character, whence $|Z| = 2$ and so G is a 2-group. We see that G/Z has no irreducible characters of degree bigger than 1, and hence, $Z = G'$. Thus, G has $|G|/2$ distinct linear characters and so the nonlinear irreducible character has degree $[G : Z]^{1/2}$, whence $Z = \mathbf{Z}(G)$. Thus, G is an extraspecial 2-group. □

Groups with exactly one conjugacy class with size divisible by a prime p have been classified in [Dolfi et al. 2009]. Next is the classification.

Lemma 1.5 [Dolfi et al. 2009, Theorem A]. *Let G be a finite group and p a prime. Then G has exactly one conjugacy class of size divisible by p if and only if G is one of the following groups*:

(i) *G is a Frobenius group with Frobenius complement of order* 2 *and Frobenius kernel of order divisible by p.*

(ii) *G is a doubly transitive Frobenius group whose Frobenius complement has a nontrivial central Sylow p-subgroup.*

(iii) *p is odd, $G = KH$, where $K = F(G)$ is an ultraspecial q-group, q prime, $H = \mathbf{C}_G(P)$ for a Sylow p-subgroup P of G, $K \cap H = \mathbf{Z}(K)$ and $G/\mathbf{Z}(K)$ is a doubly transitive Frobenius group of Dickson type.*

Given an element $g \in G$, we write $\mathrm{cl}_G(g)$ for the conjugacy class of g in G.

Lemma 1.6. *Let p be a prime number. Assume that G has a normal p-complement and that G is not nilpotent. If G has exactly one irreducible character of degree divisible by p, then it has exactly one conjugacy class of size divisible by p.*

Proof. Write $G = AN$, where A is a Sylow p-subgroup of G and N is a normal p-complement. Assume first that A is not abelian. By Lemma 1.4, A is an extraspecial 2-group. Furthermore, by hypothesis every irreducible character of N is A-invariant. We deduce that $N = 1$. This contradiction implies that A is abelian.

By hypothesis, there exists a unique A-orbit of irreducible characters of N that are not A-invariant. Let $\theta \in \mathrm{Irr}(N)$ be a character that is not A-invariant. Since θ extends to its inertia subgroup in G and our hypothesis implies that there exists a unique irreducible character of G lying over θ, we have that $G_\theta = N$. In other words, the action of A on $\mathrm{Irr}(N)$ has exactly one regular orbit, and all other orbits have size one. Since the actions on characters and classes are permutation isomorphic, the same happens for the action of A on the set of conjugacy classes of N. We deduce that there exists a unique G-conjugacy class contained in N of size a multiple of p.

Now, consider an element $g \in G \setminus N$. We must prove that p does not divide $|\mathrm{cl}_G(g)|$. Since N is a normal Hall p'-subgroup, p divides $o(g)$. If $o(g)$ is a power of p, then g belongs to some Sylow p-subgroup of G. Since the Sylow p-subgroups of G are abelian, it follows that $|\mathrm{cl}_G(g)|$ is a p'-number, as desired. Hence, we may assume that the order of g is not a prime power, and we can write $g = g_p g_{p'}$ as the product of its p-part and its p'-part, where $g_p \neq 1$ and $g_{p'} \neq 1$. Observe that $g_{p'}$ belongs to N and commutes with the nontrivial p-element g_p. By the previous paragraph, $g_{p'}$ commutes with a Sylow p-subgroup T that contains g_p. It follows that g commutes with T, and hence, $|\mathrm{cl}_G(g)|$ is a p'-number, as desired. □

Looking at the conclusion of Lemma 1.5, it is not difficult to see that only the groups in conclusions (ii) and (iii) have exactly one irreducible character whose degree is divisible by p, and that these groups satisfy conclusions (iv) and (v) of Theorem 1.1. Thus, combining Lemma 1.6 with Lemma 1.5 yields the following corollary.

Corollary 1.7. *Let p be a prime number. Assume G has a normal p-complement and G is not nilpotent. If G has exactly one irreducible character of degree divisible by p, then G is one of the groups in (iv) or (v) of Theorem 1.1.*

Let K be a normal subgroup of G. We use $\mathrm{Irr}(G \mid K)$ to denote the characters in $\mathrm{Irr}(G)$ that do not contain K in their kernels.

Lemma 1.8. *Let p be a prime number. Assume that G has a normal Sylow p-subgroup but that G is not a p-group. If G has exactly one irreducible character of degree divisible by p, then G is one of the groups in (vi) of Theorem 1.1.*

Proof. Take P to be the Sylow p-subgroup of G, and let H be a Hall p-complement for G. Observe that p divides the degree of every character in $\mathrm{Irr}(G \mid P')$. This implies that $\mathrm{Irr}(G \mid P')$ contains a unique character χ. It is not difficult to see that χ vanishes on $G \setminus P'$. Also, G has a unique orbit on the nonprincipal characters of $\mathrm{Irr}(P')$. By Theorem 6.32 of [Isaacs 1976], this implies that P' contains only one nonidentity conjugacy class of G. This implies χ is one of the characters studied by Gagola [1983]. From Lemma 2.1 of [Gagola 1983], we see that $P'H$ is a doubly transitive Frobenius group. We deduce that P' is a minimal normal in G, and P' is central in P. In the language of [Chillag and Macdonald 1984], (G, P') is a Camina pair, and using Lemma 4.2 from that paper and the fact that P' is central in P, one sees that P is semi-extraspecial. □

The solvable case. We now prove Theorem 1.1 under the addition hypothesis that G is solvable.

Theorem 1.9. *Let p be a prime, and let G be a group having exactly one irreducible character whose degree is divisible by p. If G is solvable, then G satisfies conclusions (i)–(vi) of Theorem 1.1.*

Proof. Assume that G has exactly one irreducible character of degree divisible by p. We want to prove that G is one of the groups in (i)–(vi). Let G be a minimal counterexample.

Step 1: If $\mathbf{O}_p(G) > 1$, then $\mathbf{O}_p(G)$ is a Sylow subgroup of G.

Write $V = \mathbf{O}_p(G)$, and assume that V is not a Sylow p-subgroup of G. Now, G/V has a nontrivial, nonnormal Sylow p-subgroup, so p must divide the degree of some character in $\mathrm{Irr}(G/V)$ by Itô's theorem (Theorem 12.33 of [Isaacs 1976]). Hence, G/V will have exactly one irreducible character of degree divisible by p, and thus, G/V is not a counterexample. Since G/V does not have a nontrivial normal p-subgroup, G/V is one of the groups in (iv) or (v) of Theorem 1.1.

Substep 1a: V is a minimal normal subgroup of G.

Let K be a normal subgroup of G such that V/K is a chief factor of G. Assume that $K > 1$. Since G/K is not a counterexample, $G/K = \mathsf{S}_4$ and $p = 2$ or G/K is the group in (vi) and $p = 3$. In both cases, a Sylow p-subgroup is not abelian, so we may apply Theorem 12.9 of [Manz and Wolf 1993] to see that $\mathrm{Irr}(G \mid K)$ must contain a character of degree divisible by p. Since $\mathrm{Irr}(G/K)$ already contains a character of degree divisible by p, this is a contradiction. Therefore, we must have $K = 1$, and V is a minimal normal subgroup of G.

Substep 1b: $V = F(G)$.

Assume that $V < F = F(G)$. Let $E/V = F(G/V)$. Suppose first that $F = E$. Since G/V satisfies conclusion (iv) or (v), we know that $F = V \times W$, where W is a q-group for some prime $q \neq p$, and G/VW' is a doubly transitive Frobenius group. This implies that W/W' is an elementary abelian q-group and $(1_V \times \lambda)^G \in \mathrm{Irr}(G)$ for any character $1_W \neq \lambda \in \mathrm{Irr}(W/W')$. It is not difficult to show that $(\alpha \times \lambda)^G \in \mathrm{Irr}(G)$ for every character $\alpha \in \mathrm{Irr}(V)$. Since $1_V \times \lambda$ will be in a different G-orbit than $\alpha \times \lambda$ when $\alpha \neq 1_V$, we deduce that G has more than one irreducible character of degree divisible by p, a contradiction. Thus, we may assume that $F < E$.

Recall that G/V satisfies either conclusion (iv) or conclusion (v). In both cases, we know that $G/E'V$ is a doubly transitive Frobenius group, so $E/E'V$ is a chief factor for G. If G/V satisfies conclusion (iv), then E/V is abelian, so E/V is a chief factor, and we cannot have $F < V < E$. Thus, the claim is proved in this case, and we may assume that G/V satisfies conclusion (v).

We know that E/V is a semi-extraspecial group, so every normal subgroup of E/V either contains $E'V/V$ or is contained in $E'V/V$ (by Corollary 8.3 of [Fernández-Alcober and Moretó 2001]). Since $E/E'V$ is a chief factor and $F < E$, this implies that $F \leq E'V$. As $V < F$, we still have $F = V \times W$ where $W > 1$ is a normal q-group in G. Also, because $G/E'V$ is a doubly transitive Frobenius group, we know that $G/E'V$ does not have a normal Sylow p-subgroup. Hence, G/W does not have a normal Sylow p-subgroup. By Itô's theorem, this implies that p divides the degree of some character in $\mathrm{Irr}(G/W)$. Now, G/W is not a counterexample to the theorem and it has a nontrivial normal p-subgroup, and a Sylow p-subgroup that is not normal. It follows that G/W satisfies either conclusion (ii) or conclusion (vi), and both of these have a nonabelian Sylow p-subgroup. Thus, we may apply Theorem 12.9 of [Manz and Wolf 1993] to see that $\mathrm{Irr}(G \mid W)$ has a character with degree divisible by p, a contradiction. This completes the proof of the claim that $V = F(G)$.

Substep 1c: Proof of step 1.

Now, G/V acts faithfully on V, and $\mathbf{C}_G(v)$ contains a unique Sylow p-subgroup of G for any nonidentity $v \in V$. We can apply Lemma 1.2, and deduce that either $p = 3$, $G/V = \mathrm{SL}_2(3)$ and $|V| = 9$ or G/V is a subgroup of the semilinear group on V. In the first case, we obtain that G is the group of type (vi). This is a contradiction. In the second case, we have that G/V is a metacyclic doubly transitive Frobenius group with kernel, say, K/V and that the action of K/V on V is Frobenius. By Lemma 1.3, we deduce that $|G/K| = p$. But since $|K/V| = p+1$ is prime, we deduce that $p = 2$ and $G/V = \mathrm{S}_3$. We claim that $|V| = 4$. For every nonprincipal irreducible character λ of V, the inertia subgroup of λ in G is a Sylow 2-subgroup of G. Since G has 3 Sylow 2-subgroups $\{P_1, P_2, P_3\}$, we deduce that

$$\mathrm{Irr}(F(G)) \setminus \{1_{F(G)}\} = \bigcup_{i=1}^{3} C_{\mathrm{Irr}(F(G))}(P_i),$$

and this union is disjoint. In particular, $|V| - 1 = 3 \cdot 2^a$ for some integer a. Since V is a 2-group and $|V| = 3 \cdot 2^a + 1$, we deduce that $a = 0$ and $|V| = 4$. This implies that $G = \mathrm{S}_4$, again a contradiction. This means that if $V > 1$, then V is a Sylow p-subgroup, proving step 1.

Let H be a Hall p-complement of G.

Step 2: $\mathbf{O}_p(G) = 1$.

Suppose $\mathbf{O}_p(G) > 1$. By step 1, G has a normal Sylow p-subgroup. By Lemma 1.8, G satisfies conclusion (vi) and this contradicts the choice of G as a counterexample.

Step 3: G has a normal p-complement.

Assume that G does not have a normal p-complement. By Theorem A of [Isaacs et al. 2009], G has a cyclic Sylow p-subgroup. Write $X = \mathbf{O}_{p'}(G)$ and $Y/X = \mathbf{O}_p(G/X)$. The group Y/X is isomorphic to a Sylow p-subgroup of G and G/Y is isomorphic to a p'-subgroup of $\mathrm{Aut}(Y/X)$. Hence, G/Y is a cyclic group whose order is a divisor of $p-1$. Write $Y = PX$, where P is a Sylow p-subgroup of G. Since we are assuming that $G/Y > Y/Y$, it follows that $p > 2$. Set $V = [P, X]$, and let V/W be a chief factor of G. Let r be the prime so that V/W is an elementary abelian r-group.

Substep 3a: If $W > 1$, then W is nilpotent, $W = V'$, and all nonlinear irreducible characters of V are P-invariant.

By the minimality of G as a counterexample to the theorem, G/W is one of the groups described in (i)–(vi). We know that $p > 2$, so G/W is one of the groups in (iii)–(vi). The Sylow p-subgroups of G/W are cyclic, so it is not one of the groups of type (iii) or (vi) either. The Sylow p-subgroups of G/W do not act trivially on the minimal normal subgroup V/W, so G/W cannot be of type (v). It follows that G/W is a doubly transitive Frobenius group whose complement has a cyclic normal Sylow p-subgroup. In particular, $K = F(G) \leq V$. Since G has a unique irreducible character of degree divisible by p, it is easy to see that $V' = W$. Also, since p will not divide the degrees of any of the characters in $\mathrm{Irr}(G \mid V')$, all the nonlinear characters of V are fixed by P. By Theorem A of [Isaacs 1989], W is nilpotent.

Substep 3b: If $W > 1$, then V is an r-group.

Assume that this is not true. Then there exists a normal subgroup J of G so that W/J is a chief factor of G which is a t-group, for some prime $t \neq r$. For any character $\tau \in \mathrm{Irr}(W/J)$, we see that the stabilizer G_τ contains a full Sylow p-subgroup of G. Also, τ extends to PV_τ. Since the action of P on V/W is Frobenius, it follows from our hypothesis that $V_\tau = W$. Hence, the action of V/W on W/J is Frobenius, so V/W is cyclic of order r and G/V is cyclic. It follows that G has a normal p-complement, a contradiction. This proves the claim that W is an r-group. As a consequence, V is an r-group.

Substep 3c: $W = 1$.

Suppose $W > 1$. Fix a character $\theta \in \mathrm{Irr}(V \mid W) = \mathrm{Irr}(V \mid V')$, so θ is nonlinear. We have seen that θ must be P-invariant. So there exists a character $\mu \in \mathrm{Irr}(W)$ that is a constituent of θ_W and is P-invariant. By Problem 13.10 of [Isaacs 1976], there exists a unique irreducible constituent of μ^V that is P-invariant. But all the members of $\mathrm{Irr}(V \mid W)$ are P-invariant. It follows that $\mu^V = e\theta$ for some integer e. In particular, θ vanishes on the set $V \setminus W$. By Lemma 3.1 of [Lewis et al. 2005], we see that V is a semi-extraspecial r-group. One can easily see that P acts trivially on W, so G is one of the groups in (v). This contradiction implies that $W = 1$.

Let $C = \mathbf{C}_G(V)$.

Substep 3d: G/C acts transitively on $V \setminus \{1\}$.

Observe that PV/V is a Sylow p-subgroup of G/V. Since $\mathbf{O}_p(G) = 1$, we know that P acts faithfully on V. Also, $V = \mathbf{C}_V(P) \times [V, P]$ by Fitting's lemma. We have that $\mathbf{C}_V(P)$ is normal in G (since PV is normal in G), and V is a minimal normal subgroup of G. It follows that $\mathbf{C}_V(P) = 1$, so the action of P on V is Frobenius (by Theorem A of [Isaacs et al. 2009]). It follows that the degree of any of the members of $\mathrm{Irr}(G \mid V)$ is a multiple of p. By hypothesis, $|\mathrm{Irr}(G \mid V)| = 1$. This implies that the action of G/C on $V \setminus \{1\}$ is transitive.

Now, we can apply Theorem 6.8 of [Manz and Wolf 1993] to determine the structure of G/C.

Substep 3e: $C \le X$.

On the other hand, consider any character $\theta \in \mathrm{Irr}(X \mid V)$. Since P acts Frobeniusly on V, we know that the stabilizer of any nonprincipal irreducible character of V will be contained in X, and so, all the characters in $\mathrm{Irr}(X \mid V)$ induce irreducibly to Y. In particular, $\theta^Y \in \mathrm{Irr}(Y)$. But G/Y is cyclic, and $(\theta^Y)^G$ has a unique irreducible constituent. Clifford theory implies that $\theta^G \in \mathrm{Irr}(G)$. In other words, any member of $\mathrm{Irr}(X \mid V)$ induces irreducibly to the same character of G. Also, the action of P on V is Frobenius, so C is a p'-group. This implies that $C \le X$.

Substep 3f: G/C must be one the "exceptional groups" in the conclusion of Theorem 6.8 of [Manz and Wolf 1993].

Suppose first that G/C is isomorphic to a subgroup of the semilinear group of V. This implies that G/C is metacyclic. Write $F/C = F(G/C)$, and note that G/F is cyclic. Put $U = F \cap X \trianglelefteq G$. Arguing as in the previous paragraph, any member of $\mathrm{Irr}(U \mid V)$ induces irreducibly to the same character of G. This fact and Clifford's correspondence imply for every character $\lambda \in \mathrm{Irr}(V) \setminus \{1_V\}$ that $|\mathrm{Irr}(U_\lambda \mid \lambda)| = 1$. Hence, λ is fully ramified with respect to U_λ/V. We deduce that U_λ is a q-group for some prime q. Since the Sylow q-subgroups of $\Gamma(V)$ are cyclic, it follows that U_λ/C is cyclic and is contained in the cyclic group $\mathbf{O}_q(G/C)$. In particular, $U_\lambda \trianglelefteq G$. Since G is transitive on $V \setminus \{1\}$, we deduce that $U_\lambda = C$. Because C/V is fully ramified, we obtain $Z(C) \le V$. But V is minimal normal in G, so $Z(C) = V$.

We claim that $C = V$. Assume that this is not true. Thus, $C' \cap V > 1$, so $V \le C'$. But we know that P acts trivially on C/C'. Hence, it acts trivially on C and this is a contradiction. This implies that $C = V$, and the action of G/V on V is Frobenius.

This proves that G is a double transitive Frobenius group whose complement has a cyclic normal Sylow p-subgroup; that is, G is of type (iv). This is a contradiction. By Theorem 6.8 of [Manz and Wolf 1993], G/C is one of the exceptional groups in that theorem.

Substep 3g: Proof of step 3.

The group G/C cannot be one of the groups in conclusion (a) of Theorem 6.8 of [Manz and Wolf 1993] because $F(G/C)$ is not a nonabelian group of prime power order. Also, it cannot be one of the groups in conclusion (b) because $F(G/C)$ contains a noncentral cyclic Sylow p-subgroup. Therefore, G has a normal p-complement.

Step 4: Final contradiction.

By Corollary 1.7, G is one of the groups of type (iv) or (v). This is the final contradiction. □

Solvable examples. We now construct solvable groups that appear in conclusion (vi) of Theorem 1.1. In particular, for every two-transitive Frobenius group of Dickson type D, we find a group G satisfying conclusion (vi) of Theorem 1.1 so that $\mathbf{Z}(P)H \cong D$. We will also present an example that satisfies conclusion (vi) of Theorem 1.1 where P is not ultraspecial.

We start with examples that appeared in [Gagola 1983] and in [Isaacs 2011]. Let p be a prime and let n be a positive integer. Write F for the finite field of order p^n. Take K to be the matrix group

$$\left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & d \end{pmatrix} : a, b, c \in F; d \in F^* \right\}.$$

Set

$$P = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix} : a, b, c \in F \right\}.$$

It is shown in [Gagola 1983] that P is a normal Sylow p-subgroup of K. It is well known that P is an ultraspecial group of order p^{3n}. Set

$$L = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{pmatrix} : d \in F^* \right\}.$$

Observe that L is a Hall p-complement of K. Also, L is cyclic. It is not hard to see that L acts Frobeniusly on

$$\mathbf{Z}(P) = \left\{ \begin{pmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : b \in F \right\},$$

but that L does not act Frobeniusly on P. We should note that in fact K satisfies conclusion (vi) of Theorem 1.1. However, we need to produce more complicated examples. To do this, let $\mathscr{G}$ be the Galois group for F over Z_p. We can define an action $\mathscr{G}$ on K as follows: if $\sigma \in \mathscr{G}$, then σ acts on a typical element of K by acting on each of the entries of K. Notice that P and L are invariant under the action of $\mathcal{G}$. Also, note that the semidirect product of $\mathcal{G}$ acting on L is isomorphic to the affine group on F. (See page 38 of [Manz and Wolf 1993].) We take Γ to be the semidirect product of $\mathcal{G}$ acting on K. (We note that $\mathbf{Z}(P)L\mathcal{G}$ is isomorphic to the affine semilinear group on F which is also defined on page 38 of [Manz and Wolf 1993].)

Suppose $D = NH^*$ is a two-transitive Frobenius group of Dickson type of order $p^n(p^n-1)$, where N is the Frobenius kernel and H^* is the Frobenius complement. It is well known that H^* can be embedded in the affine group of F and that NH^* is isomorphic to a subgroup of the semilinear affine group of F. Thus, H^* is isomorphic to $H \subseteq L\mathcal{G}$ and NH is isomorphic to $\mathbf{Z}(P)H$. We set $G = PH$, and it is not difficult to see that G is the desired group.

To find an example of a group satisfying conclusion (vi) of Theorem 1.1 where P is not ultraspecial, we take Γ as above, but specialize $p = 2$ and $n = 3$. This implies that $|P| = 2^9$ and $|\mathcal{G}| = 3$. We define $G^* = P\mathcal{G}$. Observe that $\mathcal{G}$ centralizes a subgroup Z of $Z(P)$ having order 2. (The fixed field under the Galois group has order 2.) It follows that Z is in the center of G^*. Let $G = G^*/Z$. Since P/Z is a nonabelian quotient of a semi-extraspecial group, it is semi-extraspecial. Since $|P : \mathbf{Z}(P)| = 2^6$ and $|\mathbf{Z}(P) : Z| = 2^2$, it is not ultraspecial. Observe that $\mathcal{G}$ acts Frobeniusly on $\mathbf{Z}(P)/Z$, so $\mathbf{Z}(P)\mathcal{G}/Z$ is a doubly transitive Frobenius group. It follows from Lemma 2.2 of [Chillag and Macdonald 1984] that G satisfies conclusion (vi) of Theorem 1.1, and this yields our example with the normal Sylow subgroup not being ultraspecial.

We will construct examples where $D \cong \mathbf{Z}(P)H$ when D is a two-transitive solvable Frobenius of exceptional group later. The technique for constructing groups that satisfy conclusion (vi) of Theorem 1.1 when $\mathbf{Z}(P)H$ is exceptional is the same for solvable and nonsolvable groups. Thus, we hold off on that construction until after we handle the proof of Theorem 1.1 in the case where the group is p-solvable, but not solvable.

Abelian Sylow p-subgroup. We now prove Theorem 1.1 in the case that G is p-solvable but not solvable and the Sylow p-subgroup of G is abelian.

The following result is the only one in this section that uses the classification of finite simple groups.

Lemma 1.10. *Let p be prime. Let S be a finite nonabelian almost-simple group not divisible by p and A a nontrivial p-group of automorphisms of S. Then A is cyclic. Moreover, A has at least two nontrivial orbits on the irreducible characters of S.*

Proof. It follows by the classification of finite simple groups and their automorphism groups that $F^*(S)$ is a finite group of Lie type and A is a cyclic group of field automorphisms. By Brauer's lemma, it suffices to prove the same statement for conjugacy classes. We can appeal to [Dolfi et al. 2009, Lemma 3.1] to complete the proof. □

We apply the previous lemma to obtain the following.

Lemma 1.11. *Let G be a finite group with a minimal normal subgroup N that is a direct product of t copies of a nonabelian finite simple group. Assume that $\mathbf{C}_G(N) = 1$ and that either $t > 1$ or that there is a prime p dividing $|G|$ that does not divide $|N|$. Then G has at least 2 nontrivial orbits on conjugacy classes of N of size a multiple of p and on irreducible characters of N of degree divisible by p.*

Proof. If p divides $|N|$ and $t > 1$, the result is clear (indeed, there will be at least t such orbits).

If $t = 1$ and p does not divide $|N|$, then Lemma 1.10 applies. □

We now prove Theorem 1.1 under the hypothesis that G is p-solvable but not solvable and that G has an abelian Sylow p-subgroup.

Theorem 1.12. *Let p be a prime, and let G be a group having exactly one irreducible character whose degree is divisible by p. Suppose G is p-solvable, but not solvable. If G has an abelian Sylow p-subgroup, then G is a doubly transitive Frobenius group whose Frobenius kernel R is an elementary abelian r-group of order r^2 where either $(p, r) = (7, 29)$ or $(29, 59)$, and a Frobenius complement $H \cong \mathrm{SL}_2(5) \times \mathbb{Z}/p$. In particular, G satisfies conclusion (iv) of Theorem 1.1.*

Proof. Let P be a Sylow p-subgroup of G. By the Itô–Michler theorem, P is not normal. Since P is not normal in G, we have that P does not centralize $F^*(G)$.

First assume that R is a nilpotent normal p'-subgroup of G with $[P, R] \neq 1$. So P does not commute with $R/\Phi(R)$. If γ is a nontrivial linear character of R vanishing on $\Phi(R)$ not fixed by P, then every constituent of γ_R^G has dimension divisible by p, whence G acts transitively on the nontrivial linear characters of $R/\Phi(R)$ not fixed by P. This implies that on a simple G-quotient M of $R/\Phi(R)$, we have G acting transitively on the nontrivial elements, whence [Liebeck 1987, Appendix 1] and the fact that G is p-solvable implies that the action of G on M is $\mathrm{SL}_2(5) \times C$, where C is a cyclic group of order divisible by p, $|M| = r^2$ with $(p, r) = (7, 29)$ or $(29, 59)$. Write $M = R/R_0$. If P does not centralize R_0, then we can induce a nontrivial linear character of R_0 and obtain a different irreducible character of degree divisible by p (it will have a different kernel), a contradiction. Thus, P commutes with R_0. So P acts as nontrivial scalars on M and so M does not commute with $[R, R_0]$ unless it is trivial. So $R_0 = 1$ and $|R| = r^2$.

Suppose P does not commute with $F(G)$. We claim that this implies that $F^*(G)$ has order r^2 as above. If $E(G) \neq 1$, then clearly there is a character of $F^*(G)$ nontrivial on $E(G)$ that is not fixed by P (if P centralizes $E(G)$, then just take some character of $F^*(G)$ that is nontrivial on $\mathbf{O}_r(G)$ and $E(G)$; if P does not centralize $E(G)$, just take a character of $E(G)$ not fixed by P) and inducing this up gives a contradiction.

So $F^*(G)$ of order r^2 implies that $F^*(G)$ is its own centralizer and we already saw that $G/F^*(G) \cong \mathrm{SL}_2(5) \times \mathbb{Z}/p$, whence the result.

So we may assume that P commutes with $F(G)$. Then since P is not normal in G, P does not commute with $E(G)$. If $F^*(G) \neq E(G)$, then as above there will be at least two G-orbits of characters of G of degree divisible by p with distinct kernels. So $F^*(G) = E(G)$. We show that this cannot happen. There is no harm in passing to $G/\mathbf{Z}(E(G))$, and so we may assume that $E(G)$ is a direct product of nonabelian simple groups. The result now follows by Lemma 1.11. □

Nonabelian Sylow p-subgroup. In this subsection, we complete the proof of Theorem 1.1 by proving it under the hypothesis that G is p-solvable but not solvable and has a nonabelian Sylow p-subgroup. We will show that G has to satisfy conclusion (vi) of Theorem 1.1. We also obtain further restrictions on the structure of G in this situation.

Lemma 1.13. *Let p be a prime, and let G be a group having exactly one irreducible character whose degree is divisible by p. Suppose that G is p-solvable, but not solvable. Let P be a Sylow p-subgroup of G. If P is nonabelian, then P is normal in G.*

Proof. Suppose not. Suppose that $\mathbf{O}_p(G) \neq 1$ and let Q be a minimal normal p-subgroup of G. If P/Q is nonabelian, then by induction P/Q is normal in G/Q, whence P is normal.

So P/Q is abelian, and recall that P is not normal in G. By the Itô–Michler theorem, G/Q has a character of degree divisible by p, and by hypothesis this character is unique. Since G/Q is p-solvable and not solvable, we may apply Theorem 1.12, and G/Q is as given there (in particular, $p = 7$ or 29). Also, there is a normal subgroup QR where R has order r^2 (with $r = 29$ or 59 depending upon p).

If R does not centralize Q, then as the element of order p in G/Q acts centrally on R (that is, as scalars), we see that on Q, the dimension of the fixed space of a Sylow p-subgroup of G on Q is $(1/p)\dim Q$ (Q is a free module for P/Q)— see Theorem 15.16 of [Isaacs 1976]. The number of Sylow p-subgroups in G is $|R| = r^2$ (since $G = N_G(P)R$). Thus, the total number of points fixed by a Sylow p-subgroup of G on Q is less than $r^2 p^{d/p}$, where $|Q| = p^d$.

Easily, we see that for both choices of r and p that $r^2 p^{d/p} < p^d$, and so there exists an element of Q (and similarly Q^*) that is fixed by no p-element outside Q. Thus, there is a linear character χ of Q whose G-orbit has size a multiple of p, whence any irreducible constituent of χ^G has dimension a multiple of p and so there are at least two irreducible characters of degree a multiple of p (one with Q in the kernel, one with Q not in the kernel).

Suppose that R centralizes Q. Then P is central in G/RQ and since Q is an irreducible G/RQ-module, P must act trivially on Q, whence $Q \le \mathbf{Z}(P)$. Since P/Q is cyclic, this implies that P is abelian, a contradiction.

So we may assume that $\mathbf{O}_p(G) = 1$. Let N be a minimal normal subgroup. So N is a p'-group. Let χ be any character of N. If P is not contained in the inertia group of χ, then every irreducible constituent of χ_N^G has dimension divisible by p. Since some irreducible character of G/N has dimension divisible by p, we have a contradiction. Indeed, the argument shows that P centralizes $\mathbf{O}_{p'}(G)$. Thus, $F^*(G)$ cannot be a p'-group, whence $\mathbf{O}_p(G) \neq 1$. This completes the proof. □

We now come to the proof of Theorem 1.1 under the additional hypothesis that G is p-solvable but not solvable and that G has a nonabelian Sylow p-subgroup. Combining Theorems 1.9, 1.12, and 1.14, we have a proof of Theorem 1.1.

Theorem 1.14. *Let p be a prime, and let G be a group having exactly one irreducible character whose degree is divisible by p. Suppose that G is p-solvable, but not solvable. If G has a nonabelian Sylow p-subgroup, then G satisfies conclusion (vi) of Theorem 1.1. Furthermore, $p = 11, 29$, or 59; $|P'| = p^2$; and $G/P \cong \mathrm{SL}_2(5) \times \mathbb{Z}/c(p)$, where $c(11) = 1$, $c(29) = 7$, and $c(59) = 29$.*

Proof. By Lemma 1.13, P is normal. Applying Lemma 1.8, G satisfies conclusion (vi) of Theorem 1.1. Let P be the Sylow p-subgroup, and let H be a p-complement in G. We know that $P'H$ is a doubly transitive Frobenius group. It follows that H is a Frobenius complement. Since H is nonsolvable, it follows by [Passman 1968, 20.2] that $|\mathbf{Z}(P)| = p^2$ with $p = 11, 29$, or 59 and $H = \mathrm{SL}_2(5) \times C$, where C is cyclic of order $c(p)$. □

Examples with exceptional doubly transitive Frobenius groups. In this subsection, we consider groups G that have the form of a coprime semidirect product of a group H acting on a semi-extraspecial p-group P where $|\mathbf{Z}(P)| = p^2$ and H is acting faithfully and irreducibly on $\mathbf{Z}(P)$. We find the possible values for $|P : P'|$ when H is either $\mathrm{SL}_2(3)$ or $\mathrm{SL}_2(5)$, and we find examples showing that each of the possible values occur. This will yield examples where G is a group that satisfies conclusion (vi) of Theorem 1.1 where HP' is any of the exceptional two-transitive Frobenius groups. In particular, this gives examples of groups G where G is a p-solvable group that is not solvable with a nonabelian Sylow subgroup P and exactly one irreducible character whose degree is divisible by p.

The key to this solution is trying to find $\mathbb{F}_p H$-modules V so that $\bigwedge^2(V)$ contains a 2-dimensional H-submodule W such that $W^\#$ consists of nondegenerate alternating forms on V. There is also is an intriguing connection with the McKay correspondence and the Dynkin diagram of affine extended E_8 when $H = \mathrm{SL}_2(5)$ and of affine extended E_6 when $H = \mathrm{SL}_2(3)$. (For more details on the McKay correspondence and Dynkin diagrams, see [McKay 1980] or [Steinberg 1985].) This was used in our initial approach to the problem but it is not needed in the solution.

In this section we address a slightly more general problem than is addressed in the remainder of the paper. We assume we have a group H with $\mathbf{Z}(H) \leq H'$ and $|Z(H)| = 2$ that is acting faithfully on a p-group P that satisfies the following conditions. In the examples needed for this paper, we will have $H = \mathrm{SL}_2(3)$ or $H = \mathrm{SL}_2(5)$.

Our hypotheses are as follows:

(i) p is a prime that is coprime to $|H|$.

(ii) P is a semi-extraspecial p-group with $|P : P'| = p^a$ and $Z = \mathbf{Z}(P) = P'$ has order p^2.

(iii) H acts on P with Z a faithful irreducible H-module (and we fix the isomorphism type).

Notice that H acting faithfully on Z implies that H is isomorphic to a subgroup of $\mathrm{SL}_2(p)$. For $H = \mathrm{SL}_2(3)$, this does not imply any further restrictions on p, but when $H = \mathrm{SL}_2(5)$, this implies that $p \equiv \pm 1 \bmod 5$. Set $V = P/Z$, and we view V as a module for H. We will determine the set of all positive integers a so that $|V| = p^a$ when $H = \mathrm{SL}_2(3)$ and $H = \mathrm{SL}_2(5)$. The particular values that occur depend on the residue class of p modulo 12 when $H = \mathrm{SL}_2(3)$ and modulo 60 when $H = \mathrm{SL}_2(5)$. In the next section, we will construct enough examples to show that there exists an example for each of the possible dimensions. The same techniques would essentially allow us to classify all the possible groups, but we do not pursue this here. Let V_1 and V_{-1} denote the eigenspaces of the action of $\mathbf{Z}(H)$ on V. Let P_1 and P_{-1} denote the inverse images of V_1 and V_{-1} in P.

Lemma 1.15. *Assume P and H satisfy the given hypotheses and P_1 and P_{-1} are defined as above. Then P_1 and P_{-1} are abelian, $a = 2b$ is even, and $|P_1| = |P_{-1}| = p^{b+2}$.*

Proof. Note that $Z(H)$ acts trivially on $[P_\epsilon, P_\epsilon]$ for $\epsilon = 1$ or $\epsilon = -1$. Since $Z(H)$ does not active trivially on Z, we must have $[P_\epsilon, P_\epsilon] \cap Z = 1$. Thus, P_1 and P_{-1} are abelian. Since P/Z_1 is extraspecial, this implies that $|P_{\pm 1} : Z| \leq p^{a/2}$. Since $p^a = |V_1||V_{-1}|$, we have that $a = 2b$ for some integer b and $|P_{\pm 1} : Z| = p^b$. The conclusion now follows. □

Given a subgroup $1 < Z_1 < Z$, we obtain an element of $(\bigwedge^2 V)^* = \mathrm{Hom}\big(\bigwedge^2 V, \mathbb{F}_p\big)$ (namely the homomorphism $aZ \wedge bZ \mapsto [a, b]Z_1$). The condition that P/Z_1 is extraspecial is equivalent to saying that this element is nondegenerate. The subspace of $\big(\bigwedge^2 V\big)^*$ generated by these elements yields a 2-dimensional irreducible

H-invariant subspace of $\left(\bigwedge^2 V\right)^*$ such that each nonzero element is nondegenerate. In the next section, we will use the fact that if such a 2-dimensional space exists, then we can construct P satisfying the above conditions.

The existence of a group P such that H acts as above implies that an H-invariant subspace W of $\left(\bigwedge^2 V\right)^*$ exists such that W is two dimensional and every nonzero element of W is nondegenerate. We choose a basis for V consisting of bases for V_1 and V_{-1}. Representing the elements in W as matrices in terms of this basis for V, we obtain matrices of the form

$$f(D) := \begin{pmatrix} 0 & D \\ -D^T & 0 \end{pmatrix},$$

where D is some element of $M_n(\mathbb{F}_p)$. The fact that the nonzero elements of W are nondegenerate is equivalent to saying that D is nonsingular for all the matrices D such that $f(D) \in W$.

We can conjugate and find a basis for W of the form $\{f(I), f(A)\}$ for some matrix A. Let $\tilde{W}$ be the $\mathbb{F}_p$-span of $\{I, A\}$. Observe that f is an isomorphism of the vector spaces W and $\tilde{W}$. Moreover, the condition that all of the nonzero matrices in $\tilde{W}$ are nonsingular is precisely equivalent to the condition that A has no eigenvalues in $\mathbb{F}_p$.

Let τ denote the inverse transpose map. We may write the action of H on W as $\operatorname{diag}(\rho_1(h), \rho_2(h)^\tau)$, where ρ_1 is the representation of H afforded by V_1 and ρ_2 is the representation of H afforded by V_2. It follows that ρ_1 has a kernel containing $\mathbf{Z}(H)$ and that ρ_2 is a faithful representation of H. The fact that W is H-invariant is precisely equivalent to saying that

$$\rho_1(h) W \rho_2(h)^{-1} = W$$

for every element $h \in H$. Since we are taking $\{f(I), f(A)\}$ as the basis for W, we see that $\rho_1(h)\rho_2(h)^{-1} = \rho_1(h) I \rho_2(h)^{-1} \in \tilde{W}$ and $\rho_1(h) A \rho_2(h)^{-1} \in \tilde{W}$ for all $h \in H$. Because $\rho_1(h) I \rho_2(h)^{-1} \in \tilde{W}$, there is a map ϕ from H to $\tilde{W}$ define by $\phi(h) = \rho_1(h) I \rho_2(h)^{-1}$. This satisfies $\rho_1(h) I = \phi(h)\rho_2(h)$ for all $h \in H$, and substituting, we obtain $\rho_1(h) A \rho_2(h)^{-1} = \phi(h)\rho_2(h) A \rho_2(h)^{-1}$. Since $\phi(h) \in \tilde{W}$, this implies that $\rho_2(h) A \rho_2(h)^{-1} \in \tilde{W}$. It follows that conjugation by $\rho_2(h)$ preserves the algebra $R = \mathbb{F}_p[A]$, and this conjugation action is therefore an algebra automorphism of R. We define $\rho : H \to \tilde{W}$ by $\rho(h) = \rho_1(h) A \rho_2(h)^{-1}$, and note that $\rho(h) = \phi(h)\rho_2(h) A \rho_2(h)^{-1}$. Since the nonzero elements of $\tilde{W}$ are nonsingular (i.e., invertible), it follows that $\phi(h) \in R^*$ for all $h \in H$, and this implies that the image of H under ρ is contained in $R^* \operatorname{Aut}(R)$, where R^* is the set of units in R.

Note that $\rho(H)$ acts by permutations on the primitive idempotents of R. Thus, we can decompose $R = R_1 \times \cdots \times R_m$, where each R_i is a subspace of R generated by an orbit of $\rho(H)$ on the primitive idempotents of R. We have $\dim R = \sum_{i=1}^m \dim R_i$.

We now work with R_1. We know that $\rho(H)$ acts transitively on the primitive idempotents of R_1. We have $R_1 = K[t_1]/t_1^e \times \cdots \times K[t_s]/t_s^e$, where K is a nontrivial extension field of $\mathbb{F}_p$. The fact that K is a nontrivial extension comes from the fact that the eigenvalues of A do not lie in $\mathbb{F}_p$. Since $Z(H)$ acts like -1, it stabilizes the subspaces generated by each of the t_i's.

Let H_1 denote the stabilizer of the subspace generated by t_1 which we call L_1, and note that $s = |H : H_1|$. Then $\rho_2(H_1)$ acts on $L_1 := K[t_1]/t_1^e$, and again, the element of order 2 in $Z(H) \leq H_1$ acts as -1. Thus, H_1 acts faithfully on L_1. Moreover, the action of H_1 on L_1 embeds in $L_1^* \operatorname{Aut}(L_1)$. Because H_1 has order prime to p, we see that in fact H_1 embeds in $K^* \operatorname{Aut}(K)$, and this implies that H_1 is either cyclic or metacyclic. If $|H_1| = 2$, then since K is a nontrivial extension of $\mathbb{F}_p$, we see that H_1 does not act irreducibly on L, and so we conclude that $|H_1| \neq 2$.

Suppose H_1 is cyclic of order $2q$, where q is an odd prime. If q divides $p-1$, then H_1 will not act irreducibly on L_1, so q does not divide $p-1$. Recall that H_1 is isomorphic to a subgroup of $\mathrm{SL}_2(p)$. Thus, it must be that q divides $p+1$. This implies that $[K : \mathbb{F}_p]$ is even and $p \equiv -1 \bmod q$.

Now, suppose H_1 is cyclic of order 4. If 4 does not divide $p-1$, then it is not difficult to see that $[K : \mathbb{F}_p]$ must be even. Suppose 4 does divide $p-1$. For H_1 to be acting irreducibly on L_1, it must be that a generator of H_1 is the product of a nontrivial element of L_1^* with a nontrivial element of $\operatorname{Aut}(L_1)$. In particular, this implies that 2 divides $[K : \mathbb{F}_p]$.

If H_1 is nonabelian, then since H_1 is metacyclic and contained in $\mathrm{SL}_2(p)$, we conclude that H_1' is cyclic and has index 2. This implies that 2 divides $[K : \mathbb{F}_p]$. If H_1 contains a subgroup isomorphic to the quaternions, then $p \equiv -1 \bmod 4$ since $K^* \operatorname{Aut}(K)$ does not contain any subgroups isomorphic to the quaternions when $p \equiv 1 \bmod 4$. We now assume that H_1 does not contain any subgroups isomorphic to the quaternions. Since H_1 is a subgroup of $\mathrm{SL}_2(p)$, this implies that the Sylow 2-subgroups of H_1 are abelian, and so $|H_1|$ is divisible by some odd prime q. Since H_1 is acting irreducibly, we see that $H_1/\mathbf{Z}(H)$ is contained in a dihedral group of order $2(p+1)$. This implies H_1' contains no elements of odd order whose order divides $p-1$. In particular, we must have q divides $p+1$ so $p \equiv -1 \bmod q$.

Theorem 1.16. *Let P be an semi-extraspecial p-group with $|\mathbf{Z}(P)| = p^2$ and $p > 5$ is a prime. Let a be the even integer so that $|P : \mathbf{Z}(P)| = p^a$. Suppose $H = \mathrm{SL}_2(5)$ acts via automorphisms on P such that H is acting faithfully and irreducibly on $Z(H)$. Then $p \equiv \pm 1 \bmod 5$ and the following holds*:

(i) *If $p \equiv 1 \bmod 60$, then $a = 120x$, where x is a positive integer.*

(ii) *If $p \equiv 11 \bmod 60$, then $a = 40x + 60y$, where x and y are nonnegative integers whose sum is positive.*

(iii) *If* $p \equiv 19 \bmod 60$, *then* $a = 24x + 60y$, *where* x *and* y *are nonnegative integers whose sum is positive.*

(iv) *If* $p \equiv 29 \bmod 60$, *then* $a = 24x + 40y$, *where* x *and* y *are nonnegative integers whose sum is positive.*

(v) *If* $p \equiv 31 \bmod 60$, *then* $a = 60x$, *where* x *is a positive integer.*

(vi) *If* $p \equiv 41 \bmod 60$, *then* $a = 40x$, *where* x *is a positive integer.*

(vii) *If* $p \equiv 49 \bmod 60$, *then* $a = 24x$, *where* x *is a positive integer.*

(viii) *If* $p \equiv 59 \bmod 60$, *then* $a = 24x + 40y + 60z$, *where* x, y, *and* z *are nonnegative integers whose sum is positive.*

Proof. We consider the possibility for a subgroup H_1 of $\mathrm{SL}_2(5)$ within the context of the previous argument. Since H_1 is cyclic or metacyclic and $\mathbf{Z}(H) \leq H_1$, we see that s is the index of a proper subgroup of A_5 and $s \neq 5$, whence $s = 6$, 10, 12, 15, 20, 30, or 60. Since $|H_1| \neq 2$, we conclude that $s \neq 60$. We see that one of the following holds:

(i) $s = 6$, $H_1 = 5.4$, $[K : \mathbb{F}_p]$ is even, $p \equiv -1 \bmod 5$, and $\dim R_1$ is a multiple of 12.

(ii) $s = 10$, $H_1 = 3.4$, $[K : \mathbb{F}_p]$ is even, $p \equiv -1 \bmod 3$, and $\dim R_1$ is a multiple of 20.

(iii) $s = 12$, H_1 is cyclic of order 10, $[K : \mathbb{F}_p]$ is even, $p \equiv -1 \bmod 5$, and $\dim R_1$ is a multiple of 24.

(iv) $s = 15$, $H_1 = Q_8$, $[K : \mathbb{F}_p]$ is even, $p \equiv -1 \bmod 4$, and $\dim R_1$ is a multiple of 30.

(v) $s = 20$, H_1 is cyclic of order 6, $[K : \mathbb{F}_p]$ is even, $p \equiv -1 \bmod 3$ and $\dim R_1$ is a multiple of 40.

(vi) $s = 30$, H_1 is cyclic of order 4, $[K : \mathbb{F}_p]$ is even, and $\dim R_1$ is a multiple of 60.

Since $\dim V_1 = \sum \dim R_i$, this gives possible dimensions for V_1 depending upon the congruences of p modulo 30. Note that $a = \dim V = 2 \dim V_1$. This gives the stated conclusion. □

We saw in Theorem 1.14 the primes p that arise in our case are $p = 11, 29$, or 59, and for these primes we obtain the following possibilities.

Corollary 1.17. *Assume the hypotheses of Theorem 1.16.*

(i) *If* $p = 11$, *then* $a = 40x + 60y$ *for some nonnegative integers* x *and* y.

(ii) *If* $p = 29$, *then* $a = 24x + 40y$ *for some nonnegative integers* x *and* y.

(iii) *If* $p = 59$, *then* $a = 24x + 40y + 60z$ *for some nonnegative integers* x, y, *and* z.

In the solvable examples, we have the group $\mathrm{SL}_2(3)$ acting on P. Thus, we have the following.

Theorem 1.18. *Let P be an semi-extraspecial p-group with $|\mathbf{Z}(P)| = p^2$ and $p > 3$ is a prime. Let a be the even integer so that $|P : \mathbf{Z}(P)| = p^a$. Suppose $H = \mathrm{SL}_2(3)$ acts via automorphisms on P such that H is acting faithfully and irreducibly on $Z(H)$. Then the following holds*:

(i) *If $p \equiv 1 \bmod 12$, then $a = 24x$, where x is a positive integer.*

(ii) *If $p \equiv 5 \bmod 12$, then $a = 16x + 24y$, where x and y are nonnegative integers whose sum is positive.*

(iii) *If $p \equiv 7 \bmod 12$, then $a = 12x$, where x is a positive integer.*

(iv) *If $p \equiv 11 \bmod 12$, then $a = 12x + 16y$, where x and y are nonnegative integers whose sum is positive.*

Proof. We consider the possibility for a subgroup H_1 of $\mathrm{SL}_2(3)$ within the context of the previous argument. Since H_1 is cyclic or metacyclic and $\mathbf{Z}(H) \leq H_1$, we see that s is the index of a proper subgroup of A_4, whence $s = 3$, 4, 6, or 12. Since $|H_1| \neq 2$, we conclude that $s \neq 12$. We see that one of the following holds:

(i) $s = 3$, $H_1 = Q_8$, $[K : \mathbb{F}_p]$ is even, $p \equiv -1 \bmod 4$, and $\dim R_1$ is a multiple of 6.

(ii) $s = 4$, H_1 is cyclic of order 6, $[K : \mathbb{F}_p]$ is even, $p \equiv -1 \bmod 3$, and $\dim R_1$ is a multiple of 8.

(iii) $s = 6$, H_1 is cyclic of order 4, $[K : \mathbb{F}_p]$ is even, and $\dim R_1$ is a multiple of 12.

Since $\dim V_1 = \sum \dim R_i$, this gives possible dimensions for V_1 depending upon the congruences of p modulo 30. Note that $a = \dim V = 2 \dim V_1$. This gives the stated conclusion. □

We will also construct examples for the solvable exceptional 2-transitive Frobenius groups. (For a list of the exceptional 2-transitive Frobenius groups, see Remark XII.9.5 in [Huppert and Blackburn 1982]. We note that matrix generators for these groups seem to be correct, even though the descriptions in (b) and (d) are not. In particular, in (b) $H \cong \widetilde{\mathrm{GL}_2(3)}$, which is isoclinic to but not isomorphic to $\mathrm{GL}_2(3)$, and in (d) $H \cong \widetilde{\mathrm{GL}_2(3)} \times \mathbb{Z}/11$. The matrix groups are also listed in [Huppert 1957].) The primes p that arise are $p = 5, 7, 11$, and 23. For these primes, we obtain the following.

Corollary 1.19. *Assume the hypotheses of Theorem 1.18.*

(i) *If $p = 5$, then $a = 16x + 24y$, where x and y are nonnegative integers whose sum is positive.*

(ii) *If $p = 7$, then $a = 12x$, where x is a positive integer.*

p	H	a	a
5	$\mathrm{SL}_2(3)$	$16x+24y$	$16w,\ w\geq 2$
7	$\widetilde{\mathrm{GL}_2(3)}$	$12x$	$12w,\ w\geq 1$
11	$\mathrm{SL}_2(3)\times\mathbb{Z}/5$	$12x+16y$	$4w,\ w\neq 1,2,5$
11	$\mathrm{SL}_2(5)$	$40x+60y$	$20w,\ w\geq 2$
23	$\widetilde{\mathrm{GL}_2(3)}\times\mathbb{Z}/11$	$12x+16y$	$4w,\ w\neq 1,2,5$
29	$\mathrm{SL}_2(5)\times\mathbb{Z}/7$	$24x+40y$	$8w,\ w\geq 3,\ w\neq 4,7$
59	$\mathrm{SL}_2(5)\times\mathbb{Z}/29$	$24x+40y+60z$	$4w,\ w\geq 6,$ $w\neq 7,8,9,11,13,14,17,19,23,29$

Table 1. Values in Theorem 1.20.

(iii) *If $p=11$, then $a=12x+16y$, where x and y are nonnegative integers whose sum is positive.*

(iv) *If $p=23$, then $a=12x+16y$, where x and y are nonnegative integers whose sum is positive.*

We will now construct examples showing that there are groups satisfying our hypotheses.

Theorem 1.20. *For each prime p, group H, and value a in Table 1, there exist groups P and G so that P is a semi-extraspecial p-group with $|P'|=p^2$ and $|P:P'|=p^a$ such that H acts faithfully by automorphisms on P and G is the resulting semidirect product. In addition, $\mathrm{Irr}(G)$ has a unique character with degree divisible by p and $P'H$ is a two-transitive Frobenius group.*

Proof. Suppose first that P and Q are semi-extraspecial p-groups with $|\mathbf{Z}(P)|=|\mathbf{Z}(Q)|=p^2$, and $\dim P/\mathbf{Z}(P)=a_1$ and $\dim Q/\mathbf{Z}(Q)=a_2$. Suppose that the group H acts via automorphisms on P and Q so that the action of H on $\mathbf{Z}(P)$ is isomorphic to its action on $\mathbf{Z}(Q)$ and this action is faithful and irreducible. Let R be the central product of PQ (that is, identifying $\mathbf{Z}(Q)$ and $\mathbf{Z}(P)$). Then $|\mathbf{Z}(R)|=p^2$ and R/Z_1 is extraspecial for any proper nontrivial subgroup Z_1 of $\mathbf{Z}(R)$ since R/Z_1 is the central product of the two extraspecial groups P/Z_1 and Q/Z_1. It follows that R is semi-extraspecial and $\dim R/\mathbf{Z}(R)=a_1+a_2$. Note that H will act via automorphisms on R and that the action of H on $\mathbf{Z}(R)$ is faithful and irreducible. This shows that the set of dimensions that can occur is closed under sums. In particular, the entries in the first column for values for a are linear combinations, and it suffices to show that there exist examples for the generators of these linear combinations.

Let p and H be entries for the first two columns of some row in Table 1, and let a be one of the generators of the third column. Observe that H is isomorphic to a subgroup of $\mathrm{GL}_2(p)$, so it has a faithful 2-dimensional module Z over $\mathbb{F}_p$. Let H_2 be a subgroup of H containing $Z(H)$ whose index is $a/4$. We have $H_2=H_1\times C$.

We take $C = Z/c$, where $c = 1$ when p is 5 or 7, or $p = 11$ and H is not solvable. Otherwise, $c = 5$ when $p = 11$ and H is solvable, $c = 11$ when $p = 23$, $c = 7$ when $p = 29$, and $c = 29$ when $p = 59$. When H is not solvable, then H_1 will have order $4r$, where $r = 2$, 3, or 5, and we note that p satisfies $p \equiv -1 \bmod 2r$. When H is solvable, we have $|H_1| = 4$ or 6 when $p = 5$, $|H_1| = 16$ when $p = 7$, $|H_1| = 6$ or 8 when $p = 11$, and $|H_1| = 12$ or 16 when $p = 23$.

Let X be the 2-dimensional $\mathbb{F}_p H_2$-module where the image of H_1 is generated by an involution of determinant -1 and the module for C is the restriction of Z. Let Y be the 2-dimensional $\mathbb{F}_p H_2$-module which is the restriction of Z to H_2. We can view H_1 as being contained in a subgroup of order $2(p+1)$ in $\mathrm{GL}_2(p)$. Notice also that C will be isomorphic to a subgroup of the center of $\mathrm{GL}_2(p)$, and thus $H_2 = H_1 \times C$ can be thought of as normalizing a given nonsplit torus in $\mathrm{GL}_2(p)$. It follows that $\mathrm{Hom}(X, Y)$ contains a 2-dimensional invariant H_2-submodule $U \cong Y$ with U consisting of invertible elements (namely the algebra generated by the nonsplit torus which is a quadratic extension of $\mathbb{F}_p$).

It is straightforward to compute (by Frobenius reciprocity) that U^H contains a unique submodule W isomorphic to Z. Note that $U^H \leq \mathrm{Hom}(X, Y)^H_{H_2} \leq \mathrm{Hom}(X^H, Y^H)$. Since U consists of invertible elements, it is straightforward to see that W does as well. We take $V = X^H_{H_2} + Y^H_{H_2}$. We identify $\mathrm{Hom}\big(X^H_{H_2}, Y^H_{H_2}\big)$ with a submodule of $\big(\bigwedge^2 V\big)^*$ as follows. Let $\{e_1, \dots, e_n\}$ be a basis for X^H and $\{f_1, \dots, f_n\}$ be a basis for y^H. Given an element $\phi \in \mathrm{Hom}(X^H, y^H)$, we define the matrix $D = D_\phi$ whose (i, j)-entry is $\phi(e_i, f_j)$. This defines an element

$$\begin{pmatrix} 0 & D \\ -D^T & 0 \end{pmatrix} \in \big(\textstyle\bigwedge^2 V\big)^*.$$

This yields W embedded as a submodule of $\big(\bigwedge^2 V\big)^*$, and as we mentioned earlier, this allows us to construct a semi-extraspecial p-group P so that $P/\mathbf{Z}(P) \cong V$ and $\mathbf{Z}(P) \cong W \cong Z$ (as H-modules) with the desired properties. □

This shows that all possible dimensions given in the previous section do occur and completes the proof of the main theorem.

2. Non-p-solvable groups

The goal of this section is to prove the following theorem. Notice that this includes the non-p-solvable portion of the main theorem. We do get somewhat more information. In particular, we determine precisely the degree of the one character whose degree is divisible by p.

Theorem 2.1. *Let p be a prime and let G be a finite non-p-solvable group. Then G has exactly one complex irreducible character of degree (say D) divisible by p if and only if one of the following holds.*

(i) *$D = q = p^a \geq 4$, a power of p. Furthermore, $G = \mathrm{PSL}_2(q)$, $\mathrm{SL}_2(q)$, or $G/V = \mathrm{SL}_2(q)$, where V is a minimal normal elementary abelian p-subgroup of G having order q^2 and can be viewed as a 2-dimensional irreducible module of G/V over $\mathrm{End}_{G/V}(V) \cong \mathbb{F}_{q^2}$.*

(ii) $(G, p, D) = (\mathsf{S}_5, 3, 6)$, $(M_{11}, 3, 45)$.

For brevity, we will call G an NP1-group (for a fixed prime p) if G is a finite non-p-solvable group with a unique irreducible character χ of degree divisible by p, in which case we let $D := \chi(1)$.

We make use of results regarding character degree graphs. The graph $\Delta(G)$ is the graph with vertex set $\rho(G)$, which is the set of primes dividing degrees in $\{\chi(1) \mid \chi \in \mathrm{Irr}(G)\}$. There is an edge between distinct primes p and q if pq divides $\chi(1)$ for some $\chi \in \mathrm{Irr}(G)$. We will make use of the results regarding nonsolvable groups G where $\Delta(G)$ is disconnected in [Lewis and White 2003] and [Manz et al. 1988]. We prove the following statement combining the results on disconnected graphs with some results of [Isaacs et al. 2009]:

Proposition 2.2. *Let G be an NP1-group. Then one of the following holds.*

(i) *$D = q = p^a \geq 4$ is a power of p, and $G/K = \mathrm{PSL}_2(q)$. Furthermore, $G = \mathrm{PSL}_2(q)$, $\mathrm{SL}_2(q)$, or $G/V = \mathrm{SL}_2(q)$, where V is a minimal normal elementary abelian p-subgroup G of order q^2 and can be viewed as a 2-dimensional irreducible module of G/V over $\mathrm{End}_{G/V}(V) \cong \mathbb{F}_{q^2}$.*

(ii) *Let $U = \mathbf{O}_p(G)$ and $K/U = \mathbf{O}_{p'}(G/U)$. Then K is the p-solvable radical of G and U is abelian. If $N \leq K$ is a normal subgroup of G and $1_N \neq \lambda \in \mathrm{Irr}(N)$, then the inertia group $I_G(\lambda)$ of λ has index coprime to p in G. Finally, $(G/K, p, D)$ is either $(\mathsf{S}_5, 3, 6)$ or $(M_{11}, 3, 45)$.*

Proof. By Theorems A and C of [Isaacs et al. 2009], we know that if $U = \mathbf{O}_p(G)$ and $K/U = \mathbf{O}_{p'}(G/U)$, then K is the p-solvable radical of G and U is abelian. We also obtain from there the fact that if $N \leq K$ is a normal subgroup of G and $1_N \neq \lambda \in \mathrm{Irr}(N)$, then the inertia group $I_G(\lambda)$ of λ has index coprime to p in G. Finally, that result also asserts that $\mathrm{soc}(G/K) = S/K =: \bar{S}$ is a nonabelian simple group of order divisible by p, and p is coprime to $|G/S|$. By the Itô–Michler theorem, we may assume that the unique irreducible character χ of degree divisible by p of G is actually a character of G/K. In particular, every irreducible character of K has degree coprime to p, whence the (unique) Sylow p-subgroup U of K must be abelian again by the Itô–Michler theorem. Next, let $N \leq K$ be normal in G and $1_N \neq \lambda \in \mathrm{Irr}(N)$. Consider any $\rho \in \mathrm{Irr}(G|\lambda)$. Since ρ_N contains $\lambda \neq 1_N$, $\mathrm{Ker}(\rho)$ cannot contain N, and so does not contain K. Thus $\rho \neq \chi$ and so has degree coprime to p. It follows that $[G : I_G(\lambda)]$ is coprime to p.

Next, by Corollary 7.5 of [Isaacs et al. 2009] applied to G/K, one of the following possibilities occurs:

(a) $\bar{S} \cong \mathrm{PSL}_2(q)$ with $p|q$, and $D = q$.

(b) $\bar{S} \cong \mathrm{PSL}_2(q)$, $q = r^f$ for some prime $r \neq p$, $2 < p|(q-\epsilon)$ for some $\epsilon = \pm 1$, and $D|(q-\epsilon)f$.

(c) $(G/K, p, D) \in \{(M_{11}, 3, 45), (J_1, 3, 120), (J_1, 5, 120)\}$.

Note that J_1 has 3 irreducible characters of degree 120. Hence, in the case of (c) we arrive at conclusion (ii).

Suppose now that we are in the case of (a). Consider the Steinberg character St of $S/K = \mathrm{PSL}_2(q)$, of degree q. By the main result of [Feit 1993], St extends to G/K. Hence $D = q$, and since p divides the degree of no other character in $\mathrm{Irr}(G)$, this implies that p is an isolated vertex in $\Delta(G)$, and so $\Delta(G)$ is disconnected. In [Manz et al. 1988], it is shown that $\Delta(G)$ has at most three connected components. In Theorems 4.1 and 6.4 of [Lewis and White 2003], it is shown that if G is nonsolvable and $\Delta(G)$ has two or three connected components, then G has a normal subgroup N so that G/N is abelian and N is either $\mathrm{PSL}_2(q)$, $\mathrm{SL}_2(q)$, or there is a normal subgroup V so that $N/V \cong \mathrm{SL}_2(q)$ and V is elementary abelian of order q^2 and N/V acts transitively on the nonidentity elements of V. As before, the Steinberg character of N will extend to G, and, by Gallagher's theorem, G has at least $|G/N|$ irreducible characters of degree divisible by q. Since G is an $\mathsf{NP1}$-group, it follows that $|G/N| = 1$, whence $G = N$ and we arrive at conclusion (i).

From now on, we may assume that (b) holds. Since p is coprime to $|G/S|$, all irreducible constituents α_i of χ_S have degree divisible by p. On the other hand, if $\beta \in \mathrm{Irr}(\bar{S})$ has degree divisible by p, then any $\gamma \in \mathrm{Irr}(G/K|\beta)$ also has degree divisible by p, and so $\gamma = \chi$ as G is an $\mathsf{NP1}$-group. Thus β must be one of the α_i. We have shown that G/S, and so $\mathrm{Aut}(\bar{S})$ as well, acts transitively on the set of irreducible characters of $\bar{S}$ of degree divisible by p. Also recall that $2 < p|(q-\epsilon)$. Now we distinguish the following subcases.

(b1) $q \equiv -\epsilon \pmod 4$. In this case, p divides $(q-\epsilon)/2$, which is odd, and $\bar{S}$ has two irreducible (Weil) characters of degree $(q-\epsilon)/2$. Now if $q \geq 7$, then $\bar{S}$ also has irreducible characters of degree $(q-\epsilon)$ which are obviously not $\mathrm{Aut}(\bar{S})$-conjugate to the ones of degree $(q-\epsilon)/2$. Hence $q = 5$, $\epsilon = -1$, $p = 3$, $\bar{S} \cong \mathsf{A}_5$, and $\mathsf{A}_5 \leq G/K \leq \mathsf{S}_5$. Since A_5 has two irreducible characters of degree 3, we conclude that $(G/K, p, D) = (\mathsf{S}_5, 3, 6)$, as stated in (ii).

(b2) $q \equiv \epsilon \pmod 4$. In this case, $4p \geq 12$ divides $q-\epsilon$, so $q \geq 11$. Then $\bar{S}$ has $(q-2+\epsilon)/4$ irreducible characters of degree $q-\epsilon$, and each of them extends to $\mathrm{PGL}_2(q)$. Notice that $\mathrm{Aut}(\bar{S}) = \mathrm{PGL}_2(q) : C_f$, so the $\mathrm{Aut}(S)$-orbit of any such θ

has length at most f. Since $q = r^f \geq 11$, we have that $(q-2+\epsilon)/4 > f$, and so $\mathrm{Aut}(\bar{S})$ cannot act transitively on the characters of $\bar{S}$ of degree divisible by p.

(b3) $q = 2^f$ with $f \geq 3$. Then $\bar{S}$ has $(q-1+\epsilon)/2$ irreducible characters of degree $q-\epsilon$. Notice that $\mathrm{Aut}(\bar{S}) = \bar{S} : C_f$, so the $\mathrm{Aut}(S)$-orbit of any such θ has length at most f. The transitivity of $\mathrm{Aut}(\bar{S})$ on the set of characters of $\bar{S}$ of degree divisible by p now implies that $f = 3$, $\epsilon = -1$, $p = 3$, and $\bar{S} = \mathrm{SL}_2(8) \leq G/K \leq \bar{S} : C_3$. But $\mathrm{SL}_2(8)$ has 3 irreducible characters of degree 9, and $\mathrm{Aut}(\mathrm{SL}_2(8)) = \bar{S} : C_3$ has irreducible characters of both degrees 21 and 27. This contradiction completes the proof of the proposition. □

We now show that the groups satisfying conclusion (i) of Theorem 2.1 have only one irreducible character whose degree is divisible by p.

Proposition 2.3. *Let $D = q = p^f \geq 4$ and assume that the finite group G satisfies conclusion* (i) *of Theorem 2.1. Then G has exactly one irreducible character θ of degree divisible by p. Furthermore, $\theta(1) = D$.*

Proof. (1) The statement is clear if $G = \mathrm{PSL}_2(q)$ or $\mathrm{SL}_2(q)$. So we consider the third possibility, with $S := G/V \cong \mathrm{SL}_2(q)$ and $|V| = q^2$. Note that S has a unique irreducible character of degree divisible by p, namely the Steinberg character of S, of degree q. So it suffices to show that any $\chi \in \mathrm{Irr}(G \mid V)$ has degree $q^2 - 1$.

Fix a nontrivial character $\lambda \in \mathrm{Irr}(V)$. Since S acts transitively on the nonzero vectors of V and V^*, the stabilizer of λ in G is a subgroup I of order $|G|/(q^2-1) = q^3$; in particular, $P := I/V \in \mathrm{Syl}_p(S)$. Also we have that $\chi = \mu^G$ for some $\mu \in \mathrm{Irr}(I|\lambda)$. Thus the map $\mu \mapsto \mu^G$ yields a bijection between $\mathrm{Irr}(I|\lambda)$ and $\mathrm{Irr}(G \mid V)$.

(2) It suffices to show that there is at least one $\chi \in \mathrm{Irr}(G \mid V)$ of degree $q^2 - 1$. Indeed, if this is the case then there exists $\nu \in \mathrm{Irr}(I)$ such that $\nu|_V = \lambda$. Now by Gallagher's theorem, $\mathrm{Irr}(P|\lambda) = \{\nu\gamma \mid \gamma \in \mathrm{Irr}(I/V)\}$ consists of exactly q characters, all of degree 1 as P is elementary abelian of order q. It follows that all characters in $\mathrm{Irr}(G \mid V)$ have degree $q^2 - 1$.

(3) Assume that the extension $G = VS$ is split. Then I splits over V and so $I/\mathrm{Ker}(\lambda)$ is a split extension of $V/\mathrm{Ker}(\lambda) \cong C_p$ by P. Since P fixes λ, P centralizes $V/\mathrm{Ker}(\lambda)$, whence $I/\mathrm{Ker}(\lambda)$ is abelian. Hence all irreducible characters of $I/\mathrm{Ker}(\lambda)$ are of degree 1 and so are all characters in $\mathrm{Irr}(I|\lambda)$, whence we are done as in (2).

Suppose now that the extension $G = VS$ is nonsplit. It follows that $H^2(S, V) \neq 0$. By a result of McLaughlin (see [Sah 1977, Proposition 4.4]), this can happen only when $p = 2$ and $f \geq 3$. In the exceptional case, $H^2(S, V) = \mathbb{F}_q$, and so there is a unique (up to isomorphism) such nonsplit extension. By [Kostrikin and Tiep 1994, Theorem 1.3.7], this nonsplit extension can be realized as the commutator subgroup of the group of all automorphisms of the standard orthogonal decomposition of the complex simple Lie algebra $\mathscr{L}$ of type A_{q-1}. Now the action of G on $\mathscr{L}$ gives rise to a faithful irreducible representation of degree $\dim(\mathscr{L}) = q^2 - 1$, and so we are done by (2). □

In what follows we will keep the notation from conclusion (ii) of Proposition 2.2.

Lemma 2.4. *Let $p = 3$ and suppose that G is an* NP1*-group with a normal subgroup $N \cong \mathsf{S}_5$ or M_{11}. Then $G = N$.*

Proof. Certainly $N' \lhd G$ and N' has an irreducible real-valued character α of odd degree $\alpha(1) = 3$ or 45. Since N' is perfect, $o(\alpha) = 1$. By [Navarro and Tiep 2008, Theorem 2.3], α extends to $I := I_G(\alpha)$. Hence, by Gallagher's theorem, there is a bijection between $\mathrm{Irr}(I|\alpha)$ and $\mathrm{Irr}(I/N')$. On the other hand, $\mathrm{Irr}(G|\alpha)$ lies in a bijective correspondence with $\mathrm{Irr}(I|\alpha)$, and every $\chi \in \mathrm{Irr}(G|\alpha)$ has degree divisible by 3. It follows that $|\mathrm{Irr}(I/N')| = 1$, and so $I = N' \leq N$. In particular, $I = I_N(\alpha)$. Now observe that N, and so G, acts transitively on the set X of irreducible characters of N' of degree equal to $\alpha(1)$. It follows that $[G : I] = |X| = [N : I]$, whence $G = N$. □

In the situation of Proposition 2.2(ii), we set $G_1 = G$ if $G/K = M_{11}$, and let G_1 be the unique subgroup of index 2 containing K if $G/K = \mathsf{S}_5$ (so that $G_1/K = \mathsf{A}_5$).

Lemma 2.5. *Let G be an* NP1*-group satisfying conclusion* (ii) *of Proposition 2.2. Let $R \lhd G$ be such that G/R is* 3*-solvable. Then $R \geq G_1$.*

Proof. Note that G/KR is a 3-solvable quotient of $G/K = \mathsf{S}_5$ or M_{11}, hence $KR \geq G_1$. Assume first that $KR = G$. Then $G/(K \cap R)$ contains the normal subgroup $R/(K \cap R) \cong G/K$, and $G/(K \cap R)$ is certainly an NP1-group. By Lemma 2.4, $R/(K \cap R) = G/(K \cap R)$ and so $R = G$.

We may now assume that $G/K = \mathsf{S}_5$ and $KR = G_1$. Again $H := G/(K \cap R)$ is an NP1-group with the normal subgroup $M := R/(K \cap R) \cong \mathsf{A}_5$. Then H has a unique irreducible character ρ of degree divisible by 3. It follows that H acts transitively on the set $\{\alpha, \beta\}$ of irreducible characters of degree 3 of M; in particular, $I := I_H(\alpha)$ has index 2 in H. As in the proof of Lemma 2.4, we see that the real character α extends to I, and there is a bijection between $\mathrm{Irr}(H|\alpha)$ and $\mathrm{Irr}(I/M)$. Since H is an NP1-group, it follows that $I = M$, and so $[G : R] = [H : M] = [H : I] = 2$. Recall that $G_1 = KR \geq R$ and $[G : G_1] = 2$. Hence, we must have that $R = G_1$. □

Lemma 2.6. *Let G be an* NP1*-group satisfying conclusion* (ii) *of Proposition 2.2. Then K is solvable.*

Proof. Assume the contrary: L/M is a nonabelian chief factor of G in K/U. Certainly G/M is also an NP1-group. So, modding out by M we may assume that $M = 1$ and that $L = T_1 \times \cdots \times T_n$ is a minimal normal subgroup of G, with $T_i \cong T$, a nonabelian simple $3'$-group. Now G acts on the set $\{T_1, \ldots, T_n\}$ inducing a transitive subgroup X of S_n.

First suppose that p divides $|X|$. By [Casolo and Dolfi 2009, Lemma 8] we can relabel the factors T_i and find $1 \leq k \leq k+l \leq n$ such that

$$\mathrm{Stab}_X(\{1, 2, \ldots, k\}, \{k+1, k+2, \ldots, k+l\})$$

has index divisible by p in X. It follows that

$$N := \mathbf{N}_G(T_1 \times T_2 \times \cdots \times T_k, T_{k+1} \times T_{k+2} \times \cdots \times T_{k+l})$$

has index divisible by p in G. Also, since T is simple nonabelian, we can find $\alpha, \beta \in \mathrm{Irr}(T)$ of distinct degrees larger than 1. Consider the irreducible character

$$\gamma := (\underbrace{\alpha \times \cdots \times \alpha}_{k}) \times (\underbrace{\beta \times \cdots \times \beta}_{l}) \times (\underbrace{1_T \times \cdots \times 1_T}_{n-k-l})$$

of L. Then $I_G(\gamma) \leq N$ and so has index divisible by p in G, contradicting Proposition 2.2.

We have X is a p'-group. Thus $X = G/Y$ is a $3'$-group, if Y denotes the kernel of the action of G on $\{T_1, \ldots, T_n\}$. By Lemma 2.5, $Y \geq G_1$; in particular, $|X| \leq 2$ and so $n \leq 2$. Recall that T is a simple nonabelian p'-group. Hence the condition $n \leq 2$ now implies that $\mathrm{Aut}(L) \leq \mathrm{Aut}(T) \wr \mathrm{S}_2$ is 3-solvable, whence $G/\mathbf{C}_G(L)$, as a subgroup of $\mathrm{Aut}(T)$, is also 3-solvable. Again by Lemma 2.5, $\mathbf{C}_G(L) \geq G_1$. Since $[G : G_1] \leq 2$, $\mathbf{C}_G(L)$ contains L and so L is abelian, a contradiction. □

Lemma 2.7. *Let $S = M_{11}$, $p = 3$, and let r be a prime. Let G be an* NP1*-group with a minimal normal r-subgroup V such that $G/V = S$. Then $V = 1$.*

Proof. Assume the contrary: $V \neq 1$. Identify $\mathrm{Irr}(V)$ with the dual module V^*. By Proposition 2.2, every nonzero $v \in V^*$ is fixed by a Sylow 3-subgroup of S. We will view V as an absolutely irreducible $\mathbb{E}S$-module of dimension n, where $\mathbb{E} := \mathrm{End}_S(V) = \mathbb{F}_t$ for some power t of r. If $n = 1$, then $V \leq \mathbf{Z}(G)$, $V = C_p$, and $G \cong V \times S$ (as S has trivial Schur multiplier), contradicting Lemma 2.4. So we will assume $n \geq 2$ and estimate the number N of nonzero elements of V^* that are fixed by a Sylow 3-subgroup P of S. Let $g \in S$ have order 3. As mentioned in [Guralnick and Saxl 2003], S is generated by three conjugates of g. Hence by [Guralnick and Tiep 2005, Lemma 3.2], g can fix at most $t^{\lfloor 2n/3 \rfloor}$ elements of V^*. It follows that $N \leq t^{\lfloor 2n/3 \rfloor} - 1$. Also note that S has 55 Sylow 3-subgroups.

Consider the case $r \neq 3$. Then $n \geq 9$ if $r = 11$ and $n \geq 10$ otherwise (see [Jansen et al. 1995]). Since $55N \geq t^n - 1 = |V^* \setminus \{0\}|$, we must have that $r = 2$ and $n \leq 21$. In the remaining cases, using Brauer characters as given in [Jansen et al. 1995], we see that g can fix at most 2^4 elements of V^*. Thus $N \leq 2^4 - 1 < (2^{10} - 1)/55 \leq (t^n - 1)/55$, again a contradiction.

Now suppose $r = 3$; in particular, $n \geq 5$. If $t \geq 9$, or if $t = 3$ but $n \geq 10$, then again $55N < t^n - 1$, again a contradiction. Thus $t = 3$ and $n = 5$. Note that M_{11}

has two irreducible 5-dimensional modules over $\mathbb{F}_3$, which are dual to each other. Moreover, the S-orbits on them are of lengths 1, 22, and 220, resp. 1, 110, and 132 (see [Liebeck 1987, Table 14]). So V^* must be isomorphic to the former module. Direct calculations using [GAP 2004] done by T. Breuer show that in this case the point stabilizer for the vectors in the orbit of length 22 are isomorphic to A_6 (and isomorphic to $\mathsf{S}_3 \times \mathsf{S}_3$ for the orbit of length 220). Consider a vector v from the first orbit and view it as $\lambda \in \mathrm{Irr}(V)$, with stabilizer I in G. Then $I/\mathrm{Ker}(\lambda) \simeq C_3 \cdot \mathsf{A}_6$, which may be split or nonsplit. In either case, there is $\mu \in \mathrm{Irr}(I|\lambda)$ of degree 9. Now $\mu^G \in \mathrm{Irr}(G|\lambda)$ has degree $9 \cdot 22 = 198$, a contradiction (since $G/V = S$ has an irreducible character of degree 45). □

Lemma 2.8. *Let $S = \mathsf{A}_5$, $p = 3$, and let r be a prime. Let H be a finite group with a minimal normal r-subgroup V such that $H/V = S$. Assume that the kernel of every irreducible character of degree divisible by 3 of H contains V. Then $V = 1$.*

Proof. Assume the contrary: $V \neq 1$. The condition on H implies by Clifford's theorem that the inertia group of every nontrivial $\lambda \in \mathrm{Irr}(V)$ has $3'$-index in H. Identifying $\mathrm{Irr}(V)$ with V^*, we see that every nonzero $v \in V^*$ is fixed by a Sylow 3-subgroup of S. View V as an absolutely irreducible $\mathbb{E}S$-module of dimension n, where $\mathbb{E} := \mathrm{End}_S(V) = \mathbb{F}_t$ for some power t of r. If $n = 1$, then $V \leq \mathbf{Z}(G)$, $V = C_3$, and so $G \cong V \times S$ (as the Schur multiplier of S equals C_2). In this case, G has irreducible characters of degree 3 which are nontrivial at V, a contradiction. So we will assume $n \geq 2$. By [Guralnick and Saxl 2003, Lemma 3.1], S is generated by two conjugates of a Sylow 3-subgroup P of S. It follows by [Guralnick and Tiep 2005, Lemma 3.2] that P can fix at most $t^{\lfloor n/2 \rfloor}$ elements of V^*. Since S has 10 Sylow 3-subgroups, we must have

$$10(t^{\lfloor n/2 \rfloor} - 1) \geq |V^* \setminus \{0\}| = t^n - 1;$$

in particular, $t^n \leq 81$.

Now if $r \geq 5$, then $n \geq 3$, a contradiction. Suppose $r = 3$. Then $n = 3$ or 4. In the former case, $t \geq 9$ (see [Jansen et al. 1995]), a contradiction. So $n = 4$ and $t = 3$. We can now realize V^* as the deleted permutation module

$$\mathbb{F}_3^4 = \left\{ \sum_{i=1}^{5} a_i e_i \;\middle|\; a_i \in \mathbb{F}_3, \sum_{i=1}^{5} a_i = 0 \right\}$$

for A_5. But in this case the A_5-orbit of $v := e_1 + e_2 - e_3 - e_4 \in V^*$ has length 30, a contradiction.

Suppose now $r = 2$. If $n = 2$, then P has no nonzero fixed points on V^*, a contradiction. Thus $n = 4$, $t = 2$. In this case, either $H \cong V \times \mathsf{A}_5$ or H is perfect. In the former case, H has irreducible characters of degree 3 which are nontrivial at V, a contradiction. In the latter case, using [GAP 2004] one can check that H has irreducible characters of degree 15 which are nontrivial at V, again a contradiction. □

Proposition 2.9. *Let G be an* NP1*-group satisfying conclusion* (ii) *of Proposition 2.2. Then G satisfies conclusion* (ii) *of Theorem 2.1. Conversely, if $G = \mathsf{S}_5$ or M_{11}, then G is an* NP1*-group.*

Proof. Suppose G satisfies Proposition 2.2(ii). By Proposition 2.2 and Lemma 2.6, K is solvable. Let G be a minimal counterexample, so that $K \neq 1$. By minimality, K is a minimal normal r-subgroup for some prime r. By Lemma 2.7, $G/K = \mathsf{S}_5$. Since G/K already has an irreducible character of degree 6, every irreducible character of G of degree divisible by 3 must be nontrivial at K. The same is true for the normal subgroup G_1 of G (recall $G_1/K \cong \mathsf{A}_5$). Modding out by a suitable normal subgroup inside K, we may assume that K is a minimal normal subgroup of G_1. Now we can apply Lemma 2.8 to get a contradiction. The converse statement is obvious. □

We have completed the proof of Theorem 2.1 and of the main theorem.

References

[Casolo and Dolfi 2009] C. Casolo and S. Dolfi, "Products of primes in conjugacy class sizes and irreducible character degrees", *Israel J. Math.* **174** (2009), 403–418. MR 2011e:20039 Zbl 1191.20024

[Chillag and Macdonald 1984] D. Chillag and I. D. Macdonald, "Generalized Frobenius groups", *Israel J. Math.* **47**:2-3 (1984), 111–122. MR 85h:20035 Zbl 0533.20012

[Dolfi et al. 2009] S. Dolfi, A. Moretó, and G. Navarro, "The groups with exactly one class of size a multiple of p", *J. Group Theory* **12**:2 (2009), 219–234. MR 2010b:20031 Zbl 1166.20018

[Dolfi et al. 2013] S. Dolfi, G. Navarro, and P. H. Tiep, "Finite groups whose same degree characters are Galois conjugate", *Israel J. Math.* **198**:1 (2013), 283–331. MR 3096641

[Feit 1993] W. Feit, "Extending Steinberg characters", pp. 1–9 in *Linear algebraic groups and their representations* (Los Angeles, 1992), edited by R. S. Elman et al., Contemp. Math. **153**, Amer. Math. Soc., Providence, RI, 1993. MR 94k:20012 Zbl 0823.20013

[Fernández-Alcober and Moretó 2001] G. A. Fernández-Alcober and A. Moretó, "Groups with two extreme character degrees and their normal subgroups", *Trans. Amer. Math. Soc.* **353**:6 (2001), 2171–2192. MR 2002a:20007 Zbl 0968.20005

[Gagola 1983] S. M. Gagola, Jr., "Characters vanishing on all but two conjugacy classes", *Pacific J. Math.* **109**:2 (1983), 363–385. MR 85e:20009 Zbl 0536.20005

[GAP 2004] The GAP group, *GAP — groups, algorithms, and programming, version 4.4*, 2004, Available at http://www.gap-system.org.

[Guralnick and Saxl 2003] R. M. Guralnick and J. Saxl, "Generation of finite almost simple groups by conjugates", *J. Algebra* **268**:2 (2003), 519–571. MR 2005f:20057 Zbl 1037.20016

[Guralnick and Tiep 2005] R. M. Guralnick and P. H. Tiep, "The non-coprime $k(GV)$ problem", *J. Algebra* **293**:1 (2005), 185–242. MR 2006g:20018 Zbl 1083.20006

[Huppert 1957] B. Huppert, "Zweifach transitive, auflösbare Permutationsgruppen", *Math. Z.* **68** (1957), 126–150. MR 20 #904 Zbl 0079.25502

[Huppert and Blackburn 1982] B. Huppert and N. Blackburn, *Finite groups, III*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **243**, Springer, Berlin, 1982. MR 84i:20001b Zbl 0514.20002

[Isaacs 1976] I. M. Isaacs, *Character theory of finite groups*, Pure and Applied Mathematics **69**, Academic Press, New York, 1976. MR 57 #417 Zbl 0337.20005

[Isaacs 1989] I. M. Isaacs, "Coprime group actions fixing all nonlinear irreducible characters", *Canad. J. Math.* **41**:1 (1989), 68–82. MR 90j:20038 Zbl 0686.20002

[Isaacs 2011] I. M. Isaacs, "Bounding the order of a group with a large character degree", *J. Algebra* **348** (2011), 264–275. MR 2012j:20027 Zbl 1242.20011

[Isaacs et al. 2009] I. M. Isaacs, A. Moretó, G. Navarro, and P. H. Tiep, "Groups with just one character degree divisible by a given prime", *Trans. Amer. Math. Soc.* **361**:12 (2009), 6521–6547. MR 2010i:20008 Zbl 1203.20005

[Jansen et al. 1995] C. Jansen, K. Lux, R. Parker, and R. Wilson, *An atlas of Brauer characters*, London Mathematical Society Monographs. New Series **11**, Oxford University Press, New York, 1995. MR 96k:20016 Zbl 0831.20001

[Kostrikin and Tiep 1994] A. I. Kostrikin and P. H. Tiep, *Orthogonal decompositions and integral lattices*, Expositions in Mathematics **15**, de Gruyter, Berlin, 1994. MR 96f:17001 Zbl 0855.11033

[Lewis 2001] M. L. Lewis, "Bounding Fitting heights of character degree graphs", *J. Algebra* **242**:2 (2001), 810–818. MR 2003d:20010 Zbl 0989.20007

[Lewis and White 2003] M. L. Lewis and D. L. White, "Connectedness of degree graphs of nonsolvable groups", *J. Algebra* **266**:1 (2003), 51–76. MR 2004g:20008

[Lewis et al. 2005] M. L. Lewis, A. Moretó, and T. R. Wolf, "Non-divisibility among character degrees", *J. Group Theory* **8**:5 (2005), 561–588. MR 2006j:20012 Zbl 1097.20011

[Liebeck 1987] M. W. Liebeck, "The affine permutation groups of rank three", *Proc. London Math. Soc.* (3) **54**:3 (1987), 477–516. MR 88m:20004 Zbl 0621.20001

[Manz and Wolf 1993] O. Manz and T. R. Wolf, *Representations of solvable groups*, London Mathematical Society Lecture Note Series **185**, Cambridge University Press, 1993. MR 95c:20013 Zbl 0928.20008

[Manz et al. 1988] O. Manz, R. Staszewski, and W. Willems, "On the number of components of a graph related to character degrees", *Proc. Amer. Math. Soc.* **103**:1 (1988), 31–37. MR 89b:20030 Zbl 0645.20005

[McKay 1980] J. McKay, "Graphs, singularities, and finite groups", pp. 183–186 in *The Santa Cruz Conference on Finite Groups* (Santa Cruz, CA, 1979), edited by B. Cooperstein and G. Mason, Proc. Sympos. Pure Math. **37**, Amer. Math. Soc., Providence, R.I., 1980. MR 82e:20014 Zbl 0451.05026

[Navarro and Tiep 2008] G. Navarro and P. H. Tiep, "Rational irreducible characters and rational conjugacy classes in finite groups", *Trans. Amer. Math. Soc.* **360**:5 (2008), 2443–2465. MR 2008k:20014 Zbl 1137.20009

[Noritzsch 1995] T. Noritzsch, "Groups having three complex irreducible character degrees", *J. Algebra* **175**:3 (1995), 767–798. MR 96d:20010 Zbl 0839.20014

[Passman 1968] D. Passman, *Permutation groups*, W. A. Benjamin, New York and Amsterdam, 1968. MR 38 #5908 Zbl 0179.04405

[Sah 1977] C. H. Sah, "Cohomology of split group extensions, II", *J. Algebra* **45**:1 (1977), 17–68. MR 57 #3217 Zbl 0352.18021

[Seitz 1968] G. Seitz, "Finite groups having only one irreducible representation of degree greater than one", *Proc. Amer. Math. Soc.* **19** (1968), 459–461. MR 36 #5212 Zbl 0244.20010

[Steinberg 1985] R. Steinberg, "Finite subgroups of SU_2, Dynkin diagrams and affine Coxeter elements", *Pacific J. Math.* **118**:2 (1985), 587–598. MR 86g:20016 Zbl 0567.20026

Communicated by David Benson
Received 2013-02-20 Revised 2013-05-06 Accepted 2013-07-11

danielgolds@gmail.com *Center for Communications Research, 4320 Westerra Court, San Diego, CA 92121, United States*

guralnic@usc.edu *Department of Mathematics, University of Southern California, 3620 S. Vermont Ave., Los Angeles, CA 90089, United States*

lewis@math.kent.edu *Department of Mathematical Sciences, Kent State University, Kent, OH 44242, United States*

alexander.moreto@uv.es *Departament dÁlgebra, Facultat de Matemátiques, Universitat de Valéncia, Burjassot, 46100 Valéncia, Spain*

gabriel.navarro@uv.es *Departament dÁlgebra, Facultat de Matemátiques, Universitat de Valéncia, Burjassot, 46100 Valéncia, Spain*

tiep@math.arizona.edu *Department of Mathematics, University of Arizona, 617 N. Santa Rita Ave., P.O. Box 210089, Tucson, AZ 85721-0089, United States*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.429

The homotopy category of injectives

Amnon Neeman

Krause studied the homotopy category $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ of complexes of injectives in a locally noetherian Grothendieck abelian category $\mathscr{A}$. Because $\mathscr{A}$ is assumed locally noetherian, we know that arbitrary direct sums of injectives are injective, and hence, the category $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ has coproducts. It turns out that $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ is compactly generated, and Krause studies the relation between the compact objects in $\mathbf{K}(\operatorname{Inj}\mathscr{A})$, the derived category $\mathbf{D}(\mathscr{A})$, and the category $\mathbf{K}_{\mathrm{ac}}(\operatorname{Inj}\mathscr{A})$ of acyclic objects in $\mathbf{K}(\operatorname{Inj}\mathscr{A})$.

We wish to understand what happens in the nonnoetherian case, and this paper begins the study. We prove that, for an arbitrary Grothendieck abelian category $\mathscr{A}$, the category $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ has coproducts and is μ-compactly generated for some sufficiently large μ.

The existence of coproducts follows easily from a result of Krause: the point is that the natural inclusion of $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ into $\mathbf{K}(\mathscr{A})$ has a left adjoint and the existence of coproducts is a formal corollary. But in order to prove anything about these coproducts, for example the μ-compact generation, we need to have a handle on this adjoint.

Also interesting is the counterexample at the end of the article: we produce a locally noetherian Grothendieck abelian category in which products of acyclic complexes need not be acyclic. It follows that $\mathbf{D}(\mathscr{A})$ is not compactly generated. I believe this is the first known example of such a thing.

This research was partly supported by the Australian Research Council. The author would like to thank an anonymous referee for several suggestions of improvements to the manuscript.
MSC2010: primary 13D09; secondary 08B30.
Keywords: locally presentable categories, well generated categories, homotopy categories of injectives.

Introduction

The starting point of this investigation is the article by Krause [2005], in which he studied the homotopy category $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ of complexes of injectives in a locally noetherian Grothendieck abelian category $\mathscr{A}$. It turns out that $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ is compactly generated, the compact objects being injective resolutions of bounded complexes of noetherian objects. In symbols, we have an equivalence $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})^c \cong \mathbf{D}^b(\mathrm{noeth}\,\mathscr{A})$. We can consider the sequence of functors

$$\mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A}) \xrightarrow{J} \mathbf{K}(\mathrm{Inj}\,\mathscr{A}) \xrightarrow{Q} \mathbf{D}(\mathscr{A}),$$

which expresses $\mathbf{D}(\mathscr{A})$ as the quotient of $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ by the subcategory of acyclics $\mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A}) \subset \mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. It is not hard to prove that this is a localization sequence: the functors J and Q have right adjoints, denoted J_ρ and Q_ρ, respectively. Not so formal is that, as long as $\mathbf{D}(\mathscr{A})$ is compactly generated, the functors J and Q also have left adjoints J_λ and Q_λ turning this into a recollement. If we restrict J_λ and Q_λ to the subcategories of compact objects, then we have functors

$$\mathbf{D}(\mathscr{A})^c \xrightarrow{Q_\lambda} \mathbf{K}(\mathrm{Inj}\,\mathscr{A})^c \xrightarrow{J_\lambda} \mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A})^c,$$

which allow us to identify $\mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A})^c$ as the idempotent completion of the Verdier quotient $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})^c/\mathbf{D}(\mathscr{A})^c$.

In the generality above, where $\mathscr{A}$ is an arbitrary locally noetherian Grothendieck category, we understand the compact objects only in $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$, where we have $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})^c = \mathbf{D}^b(\mathrm{noeth}\,\mathscr{A})$. But in examples, we sometimes also know $\mathbf{D}(\mathscr{A})^c$; for instance, if X is a noetherian, separated scheme and $\mathscr{A}$ is the category of quasicoherent sheaves on X, we know that $\mathbf{D}(\mathscr{A})^c = \mathbf{D}^{\mathrm{perf}}(\mathrm{coh}\,X)$, the category of perfect complexes. In this special case, $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})^c = \mathbf{D}^b(\mathrm{noeth}\,\mathscr{A})$ comes down to $\mathbf{D}^b(\mathrm{coh}\,X)$, the bounded derived category of the coherent sheaves on X. The general theory gives us the sequence of functors

$$\mathbf{D}^{\mathrm{perf}}(\mathrm{coh}\,X) \xrightarrow{Q_\lambda} \mathbf{D}^b(\mathrm{coh}\,X) \xrightarrow{J_\lambda} \mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A})^c,$$

and furthermore, it informs us that this sequence identifies the category $\mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A})^c$ of compact objects in $\mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A})$ with the idempotent completion of

$$\mathbf{D}^b(\mathrm{coh}\,X)/\mathbf{D}^{\mathrm{perf}}(\mathrm{coh}\,X) = \mathbf{D}_{\mathrm{sing}}(X),$$

the singularity category of X.

Jørgensen [2005] studied the analogue where injectives are replaced by projectives. Of course, Grothendieck abelian categories do not in general have enough projectives, so he restricted himself to the case where $\mathscr{A}$ is the category of modules over some ring. Under suitable noetherian hypotheses, he proved an analogue of Krause's theorem: the homotopy category $\mathbf{K}(R\text{-Proj})$ is compactly generated, but strangely enough, the subcategory $\mathbf{K}(R\text{-Proj})^c$ of compact objects in $\mathbf{K}(R\text{-Proj})$ is naturally isomorphic to $\mathbf{D}^b(R^{\mathrm{op}}\text{-mod})^{\mathrm{op}}$, the opposite category of the bounded derived category of finitely presented R^{op}-modules. Krause's theorem, in the special case where $\mathscr{A}$ is the category of R-modules, tells us that the subcategory $\mathbf{K}(R\text{-Inj})^c$ of compact objects in $\mathbf{K}(R\text{-Inj})$ is naturally identified with $\mathbf{D}^b(R\text{-mod})$. If R is a noetherian commutative ring, then both $\mathbf{K}(R\text{-Proj})$ and $\mathbf{K}(R\text{-Inj})$ are compactly generated, but the subcategories of compact objects are naturally the opposite of each other.

Iyengar and Krause [2006] studied this further and proved, among other things, that in the presence of a dualizing complex the categories $\mathbf{K}(R\text{-Proj})$ and $\mathbf{K}(R\text{-Inj})$ are equivalent. More precisely, tensoring with the dualizing complex induces an equivalence. Of course, it must also induce an equivalence on the subcategories of compact objects; that is, it must induce an equivalence

$$\mathbf{D}^b(R^{\mathrm{op}}\text{-mod})^{\mathrm{op}} \to \mathbf{D}^b(R\text{-mod}).$$

This equivalence turns out to be the usual one of Grothendieck duality.

The results raise the obvious question: what is the right generality in which the results hold? Since Grothendieck abelian categories rarely have enough projectives, Jørgensen's results all assumed that he was working over a ring; in other words, they were restricted to the affine case of Grothendieck duality. In [Neeman 2008; 2010], I studied this problem and proved several improvements of Jørgensen's results, and Murfet [2007] carried the project further in his PhD thesis. One striking feature of my results was that much of what Jørgensen proved for $\mathbf{K}(R\text{-Proj})$ was true without the noetherian hypothesis, which raises the question: to what extent is the noetherian hypothesis necessary in Krause's results? On the face of it, the situation looks hopeless unless we assume that the category $\mathscr{A}$ is locally noetherian because without the noetherian hypothesis direct sums of injectives need not be injective. Hence, $\mathbf{K}(R\text{-Inj})$ does not obviously have coproducts in general, and without coproducts, one doesn't have a good notion of compact objects.

The first result of the current article, also found as Example 5 in [Krause 2012, pp. 778–779], addresses this:

Theorem 2.13 and Corollary 2.14. *Let $\mathscr{A}$ be any Grothendieck abelian category not necessarily locally noetherian. Then the inclusion of* $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ *into* $\mathbf{K}(\mathscr{A})$ *has a left adjoint* $I : \mathbf{K}(\mathscr{A}) \to \mathbf{K}(\operatorname{Inj}\mathscr{A})$. *It formally follows that* $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ *has coproducts.*

Remark. Krause's proof is based on the work of Bican, El Bashir, and Enochs [Bican et al. 2001], which means that it works far more generally than the argument we give here, but unfortunately, the adjoint is not very explicit. For the proof of the next theorem, we need to have a handle on this adjoint; hence, we give a different proof.

While I do not understand the situation well enough to say when $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ is compactly generated, I do have the following result:

Theorem 3.13. *There is a regular cardinal μ for which the category $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ is μ-compactly generated* (in the sense of [Neeman 2001, Definition 8.1.6] or [Krause 2001]).

In the algebrogeometric situation, it may well be that the categories $\mathbf{K}(\mathrm{Inj}\,X)$ and $K_m(\mathrm{Proj}\,X)$ can be equivalent even when X is not noetherian. Here $K_m(\mathrm{Proj}\,X)$ is Murfet's mock homotopy category of projectives; for nonaffine schemes, this is the right generalization. Both $\mathbf{K}(\mathrm{Inj}\,X)$ and $K_m(\mathrm{Proj}\,X)$ have coproducts and are μ-compactly generated for μ sufficiently large. In the noetherian case, they are equivalent whenever X has a dualizing complex, and part of the interest of the results is that they might lead to a nonnoetherian generalization.

Recall that, if $\mathscr{A}$ is locally noetherian and $\mathbf{D}(\mathscr{A})$ is compactly generated, then the natural functors

$$\mathbf{K}_{\mathrm{ac}}(\mathrm{Inj}\,\mathscr{A}) \xrightarrow{J} \mathbf{K}(\mathrm{Inj}\,\mathscr{A}) \xrightarrow{Q} \mathbf{D}(\mathscr{A})$$

have right and left adjoints giving a recollement. It turns out that the right adjoints J_ρ and Q_ρ exist much more generally for any Grothendieck abelian category. But the left adjoints don't: we will produce an example of a locally noetherian Grothendieck abelian category $\mathscr{A}$ such that the functor J does not respect products and hence cannot have a left adjoint. It will then follow, from [Krause 2005], that the category $\mathbf{D}(\mathscr{A})$ is not compactly generated.

The article is organized as follows. In Section 1, we recall that any Grothendieck abelian category is locally presentable, meaning there is a generator g and a regular cardinal α so that $\mathrm{Hom}(g,-)$ commutes with α-filtered colimits. We discuss this in some detail because we also want to prove that $\mathrm{Ext}^1(g,-)$ commutes with α-filtered colimits.

In Section 2, we prove Theorem 2.13 and Corollary 2.14; for a given $X \in \mathbf{K}(\mathscr{A})$, we give an explicit construction of $I(X) \in \mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ as a certain colimit. In Section 3, we prove Theorem 3.13, showing the μ-compact generation of $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. The essence of the proof is to study the construction of $I(X)$ more carefully and see what it does to subobjects of X.

Finally, Section 4 contains the counterexample, the locally noetherian Grothendieck abelian category in which products of acyclic complexes of injectives need not be acyclic.

1. Cardinality estimates in Grothendieck abelian categories

Throughout the section, we will assume that $\mathscr{A}$ is a Grothendieck abelian category and $g \in \mathscr{A}$ is a fixed generator.

Lemma 1.1. *Let Y be an object of $\mathscr{A}$. If the cardinality of* $\mathrm{Hom}_{\mathscr{A}}(g, Y)$ *is* $\leq \alpha$, *then Y has no more than 2^α subobjects.*

Proof. We have a map

$$\{\text{subobjects of } Y\} \xrightarrow{\Phi} \{\text{subsets of } \mathrm{Hom}(g, Y)\},$$

which takes a subobject $X \subset Y$ to the subset $\Phi(X) = \mathrm{Hom}(g, X) \subset \mathrm{Hom}(g, Y)$. The map Φ is injective because we can recover X from $\Phi(X)$: the fact that g is a generator allows us to choose an epimorphism $\coprod_\Lambda g \to X$. Then the factorization

$$\coprod_\Lambda g \to \coprod_{\mathrm{Hom}(g,X)} g \to X$$

tells us that X is the image in Y of the natural map $\coprod_{\mathrm{Hom}(g,X)} g \to Y$. □

Construction 1.2. We construct the smallest full subcategory $\mathscr{C} \subset \mathscr{A}$ satisfying:

(i) The generator g belongs to $\mathscr{C}$.

(ii) If X is an object of $\mathscr{C}$, then so are all the subquotients of $g^{\#\mathrm{Hom}(g,X)}$. Here $\#\mathrm{Hom}(g, X)$ stands for the cardinality of $\mathrm{Hom}(g, X)$, and for a cardinal α, we let g^α be the coproduct of α copies of g.

Lemma 1.3. *There is only a set of isomorphism classes of objects of* $\mathscr{C}$.

Proof. We build up $\mathscr{C}$ in countably many steps: we start with $\mathscr{C}_0 = \{g\}$ and then construct $\mathscr{C}_{n+1}$ out of $\mathscr{C}_n$ by throwing in all the subquotients of $g^{\#\mathrm{Hom}(g,X)}$ for all $X \in \mathscr{C}_n$. We let $\mathscr{C}$ be the union of the $\mathscr{C}_n$. □

Definition 1.4. We let $\alpha > \aleph_0$ be a regular cardinal such that (i) $2^{\#\mathrm{Hom}(g,Y)} < \alpha$ for all $Y \in \mathscr{C}$, and (ii) there are $< \alpha$ isomorphism classes of objects in $\mathscr{C}$.

Lemma 1.5. *Let Z be an object of $\mathscr{C}$ and $f : Z' \to Z$ an epimorphism in $\mathscr{A}$. Then there is an object $Y \in \mathscr{C}$ and a morphism $g : Y \to Z'$ so that the composite $Y \to Z' \to Z$ is epi.*

Proof. We may choose an epimorphism $\coprod_\Lambda g \to Z'$ and consider the composite epimorphism

$$\coprod_\Lambda g \to Z' \to Z.$$

Let M be the image of Λ in $\mathrm{Hom}(g, Z)$; we may choose a splitting to the surjection $\Lambda \to M$. The composite

$$\coprod_M g \to \coprod_\Lambda g \to Z' \to Z$$

is an epimorphism from a subquotient of $g^{\#\mathrm{Hom}(g,Z)}$ to Z and factors through Z'. □

Lemma 1.6. *Let $\mathscr{I}$ be a β-filtered category for some regular cardinal β, let $F : \mathscr{I} \to \mathscr{A}$ be a functor, and let $\varphi : F \to Z$ be a natural transformation from F to the constant functor that takes every $i \in \mathscr{I}$ to the object $Z \in \mathscr{A}$ and every morphism to the identity. Assume that the map $\underrightarrow{\operatorname{colim}} F \to Z$ is epi and that Z has fewer than β subobjects. Then there exists some object $i \in \mathscr{I}$ with $Fi \to Z$ epi.*

Proof. Consider the set S of subobjects of Z that are images of $Fi \to Z$ for some $i \in \mathscr{I}$. For each $X \in S$, choose an object $\rho(X) \in \mathscr{I}$ so that the image of $F(\rho(X)) \to Z$ is X. There are fewer than β such $\rho(X)$, and hence, we may choose an object $j \in \mathscr{I}$ and, for each X, a morphism $\rho(X) \to j$. Then $\operatorname{Im}(Fj \to Z)$ belongs to S and contains all the other elements of S as subobjects. The epimorphism $\underrightarrow{\operatorname{colim}}\, Fi \to Z$ factors through $\operatorname{Im}(Fj \to Z)$, and hence, $Fj \to Z$ is epi. $\square$

Lemma 1.7. *Every object of $\mathscr{C}$ is α-presentable in $\mathscr{A}$.*

Proof. Let Z be an object of $\mathscr{C}$; we need to show that $\operatorname{Hom}(Z, -)$ commutes with α-filtered colimits in $\mathscr{A}$. Let $\mathscr{I}$ be an α-filtered small category, and let $F : \mathscr{I} \to \mathscr{A}$ be a functor. We need to show that the natural map

$$\Phi : \underrightarrow{\operatorname{colim}}_{i \in \mathscr{I}} \operatorname{Hom}(Z, Fi) \to \operatorname{Hom}\big(Z, \underrightarrow{\operatorname{colim}}_{i \in \mathscr{I}} Fi\big)$$

is an isomorphism. We will prove that Φ is surjective and injective.

Let us prove the injectivity of Φ first. An element in the kernel of Φ may be represented by a morphism $\theta : Z \to Fi$ so that the composite

$$Z \to Fi \to \underrightarrow{\operatorname{colim}}_{i \in I} Fi$$

vanishes. We need to show that for some $\rho : i \to j$ the composite $Z \to Fi \xrightarrow{F\rho} Fj$ vanishes. Let $i/\mathscr{I}$ be the category whose objects are maps $\rho : i \to j$ in $\mathscr{I}$ and whose morphisms are commutative triangles

$$\begin{tikzcd}
 & j \arrow[d] \\
i \arrow[ur] \arrow[r] & j'
\end{tikzcd}$$

We have an exact sequence of functors on $i/\mathscr{I}$ that takes each object $\rho : i \to j$ to

$$0 \to \operatorname{Ker}(Z \to Fj) \to Z \to Fj.$$

The category $i/\mathscr{I}$ is filtered, and hence, the colimit in the Grothendieck category $\mathscr{A}$ is the exact sequence

$$0 \to \underrightarrow{\operatorname{colim}}_{i/\mathscr{I}} \operatorname{Ker}(Z \to Fj) \to Z \to \underrightarrow{\operatorname{colim}}_{\mathscr{I}} Fj.$$

The fact that the map $Z \to \underrightarrow{\operatorname{colim}}_{\mathscr{I}} Fj$ vanishes means that Z must be the colimit of its subobjects $\operatorname{Ker}(Z \to Fj)$ over the α-filtered category $i/\mathscr{I}$. But the cardinal α

was chosen to be larger than $2^{\# \mathrm{Hom}(g,Z)}$, and Lemma 1.1 tells us that $Z \in \mathscr{C}$ has fewer than α subobjects. By Lemma 1.6, there is some object $\rho : i \to j$ in $\mathscr{I}$ with $\mathrm{Ker}(Z \to Fj) = Z$.

Next we prove the surjectivity of Φ. Let $L = \varinjlim_{i \in I} Fi$; suppose we take an element of $\mathrm{Hom}(Z, L)$, that is, a map $\varphi : Z \to L$. For each i, we form the pullback square

$$\begin{array}{ccc} Gi & \longrightarrow & Z \\ \big\downarrow & & \big\downarrow{\scriptstyle \varphi} \\ Fi & \longrightarrow & L \end{array}$$

Then the Gi extend to a functor $G : \mathscr{I} \to \mathscr{A}$. Taking the colimit over the filtered category $\mathscr{I}$, we obtain a pullback square

$$\begin{array}{ccc} \varinjlim Gi & \longrightarrow & Z \\ \big\downarrow & & \big\downarrow{\scriptstyle \varphi} \\ L & \longrightarrow & L \end{array}$$

from which we conclude that the map $\varinjlim Gi \to Z$ is epi (actually, it's even an isomorphism). But $\mathscr{I}$ is α-filtered, and Lemma 1.1 tells us that $Z \in \mathscr{C}$ has fewer than α subobjects. By Lemma 1.6, there is an object $i \in \mathscr{I}$ so that $Gi \to Z$ is epi.

By Lemma 1.5, we may choose an object $Y \in \mathscr{C}$ and a morphism $Y \to Gi$ so that the composite $Y \to Gi \to Z$ is epi. Let X be the kernel of the epimorphism $Y \to Z$; because X is a subobject of $Y \in \mathscr{C}$, it lies in $\mathscr{C}$ and the composite $X \to Fi \to L$ vanishes. By the injectivity of Φ, there must be some $\rho : i \to j$ in $\mathscr{I}$ so that the composite $X \to Fi \xrightarrow{F(\rho)} Fj$ vanishes. But the vanishing of $X \to Y \to Fi \to Fj$ means that the map $Y \to Fj$ factors through $Y \to Z \to Fj$. We have found a $Z \to Fj$ that maps under Φ to $\varphi : Z \to L = \varinjlim_{i \in I} Fi$. □

Lemma 1.8. *Let X be an object of $\mathscr{C}$ and Z an object of $\mathscr{A}$. For every element $z \in \mathrm{Ext}^n_{\mathscr{A}}(X, Z)$, there exists an object $Y \in \mathscr{C}$, a morphism $Y \to Z$, and an element $y \in \mathrm{Ext}^n_{\mathscr{C}}(X, Y)$ so that y maps to z under the natural map.*

Proof. The case $n = 0$ is trivial; we may take y to be the identity. Suppose therefore $n > 0$; then z is represented by an extension

$$0 \to Z \to \cdots \to W \to X \to 0.$$

By Lemma 1.5, we may choose an object $W' \in \mathscr{C}$ and a morphism $W' \to W$ so that the composite $W' \to W \to X$ is epi. If X' is the kernel of $W' \to X$, then X' is a subobject of the object $W' \in \mathscr{C}$ and hence belongs to $\mathscr{C}$. And the extension $z \in \mathrm{Ext}^n_{\mathscr{A}}(X, Z)$ is equivalent to the concatenation of $0 \to X' \to W' \to X \to 0$ in $\mathrm{Ext}^1_{\mathscr{C}}(X, X')$ with an extension $z' \in \mathrm{Ext}^{n-1}_{\mathscr{A}}(X', Z)$. Induction on n now gives the result. □

Corollary 1.9. *The category $\mathscr{C} \subset \mathscr{A}$ is closed under extensions.*

Proof. Let $0 \to Z \to Y \to X \to 0$ be an extension with $X, Z \in \mathscr{C}$ and $Y \in \mathscr{A}$. By Lemma 1.8, there is an extension $0 \to A \to B \to X \to 0$ in $\mathscr{C}$ and a map $A \to Z$ connecting the extensions. Thus, Y must be the pushout in the diagram

$$\begin{tikzcd} A \arrow[r] \arrow[d] & B \arrow[d] \\ Z \arrow[r] & Y \end{tikzcd}$$

and hence, Y is a quotient of $B \oplus Z \in \mathscr{C}$ and must belong to $\mathscr{C}$. $\square$

Lemma 1.10. *For every object $X \in \mathscr{C}$ and every $n \geq 0$, the functor* $\operatorname{Ext}^n(X, -)$ *commutes with α-filtered colimits.*

Proof. The case $n = 0$ was proved in Lemma 1.7, and we will now prove the general case. Let $\mathscr{I}$ be an α-filtered category, and let $F : \mathscr{I} \to \mathscr{A}$ be a functor. We need to show that the natural map

$$\Phi : \underrightarrow{\operatorname{colim}}_{i \in \mathscr{I}} \operatorname{Ext}^n(X, Fi) \to \operatorname{Ext}^n\big(X, \underrightarrow{\operatorname{colim}}_{i \in \mathscr{I}} Fi\big)$$

is an isomorphism. We will prove that Φ is surjective and injective.

We prove surjectivity first. Given an element

$$z \in \operatorname{Ext}^n(X, \underrightarrow{\operatorname{colim}}_{i \in \mathscr{I}} Fi),$$

there is by Lemma 1.8 an object $Y \in \mathscr{C}$, an extension $y \in \operatorname{Ext}^n(X, Y)$, and a morphism $Y \to \underrightarrow{\operatorname{colim}}_{i \in \mathscr{I}} Fi$ taking y to z. By Lemma 1.7, the map from $Y \in \mathscr{C}$ to the α-filtered colimit factors through some Fi, and the surjectivity of Φ follows.

Next we prove injectivity by induction on n. Let

$$L = \underrightarrow{\operatorname{colim}}_{i \in \mathscr{I}} Fi,$$

and suppose we are given a $\theta \in \operatorname{Ext}^n(X, Fi)$ that maps to zero in $\operatorname{Ext}^n(X, L)$. In the proof of Lemma 1.8, we produced objects $W', X' \in \mathscr{C}$, an exact sequence $0 \to X' \to W' \to X \to 0$ in $\mathscr{C}$, and an element $\theta' \in \operatorname{Ext}^{n-1}(X', Fi)$ mapping to $\theta \in \operatorname{Ext}^n(X, Fi)$. In the commutative diagram with exact rows

$$\begin{tikzcd} \underrightarrow{\operatorname{colim}} \operatorname{Ext}^{n-1}(W', Fj) \arrow[r] \arrow[d, "\Phi_{W'}"'] & \underrightarrow{\operatorname{colim}} \operatorname{Ext}^{n-1}(X', Fj) \arrow[r, "\pi"] \arrow[d, "\Phi_{X'}"'] & \underrightarrow{\operatorname{colim}} \operatorname{Ext}^n(X, Fj) \arrow[d, "\Phi_X"'] \\ \operatorname{Ext}^{n-1}(W', L) \arrow[r] & \operatorname{Ext}^{n-1}(X', L) \arrow[r] & \operatorname{Ext}^n(X, L) \end{tikzcd}$$

we have $\theta \in \underrightarrow{\operatorname{colim}} \operatorname{Ext}^n(X, Fj)$ and $\theta' \in \underrightarrow{\operatorname{colim}} \operatorname{Ext}^{n-1}(X', Fj)$ satisfying $\pi(\theta') = \theta$ and $\Phi_X(\theta) = 0$. By induction on n, we know that the maps $\Phi_{X'}$ and $\Phi_{W'}$ are isomorphisms. A short diagram chase establishes that θ vanishes as an element of $\underrightarrow{\operatorname{colim}} \operatorname{Ext}^n(X, Fj)$. $\square$

Definition 1.11. Choose a regular cardinal μ so that:

(i) For each object $X \in \mathscr{C}$ and for each quotient Y of g^α, where α is as in Definition 1.4, we have $\#\operatorname{Hom}(X, Y) < \mu$.

(ii) The object g^α has $< \mu$ quotients.

(iii) The product of $\leq \alpha$ cardinals, all less than μ, is less than μ.

Remark 1.12. If we let β be any cardinal $\geq \alpha$ satisfying parts (i) and (ii) of Definition 1.11, then the successor of 2^β satisfies all three hypotheses on μ. The product of $\leq \alpha$ cardinals, all less than μ, is the product of $\leq \alpha$ cardinals all $\leq 2^\beta$ and is bounded above by $(2^\beta)^\alpha = 2^{\beta \times \alpha} = 2^\beta$.

Definition 1.13. Let μ be as in Definition 1.11. We define $\mathscr{B} \subset \mathscr{A}$ to be the full subcategory of $\mathscr{A}$ consisting of the objects Y with $\#\operatorname{Hom}(g, Y) < \mu$.

Lemma 1.14. *An object Y belongs to $\mathscr{B}$ if and only if it is the quotient to g^λ for some $\lambda < \mu$. And if Y belongs to $\mathscr{B}$, then $\#\operatorname{Hom}(X, Y) < \mu$ for all $X \in \mathscr{C}$.*

Proof. Every Y is the quotient of $g^{\#\operatorname{Hom}(g,Y)}$. If Y happens to belong to $\mathscr{B}$, then $\lambda = \#\operatorname{Hom}(g, Y) < \mu$, so Y is a quotient as specified. We need to prove the converse: any quotient of g^λ, $\lambda < \mu$, belongs to $\mathscr{B}$. Let Y be such a quotient; to show that Y belongs to $\mathscr{B}$, it suffices to prove that $\#\operatorname{Hom}(g, Y) < \mu$, but we actually want the refinement that $\#\operatorname{Hom}(X, Y) < \mu$ for every $X \in \mathscr{C}$.

Choose an epimorphism $g^\lambda \to Y$ and $X \in \mathscr{C}$. We note that Y is the α-filtered colimit of the subobjects Fi, $i \in \mathscr{I}$, where Fi is the image in Y of the map from a summand $g^{\beta_i} \subset g^\lambda$ with $\beta_i < \alpha$. By Lemma 1.7, every map $X \to Y$ factors as $X \to Fi \to Y$ for some $i \in \mathscr{I}$. But $\#\operatorname{Hom}(X, Fi) < \mu$ by Definition 1.11(i), and there are fewer than $\lambda^\alpha < \mu$ different inclusions $g^{\beta_i} \subset g^\lambda$. Therefore, there are fewer than μ factorizations $X \to Fi \to Y$. □

Proposition 1.15. *The category $\mathscr{B}$ satisfies the following properties*:

(i) *The coproduct of $< \mu$ objects in $\mathscr{B}$ lies in $\mathscr{B}$.*

(ii) *Any subquotient of an object in $\mathscr{B}$ belongs to $\mathscr{B}$.*

(iii) *Any extension of objects in $\mathscr{B}$ lies in $\mathscr{B}$.*

(iv) *$\mathscr{C}$ is contained in $\mathscr{B}$.*

(v) *For all $X \in \mathscr{C}$, all $Z \in \mathscr{B}$, and all integers $n \geq 0$, we have $\#\operatorname{Ext}^n(X, Z) < \mu$.*

Proof. (i) Let $\{Y_i \mid i \in \mathscr{I}\}$ be a set of $< \mu$ objects in $\mathscr{B}$. Each Y_i is the quotient of $g^{\#\operatorname{Hom}(g,Y_i)}$, and hence, the coproduct of the Y_i is a quotient of $g^{\sum_i \#\operatorname{Hom}(g,Y_i)}$. And $\sum_i \#\operatorname{Hom}(g, Y_i)$ is the sum of fewer than μ cardinals, each $< \mu$, and hence is $< \mu$.

(ii) If Y belongs to $\mathscr{B}$, it is a quotient of g^λ, $\lambda < \mu$, and hence, so is any of its quotients. Also $\#\operatorname{Hom}(g, Y) < \mu$, and for any subobject X, we have $\#\operatorname{Hom}(g, X) \leq \#\operatorname{Hom}(g, Y) < \mu$.

(iii) Given an extension $0 \to X \to Y \to Z \to 0$, we have an exact sequence $0 \to \mathrm{Hom}(g, X) \to \mathrm{Hom}(g, Y) \to \mathrm{Hom}(g, Z)$. If X and Z belong to $\mathscr{B}$, then $\#\mathrm{Hom}(g, X) < \mu$ and $\#\mathrm{Hom}(g, Z) < \mu$, and $\#\mathrm{Hom}(g, Y)$ is the sum of fewer than μ cardinals all smaller than μ.

(iv) Suppose $Y \in \mathscr{C}$. By the definition of α, we have $\#\mathrm{Hom}(g, Y) < \alpha$, and μ was chosen larger than α.

(v) There are fewer than α objects in $\mathscr{C}$, and for any pair of objects $X, Y \in \mathscr{C}$, we have that $\#\mathrm{Hom}(X, Y) < \mu$. Hence, there are fewer than μ sequences in $\mathscr{C}$ of length n

$$0 \to Y \to W_1 \to \cdots \to W_n \to X \to 0.$$

Some of these sequences will be exact, and up to equivalence, they define fewer than μ elements in groups $\mathrm{Ext}^n_{\mathscr{C}}(X, Y)$. By Lemma 1.8, we have that any element of $\mathrm{Ext}^n(X, Z)$ is the image under some map in $\mathrm{Hom}(Y, Z)$ of an element in $\mathrm{Ext}^n_{\mathscr{C}}(X, Y)$. But $\#\mathrm{Hom}(Y, Z) < \mu$ by Lemma 1.14; hence, there are fewer than μ ways of picking a triple

$$Y \in \mathscr{C}, \quad y \in \mathrm{Ext}^n_{\mathscr{C}}(X, Y), \quad \text{and} \quad f \in \mathrm{Hom}(Y, Z). \qquad \square$$

Remark 1.16. If $F : \mathscr{I} \to \mathscr{B}$ is a functor and $\mathscr{I}$ has fewer than μ objects, then the colimit of F belongs to $\mathscr{B}$. This is because the colimit is a quotient of the coproduct of Fi over all objects $i \in \mathscr{I}$. The coproduct belongs to $\mathscr{B}$ by Proposition 1.15(ii) and its quotient the colimit by Proposition 1.15(iii).

Remark 1.17. Let q be the coproduct of all the quotients of the generator g. Then for every $n \geq 0$ and any object $Z \in \mathscr{B}$, we have $\#\mathrm{Ext}^n(q, Z) < \mu$. The reason is the following. Write $q = \coprod_M x_m$ to express q as the coproduct of all the quotients x_m of g. Then

$$\mathrm{Ext}^n(q, Z) = \prod_M \mathrm{Ext}^n(x_m, Z)$$

with each $x_m \in \mathscr{C}$ and $Z \in \mathscr{B}$. By Proposition 1.15(v), we have that $\#\mathrm{Ext}^n(x_m, Z) < \mu$ for each x_m, and there are fewer than α objects $x_m \in \mathscr{C}$. Definition 1.11(iii) guarantees that $\#\mathrm{Ext}^n(q, Z) < \mu$.

Proposition 1.18. *The category $\mathscr{B}$ is precisely the full subcategory of μ-presentable objects of $\mathscr{A}$.*

Proof. Let Y be a μ-presentable object in $\mathscr{A}$, and let $g^\gamma \to Y$ be an epimorphism. Then Y is the μ-filtered colimit of all the subobjects $\{Fi \mid i \in \mathscr{I}\}$, where Fi is the image in Y of some summand $g^{\lambda_i} \subset g^\gamma$ with $\lambda_i < \mu$. The identity map $Y \to Y$ is a map from the μ-presentable object Y to the μ-filtered colimit $Y = \varinjlim Fi$ and hence must factor through some Fi. But then the map $g^{\lambda_i} \to Fi \to Y$ is the composite of two epimorphisms and is epi, and $\lambda_i < \mu$. Thus, $Y \in \mathscr{B}$.

Next we must prove that every $Y \in \mathcal{B}$ is μ-presentable. Choose an epimorphism $g^\lambda \to Y$ with $\lambda < \mu$; its kernel K is a subobject of a coproduct of fewer than μ copies of $g \in \mathcal{B}$ and hence belongs to $\mathcal{B}$. Thus, we may choose an epimorphism $g^\nu \to K$ with $\nu < \mu$. We therefore have a presentation of Y

$$g^\nu \to g^\lambda \to Y \to 0$$

with $\lambda, \nu < \mu$. But g is α-presentable by Lemma 1.7 and hence also μ-presentable for our choice of $\mu > \alpha$. The μ-presentability of Y follows from the presentation. □

An immediate corollary is:

Corollary 1.19. *Any Grothendieck abelian category is locally presentable* (in the sense of [Gabriel and Ulmer 1971]).

Definition 1.20. Let ν be an infinite cardinal. We define $\mathcal{A}^\nu \subset \mathcal{A}$ to be the full subcategory of all ν-presentable objects.

Remark 1.21. In Definition 1.11, we chose a regular cardinal μ, and in Proposition 1.18, we saw that $\mathcal{A}^\mu = \mathcal{B}$ with $\mathcal{B}$ as in Definition 1.13. As it happens, we chose μ sufficiently large so that $\mathcal{A}^\mu$ satisfies all the nice properties of Lemma 1.14, Proposition 1.15, and Remark 1.16.

2. An adjoint to the inclusion $\mathbf{K}(\mathrm{Inj}\,\mathcal{A}) \hookrightarrow \mathbf{K}(\mathcal{A})$

Let $\mathcal{A}$ be a Grothendieck abelian category, let $\mathbf{K}(\mathcal{A})$ be the homotopy category of chain complexes in $\mathcal{A}$, and let $\mathbf{K}(\mathrm{Inj}\,\mathcal{A})$ be the full subcategory whose objects are the chain complexes of injectives. There is an obvious inclusion $\mathbf{K}(\mathrm{Inj}\,\mathcal{A}) \hookrightarrow \mathbf{K}(\mathcal{A})$, and in this section, we will study its left adjoint.

We begin with some preliminaries.

Lemma 2.1. *Every bounded-below, acyclic complex lies in* ${}^\perp\mathbf{K}(\mathrm{Inj}\,\mathcal{A})$. *We remind the reader: this means that any chain map $A \to E$ from a bounded-below, acyclic complex to a complex of injectives is null-homotopic.*

Proof. If A vanishes in degrees $< n$, then the map $A \to E$ factors through the brutal truncation $\beta^{\geq n} E$; the factorization is the obvious

$$\begin{tikzcd}
\cdots \arrow[r] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & A^n \arrow[r] \arrow[d] & A^{n+1} \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & E^n \arrow[r] \arrow[d] & E^{n+1} \arrow[r] \arrow[d] & \cdots \\
\cdots \arrow[r] & E^{n-2} \arrow[r] & E^{n-1} \arrow[r] & E^n \arrow[r] & E^{n+1} \arrow[r] & \cdots
\end{tikzcd}$$

The map $A \to \beta^{\geq n} E$ is a chain map from an acyclic complex to a bounded-below complex of injectives and hence null-homotopic. □

The converse is not true: objects of ${}^{\perp}\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ do not have to be bounded below. But they do have to be acyclic.

Lemma 2.2. *Every object of* ${}^{\perp}\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ *is acyclic.*

Proof. Let E be an injective cogenerator of the abelian category $\mathscr{A}$. Then $\mathrm{Hom}(X, \Sigma^{-n}E) = 0$ if and only if $H^n(X) = 0$. □

Lemma 2.3. *Let* $f : X \to Y$ *be quasi-isomorphism of chain complexes, and assume* $f^i : X^i \to Y^i$ *is an isomorphism for all* $i \ll 0$. *Then, for any chain complex of injectives* E, *the natural map*

$$\mathrm{Hom}_{\mathbf{K}(\mathscr{A})}(Y, E) \xrightarrow{\mathrm{Hom}(f,E)} \mathrm{Hom}_{\mathbf{K}(\mathscr{A})}(X, E)$$

is an isomorphism.

Proof. In the triangle $A \to X \xrightarrow{f} Y \to \Sigma A$, we have that f is a quasi-isomorphism and hence A is acyclic. Furthermore, the fact that $f^i : X^i \to Y^i$ is an isomorphism for $i \ll 0$ means that A is homotopy-equivalent to a bounded-below complex, and Lemma 2.1 tells us that $A \in {}^{\perp}\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. The result now follows by applying $\mathrm{Hom}(-, E)$ to the triangle. □

Remark 2.4. In Lemma 2.3, we saw that any chain map $X \to E$ factors up to homotopy through $X \to Y$. We wish to consider the factorizations not only up to homotopy, and hence, we will work in $\mathbf{C}(\mathscr{A})$, the category of chain complexes in $\mathscr{A}$ where the morphisms are genuine chain maps, not homotopy equivalence classes.

Lemma 2.5. *Let* $f : X \to Y$ *be a morphism in* $\mathbf{C}(\mathscr{A})$ *whose mapping cone lies in* ${}^{\perp}\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. *Suppose further that in each degree* i *the map* $f^i : X^i \to Y^i$ *is a monomorphism. Let* E *be a complex of injectives*; *then the map* $\mathrm{Hom}(f, E) : \mathrm{Hom}_{\mathbf{C}(\mathscr{A})}(Y, E) \to \mathrm{Hom}_{\mathbf{C}(\mathscr{A})}(X, E)$ *is surjective. In other words, every chain map* $X \to E$ *factors through* $f : X \to Y$, *not only up to homotopy but in the category* $\mathbf{C}(\mathscr{A})$.

Proof. Suppose we are given a chain map $h : X \to E$. By Lemma 2.3, it factors up to homotopy through $f : X \to Y$, meaning there exists a $g : Y \to E$ so that h is homotopic to gf. Let $\Theta : X \Rightarrow E$ be a homotopy that works; then for every $i \in \mathbb{Z}$, we have a map $\Theta^i : X^i \to E^{i-1}$ with

$$h - gf = \Theta\partial + \partial\Theta.$$

But now $\Theta^i : X^i \to E^{i-1}$ is a morphism in $\mathscr{A}$ from X^i to the injective object E^{i-1}, and we may therefore factor it through the monomorphism $f^i : X^i \to Y^i$. Thus, we construct maps $\Phi^i : Y^i \to E^{i-1}$ with $\Phi^i f^i = \Theta^i$. If we let

$$g' = g + \Phi\partial + \partial\Phi,$$

then $h = g'f$. □

In Lemma 2.5, we showed the existence of a factorization $h = gf$. Next we worry about uniqueness.

Lemma 2.6. *Suppose $f : X \to Y$ and $h : X \to E$ are as in Lemma 2.5. Assume $g, g' :\to E$ are two morphisms with gf and $g'f$ homotopic to h, and let $\Theta : X \Rightarrow E$ be a homotopy with*

$$gf - g'f = \Theta\partial + \partial\Theta.$$

Then there exists a homotopy $\Phi : Y \Rightarrow E$ with $\Phi f = \Theta$ and so that

$$g - g' = \Phi\partial + \partial\Phi.$$

Proof. Note that Lemma 2.3 guarantees that g is homotopic to g'; the content of what we are about to prove is that the homotopy connecting them may be chosen to lift any given homotopy of gf with $g'f$.

Let us therefore choose any homotopy Φ' connecting g with g'. Then $\Phi' f$ is a homotopy connecting gf with $g'f$ as is Θ; it follows that $\Theta - \Phi' f$ is a chain map $X \to \Sigma E$. By Lemma 2.5, it has a factorization $\Theta - \Phi' f = \rho f$ with $\rho : Y \to \Sigma E$ a chain map. But then $\Phi = \Phi' + \rho$ is a homotopy of g with g', and $\Theta = \Phi f$. $\square$

Definition 2.7. Let λ be an ordinal and $\mathcal{K}$ a category. *A sequence of length λ in $\mathcal{K}$* is the following data:

(i) for every ordinal $i \leq \lambda$ an object $X_i \in \mathcal{K}$ and

(ii) for every pair of ordinals i and j with $i < j \leq \lambda$ a morphism $f_{ij} : X_i \to X_j$.

(iii) If $i < j < k \leq \lambda$, then the composite $X_i \xrightarrow{f_{ij}} X_j \xrightarrow{f_{jk}} X_k$ agrees with $f_{ik} : X_i \to X_k$.

Lemma 2.8. *Suppose X is a sequence of length λ in $\mathbf{C}(\mathcal{A})$, and assume that for every limit ordinal j we have*

$$X_j = \underrightarrow{\operatorname{colim}}_{i<j} X_i.$$

Suppose further that the mapping cone on every $X_i \to X_{i+1}$ belongs to ${}^{\perp}\mathbf{K}(\operatorname{Inj}\mathcal{A})$ and that each of the maps $X_i \to X_{i+1}$ is a degreewise monomorphism. Then the mapping cones of all $f_{ij} : X_i \to X_j$ belong to ${}^{\perp}\mathbf{K}(\operatorname{Inj}\mathcal{A})$.

Proof. We prove, by induction on $k \leq \lambda$, that the statement is true for all f_{ij} with $i \leq j \leq k$. If $k = 0$, there is nothing to prove.

Suppose the statement is true for k; we wish to prove it for $k+1$. Choose any $i < j \leq k+1$. If $i < j \leq k$, then the mapping cone on f_{ij} lies in ${}^{\perp}\mathbf{K}(\operatorname{Inj}\mathcal{A})$ by the inductive hypothesis. If $j = k+1$, then $i \leq k$ and f_{ij} can be written as the composite $f_{k,k+1} f_{ik}$. Since the mapping cones on f_{ik} and on $f_{k,k+1}$ both lie in ${}^{\perp}\mathbf{K}(\operatorname{Inj}\mathcal{A})$, so does the mapping cone on the composite $f_{ij} = f_{k,k+1} f_{ik}$.

Next suppose k is a limit ordinal and the mapping cone on f_{ij} lies in ${}^{\perp}\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ for all $i < j < k$. We need to show that the mapping cone on f_{ik} lies in ${}^{\perp}\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ for every $i < k$. Equivalently, we must prove that the induced map

$$\mathrm{Hom}(f, E) : \mathrm{Hom}_{\mathbf{K}(\mathscr{A})}(X_k, E) \to \mathrm{Hom}_{\mathbf{K}(\mathscr{A})}(X_i, E)$$

is an isomorphism for every $E \in \mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. Let us first prove the surjectivity.

Suppose we are given a chain map $h_i : X_i \to E$. By induction on j, we will factor h_i in $\mathbf{C}(\mathscr{A})$ as $X_i \xrightarrow{f_{ij}} X_j \xrightarrow{h_j} E$. If we have produced the factorization through h_j, then Lemma 2.5 permits us to factor $h_j : X_j \to E$ as $X_j \xrightarrow{f_{j,j+1}} X_{j+1} \xrightarrow{h_{j+1}} E$. For limit ordinals ν, we use the fact that $X_\nu = \underrightarrow{\mathrm{colim}}_{i<\nu}\, X_i$ to extend the factorization. This finishes the induction, and we have a factorization of h_i as $X_i \xrightarrow{f_{ik}} X_k \xrightarrow{h_k} E$. This factorization is in the category $\mathbf{C}(\mathscr{A})$, which is more than we need. It certainly reduces to a factorization in $\mathbf{K}(\mathscr{A})$.

Now we prove injectivity. Suppose we are given a chain map $h_i : X_i \to E$; we wish to prove that the factorization through $f_{ik} : X_i \to X_k$ is unique in $\mathbf{K}(\mathscr{A})$. Choose an $h_k : X_k \to E$ where the identity $h_k f_{ik} = h_i$ holds in $\mathbf{C}(\mathscr{A})$; the existence of such an h_k has just been proved. Now take any $h : X_k \to E$ with $hf_{ik} \cong h_i$, that is, with hf_{ik} homotopic to $h_i = h_k f_{ik}$; we need to prove that h is homotopic to h_k. The proof is by choosing a homotopy Θ_i connecting hf_{ik} with $h_k f_{ik}$ and then by induction on j lifting it to a homotopy connecting hf_{jk} with $h_k f_{jk}$ with $i \le j \le k$ using Lemma 2.6. □

We will construct sequences to which we will apply Lemma 2.8. The maps $X_i \to X_{i+1}$ from which these sequences are built up will be obtained as follows.

Construction 2.9. Given an object $X \in \mathbf{C}(\mathscr{A})$, an integer n, and a monomorphism $X^n \to A$ in $\mathscr{A}$, we form a map of chain complexes $f : X \to Y = B(X, n, X^n \to A)$ as follows:

(i) $f^i : X^i \to Y^i$ is the identity map $1 : X^i \to X^i$ for all $i \neq n, n+1$.

(ii) In degrees n and $n+1$, the commutative square

$$\begin{tikzcd} X^n \arrow[r] \arrow[d] & Y^n \arrow[d] \\ X^{n+1} \arrow[r] & Y^{n+1} \end{tikzcd}$$

is just the pushout square

$$\begin{tikzcd} X^n \arrow[r] \arrow[d] & A \arrow[d] \\ X^{n+1} \arrow[r] & Y^{n+1} \end{tikzcd}$$

We could specify A, up to noncanonical isomorphism, by giving its class as an extension in $\mathrm{Ext}^1(A/X^n, X^n)$. In our applications, A/X^n will be a large coproduct

$A/X^n = q^\beta$ of β copies of the object q of Remark 1.17, and hence, it will suffice to give a subset $\Lambda \subset \mathrm{Ext}^1(q, X^n)$ of cardinality β. We will let $B(X, \Lambda)$ denote the corresponding complex Y.

Remark 2.10. Suppose we are given an integer n and an object $y \in \mathscr{A}$. The trivial complex $T(y)$ is just the complex

$$\cdots \to 0 \to 0 \to y \xrightarrow{1} y \to 0 \to 0 \to \cdots$$

with the nonzero terms in degrees n and $n+1$. Assume now that we are given an object $X \in \mathbf{C}(\mathscr{A})$ and a monomorphism $X^n \to A$ in $\mathscr{A}$. The morphism $f : X \to Y = B(X, n, X^n \to A)$ of Construction 2.9 fits in a short exact sequence of complexes

$$0 \to X \xrightarrow{f} Y \to T(A/X^n) \to 0,$$

and it immediately follows that f is a monomorphism and a quasi-isomorphism. But the mapping cone in $\mathbf{K}(\mathscr{A})$ on the map f is homotopic to a bounded complex and belongs to ${}^\perp\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ by Lemma 2.1. Thus, f is a suitable building block for constructing chains of complexes as in Lemma 2.8.

Construction 2.11. Let $X \in \mathbf{C}(\mathscr{A})$ be an object. Let g be our chosen generator for the abelian category $\mathscr{A}$, and let M be the set of all the quotients of g. In Remark 1.17, we defined q to be the coproduct of them all.

For each subset $\Lambda \subset \mathrm{Ext}^1(q, X^n)$, we consider the map $X \to B(X, \Lambda)$ of Construction 2.9. In the special case where $\Lambda = \mathrm{Ext}^1(q, X^n)$ is maximal, we denote the map as $X \to B(X, n)$. In this case, we know that the functor $\mathrm{Ext}^1(x, -)$ annihilates the map $X^n \to B(X, n)^n$ whenever x is a direct summand of q, in particular for all quotients x of g.

Given $X \in \mathbf{C}(\mathscr{A})$, we inductively define a sequence of length ω in $\mathbf{C}(\mathscr{A})$. At each step, we let the map $X_i \to X_{i+1}$ be $X_i \to B(X_i, n)$ for some suitable n depending on i. The precise recipe is:

(i) $X_0 = X$, and $X_0 \to X_1$ is the map $X \to B(X, 0)$.

(ii) For an integer $i > 0$, we define $X_{2i-1} \to X_{2i}$ to be $X_{2i-1} \to B(X_{2i-1}, i)$ while $X_{2i} \to X_{2i+1}$ is set to be $X_{2i} \to B(X_{2i}, -i)$.

(iii) $X_\omega = \underrightarrow{\mathrm{colim}}\, X_n$.

Lemma 2.12. *Define the map $f_X : X \to J(X)$ to be the morphism $X \to X_\omega$ of Construction 2.11. Then f_X is a degreewise monomorphism and is annihilated degreewise by* $\mathrm{Ext}^1(x, -)$ *whenever x is a quotient of g. Furthermore, the mapping cone of f_X lies in* ${}^\perp\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$.

Proof. By construction, f_X is the colimit of degreewise monomorphisms and hence a degreewise monomorphism. The fact that f_X is annihilated by $\operatorname{Ext}^1(x, -)$ in every degree n is true because, depending on whether n is positive or negative, either the map $X_{2|n|-1} \to X_{2|n|}$ or the map $X_{2|n|} \to X_{2|n|+1}$ will induce zero in degree n under the functors $\operatorname{Ext}^1(x, -)$. That the mapping cone lies in ${}^{\perp}\mathbf{K}(\operatorname{Inj} \mathscr{A})$ comes from Lemma 2.8. $\square$

Theorem 2.13. *The natural inclusion* $\mathbf{K}(\operatorname{Inj} \mathscr{A}) \to \mathbf{K}(\mathscr{A})$ *has a left adjoint* I.

Proof. Let X be an arbitrary object of $\mathbf{C}(\mathscr{A})$. By transfinite induction, we define a chain of complexes $J^{\lambda}(X)$ for every ordinal λ. The rule is:

(i) $J^0(X) = X$.

(ii) If $J^{\lambda}(X)$ has been defined, then the map $J^{\lambda}(X) \to J^{\lambda+1}(X)$ is just $J^{\lambda}(X) \to J(J^{\lambda}(X))$.

(iii) If λ is a limit ordinal, then $J^{\lambda}(X) = \underrightarrow{\operatorname{colim}}_{i<\lambda} J^i(X)$.

Let α be the regular cardinal of Definition 1.4. Now consider the triangle

$$A(X) \to X \to J^{\alpha}(X) \to \Sigma A(X).$$

Lemma 2.8 tells us that $A(X)$ belongs to ${}^{\perp}\mathbf{K}(\operatorname{Inj} \mathscr{A})$. I assert that $J^{\alpha}(X)$ belongs to $\mathbf{K}(\operatorname{Inj} \mathscr{A})$; from the triangle and [Neeman 2001, Theorem 9.1.13], we deduce the existence of the adjoint and note that the adjoint takes X to $I(X) = J^{\alpha}(X)$.

It remains to prove the assertion: we must show that in each degree n the object $J^{\alpha}(X)^n \in \mathscr{A}$ is injective. Since α is an α-filtered colimit of the ordinals $\lambda < \alpha$, Lemma 1.10 tells us that, for each quotient x of the generator g of $\mathscr{A}$,

$$\operatorname{Ext}^1(x, J^{\alpha}(X)^n) = \underrightarrow{\operatorname{colim}}_{\lambda<\alpha} \operatorname{Ext}^1(x, J^{\lambda}(X)^n).$$

By construction, we know that the map

$$\operatorname{Ext}^1(x, J^{\lambda}(X)^n) \to \operatorname{Ext}^1(x, J^{\lambda+1}(X)^n)$$

is zero, and hence, the colimit vanishes. Thus, $\operatorname{Ext}^1(x, J^{\alpha}(X)^n) = 0$ whenever x is a quotient of the generator g, and hence, $J^{\alpha}(X)^n$ must be injective. $\square$

Corollary 2.14. *The homotopy category* $\mathbf{K}(\operatorname{Inj} \mathscr{A})$ *satisfies TR5, meaning it has coproducts.*

Proof. Given a collection of objects $\{X_\lambda \mid \lambda \in \Lambda\}$ in the category $\mathbf{K}(\operatorname{Inj} \mathscr{A})$, we can certainly form the coproduct in $\mathbf{K}(\mathscr{A})$; applying the functor I to this coproduct gives the coproduct in $\mathbf{K}(\operatorname{Inj} \mathscr{A})$. $\square$

Remark 2.15. The construction of $I(X)$ out of X was broken up into two steps. In the proof of Theorem 2.13, we constructed a sequence by letting $J^{i+1}(X) = J(J^i(X))$ for each ordinal i and by taking colimits at limit ordinals. But this hides the fact that $J(Y)$ is constructed out of Y as the colimit of a countable sequence where $Y_{i+1} = B(Y_i, n)$ for some suitable n depending on i; see Construction 2.9. If we assemble it all into one long sequence, then we define a sequence where $X_{i+1} = B(X_i, n)$ for every ordinal i but where the integer n depends on the distance of the ordinal i from its predecessor limit ordinal. And we recover the sequence $\{J^i(X)\}$ by restricting attention to X_i for limit ordinals i.

3. The μ-compact generation of $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$

In Section 2, we proved Theorem 2.13: the inclusion $\mathbf{K}(\mathrm{Inj}\,\mathscr{A}) \to \mathbf{K}(\mathscr{A})$ has a left adjoint I. In the construction, we made many choices: even though we constructed a morphism $X \to I(X) = J^\alpha(X)$ in the category $\mathbf{C}(\mathscr{A})$, the construction is not functorial in $\mathbf{C}(\mathscr{A})$. The map sending X to $I(X)$ becomes a well defined functor only in the homotopy category $\mathbf{K}(\mathscr{A})$, and $X \to I(X)$ is a natural transformation only at the homotopy level. Still X can be expressed as the colimit of all its μ-presentable subobjects with μ as in Definition 1.11, and we would like to express $I(X)$ as a μ-filtered colimit.

Construction 3.1. Recall that $\mathscr{B} \subset \mathscr{A}$ was the category $\mathscr{A}^\mu$ of μ-presentable objects in $\mathscr{A}$; see Proposition 1.18 and Definition 1.20. Let X be an object in $\mathbf{C}(\mathscr{A})$, and let $\mathscr{I}$ be a full subcategory of subobjects $Y \subset X$ with $Y \in \mathbf{C}(\mathscr{B}) \subset \mathbf{C}(\mathscr{A})$. Assume $\mathscr{I}$ is μ-filtered and its colimit is X. Construct the category $\mathscr{J}(\mathscr{I}, n)$ whose objects are subobjects Y of $B(X, n)$ with the following properties:

(i) $Y \cap X$ belongs to $\mathscr{I}$.

(ii) The map $Y \cap X \to Y$ is an isomorphism in degrees $i \neq n, n+1$.

(iii) In degree n, we have a monomorphism $Y^n/Y^n \cap X^n \to B(X,n)^n/X^n$, and from the construction of $B(X,n)$, we know that $B(X,n)^n/X^n$ is the coproduct $q^\beta = \coprod_{\mathrm{Ext}^1(q,X^n)} q$ with q as in Remark 1.17. We require that the monomorphism $Y^n/Y^n \cap X^n \to B(X,n)^n/X^n$ is the inclusion of a subcoproduct.

(iv) The square

$$\begin{tikzcd} Y^n \cap X^n \arrow[r] \arrow[d] & Y^n \arrow[d] \\ Y^{n+1} \cap X^{n+1} \arrow[r] & Y^{n+1} \end{tikzcd}$$

is a pushout.

Remark 3.2. Let us untangle what this means. In Construction 2.11, $B(X,n)$ was defined so that there is a short exact sequence in $\mathbf{C}(\mathscr{A})$

$$0 \to X \to B(X,n) \to T(q)^{\#\mathrm{Ext}(q,X^n)} \to 0$$

where $T(q)$ is the trivial complex $\cdots \to 0 \to q \xrightarrow{1} q \to 0 \to \cdots$ concentrated in degrees n and $n+1$. The conditions on the subobject $Y \subset B(X,n)$ that it must satisfy to belong to $\mathcal{J}(\mathcal{I},n)$ come down to asking that $Y \cap X$ belongs to $\mathcal{I}$ and that in the map of short exact sequences

$$\begin{tikzcd} 0 \arrow[r] & Y \cap X \arrow[r] \arrow[d] & Y \arrow[r] \arrow[d] & Y/(Y\cap X) \arrow[r] \arrow[d, "h"'] & 0 \\ 0 \arrow[r] & X \arrow[r] & B(X,n) \arrow[r] & T(q)^{\#\mathrm{Ext}(q,X^n)} \arrow[r] & 0 \end{tikzcd}$$

the monomorphism $h : Y/(Y\cap X) \to T(q)^{\#\mathrm{Ext}(q,X^n)}$ should be the inclusion of a subcoproduct. In degree n, we have a diagram

$$\begin{tikzcd} 0 \arrow[r] & Y^n \cap X^n \arrow[r] \arrow[d] & Y^n \arrow[r] \arrow[d] & q^{\#\Lambda'} \arrow[r] \arrow[d, "h"'] & 0 \\ 0 \arrow[r] & X^n \arrow[r] & B(X,n)^n \arrow[r] & q^{\#\mathrm{Ext}(q,X^n)} \arrow[r] & 0 \end{tikzcd}$$

The top row of this diagram defines a map $\varphi : \Lambda' \to \mathrm{Ext}^1(q, Y^n \cap X^n)$ giving the extension, and the fact that h is an inclusion means that the composite

$$\Lambda' \xrightarrow{\varphi} \mathrm{Ext}^1(q, Y^n \cap X^n) \to \mathrm{Ext}^1(q, X^n)$$

must be injective. Therefore, φ must be injective; Λ' is a subset of $\mathrm{Ext}^1(q, Y^n \cap X^n)$.

Lemma 3.3. *The objects of the category $\mathcal{J}(\mathcal{I},n)$ all belong to $\mathbf{C}(\mathcal{A}^\mu) = \mathbf{C}(\mathcal{B})$.*

Proof. We know that $Y \cap X$ belongs to $\mathcal{I} \subset \mathbf{C}(\mathcal{B})$ and hence all the objects $Y^i \cap X^i$ belong to $\mathcal{B}$. For $i \neq n, n+1$, we have that $Y^i = Y^i \cap X^i \in \mathcal{B}$. We need to show that $Y^n, Y^{n+1} \in \mathcal{B}$. From the pushout square

$$\begin{tikzcd} Y^n \cap X^n \arrow[r] \arrow[d] & Y^n \arrow[d] \\ Y^{n+1} \cap X^{n+1} \arrow[r] & Y^{n+1} \end{tikzcd}$$

it follows that Y^{n+1} is a quotient of $Y^n \oplus (Y^{n+1} \cap X^{n+1})$; by Proposition 1.15(ii), Y^{n+1} will belong to $\mathcal{B}$ if Y^n does.

In Remark 3.2, we saw that Y^n is an extension of $q^{\#\Lambda'}$ by $Y^n \cap X^n \in \mathcal{B}$, where Λ' can be thought of as a subset $\Lambda' \subset \mathrm{Ext}^1(q, Y^n \cap X^n)$. By Remark 1.17, we know that $\#\Lambda' \leq \#\,\mathrm{Ext}^1(q, Y^n \cap X^n) < \mu$. But q is the coproduct of the $< \alpha$ quotients x of the generator g, all of which belong to $\mathcal{C} \subset \mathcal{B}$; hence, $q^{\#\Lambda'}$ is a coproduct of $< \mu$ objects in $\mathcal{B}$ and belongs to $\mathcal{B}$. By Proposition 1.15(iii), $\mathcal{B}$ is closed under extensions, and therefore, Y^n also belongs to $\mathcal{B}$. $\square$

Lemma 3.4. *The category $\mathcal{J}(\mathcal{I}, n)$ is μ-filtered.*

Proof. Since $\mathcal{J}(\mathcal{I}, n)$ is equivalent to a partially ordered set, we need only show that every collection of fewer than μ objects in $\mathcal{J}(\mathcal{I}, n)$ is dominated by an object of $\mathcal{J}(\mathcal{I}, n)$. Suppose therefore that we are given a set $\{Y_j \mid j \in J\}$ of $< \mu$ objects of $\mathcal{J}(\mathcal{I}, n)$. The objects $Y_j \cap X$ all belong to the μ-filtered category $\mathcal{I}$, and we may therefore choose a $Z \in \mathcal{I}$ dominating them.

Take $k \in \operatorname{Ker}(\operatorname{Ext}^1(q, Z^n) \to \operatorname{Ext}^1(q, X^n))$. Now $X^n = \underrightarrow{\operatorname{colim}}_{X_i \in \mathcal{I}} X_i^n$, the category $\mathcal{I}$ is μ-filtered, and $\operatorname{Ext}^1(q, -)$ commutes with μ-filtered colimits. Hence

$$\operatorname{Ext}^1(q, X^n) = \underrightarrow{\operatorname{colim}}_{X_i \in \mathcal{I}} \operatorname{Ext}^1(q, X_i^n),$$

and the fact that $k \in \operatorname{Ext}^1(q, Z^n)$ maps to zero in $\underrightarrow{\operatorname{colim}}_{X_i \in \mathcal{I}} \operatorname{Ext}^1(q, X_i^n)$ means that we may choose some morphism $Z \to Z_k$ in $\mathcal{I}$ so that k is annihilated by $\operatorname{Ext}^1(q, Z^n) \to \operatorname{Ext}^1(q, Z_k^n)$. We can choose such a $Z \to Z_k$ for every $k \in \operatorname{Ker}(\operatorname{Ext}^1(q, Z^n) \to \operatorname{Ext}^1(q, X^n))$. But $\# \operatorname{Ext}^1(q, Z^n) < \mu$, and hence, there are $< \mu$ possible k. Since $\mathcal{I}$ is μ-filtered, the Z_k are all dominated by some object $Z' \in \mathcal{I}$. Thus, the map $Z \to Z'$ annihilates all the k; on the image $\operatorname{Im}(\operatorname{Ext}^1(q, Z^n) \to \operatorname{Ext}^1(q, (Z')^n))$, the map to $\operatorname{Ext}^1(q, X^n)$ is injective.

For each Y_j, we have that Y_j^n is an extension of $q^{\#\Lambda_j}$ by $Y_j^n \cap X^n$, where Λ_j is a subset of $\operatorname{Ext}^1(q, Y_j^n \cap X^n)$ that maps injectively to $\operatorname{Ext}^1(q, X^n)$. We may take the image of Λ_j under the composite $Y_j^n \cap X^j \to Z^n \to (Z')^n$ or more precisely under the composite

$$\operatorname{Ext}^1(q, Y_j^n \cap X^n) \to \operatorname{Ext}^1(q, Z^n) \to \operatorname{Ext}^1(q, (Z')^n).$$

The image of each Λ_j is contained in $\operatorname{Im}(\operatorname{Ext}^1(q, Z^n) \to \operatorname{Ext}^1(q, (Z')^n))$; hence, so is the union of the images $\Lambda' = \bigcup \operatorname{Im}(\Lambda_j)$. But $\operatorname{Im}(\operatorname{Ext}^1(q, Z^n) \to \operatorname{Ext}^1(q, (Z')^n))$ maps injectively to $\operatorname{Ext}^1(q, X^n)$; and hence, so does its subset Λ'. Let $Y' = B(Z', \Lambda')$.

As in Remark 2.10, for each $y \in \mathcal{A}$, let $T(y)$ be the trivial complex

$$\cdots \to 0 \to 0 \to y \to y \to 0 \to 0 \to \cdots$$

where the nonzero terms are in degrees n and $n+1$. The objects Y_j, Y', and $B(X, n) \in \mathbf{C}(\mathcal{A})$ fit into extension sequences

$$0 \to Y_j \cap X \longrightarrow Y_j \longrightarrow T(q)^{\#\Lambda_j} \longrightarrow 0,$$

$$0 \longrightarrow Z' \longrightarrow Y' \longrightarrow T(q)^{\#\Lambda'} \longrightarrow 0,$$

$$0 \longrightarrow X \longrightarrow B(X, n) \to T(q)^{\# \operatorname{Ext}^1(x, X^n)} \to 0,$$

and the extension classes are all compatible. We may choose maps of extensions

$$\begin{tikzcd}
0 \arrow[r] & Y_j \cap X \arrow[r] \arrow[d] & Y_j \arrow[r, "\delta_j"] \arrow[d, "f_j"'] & T(q)^{\#\Lambda_j} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & Z' \arrow[r] \arrow[d] & Y' \arrow[r] \arrow[d, "g"'] & T(q)^{\#\Lambda'} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & X \arrow[r] & B(X,n) \arrow[r] & T(q)^{\#\operatorname{Ext}^1(x,X^n)} \arrow[r] & 0
\end{tikzcd}$$

Now the monomorphisms $Y_j \cap X \to Z'$, $Z' \to X$, $\Lambda_j \to \Lambda'$ and $\Lambda' \to \operatorname{Ext}^1(x, X^n)$ are all given to us explicitly. The fact that the extension classes are compatible means we may choose maps f_j and g as above, but they are not unique. Let us make the choices.

Now each Y_j is a subobject of $B(X,n)$; it comes with a given monomorphism $h_j : Y_j \to B(X,n)$ making commutative the diagram

$$\begin{tikzcd}
0 \arrow[r] & Y_j \cap X \arrow[r] \arrow[d] & Y_j \arrow[r] \arrow[d, "h_j"'] & T(q)^{\#\Lambda_j} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & X \arrow[r] & B(X,n) \arrow[r] & T(q)^{\#\operatorname{Ext}^1(x,X^n)} \arrow[r] & 0
\end{tikzcd}$$

There is no reason to expect that h_j should equal gf_j. The difference $h_j - gf_j$ must however factor through a map $T(q)^{\#\Lambda_j} \to X$, and maps in $\mathbf{C}(\mathscr{A})$ of the form $T(y) \to W$ are in bijection with maps $y \to W^n$. Thus, $h_j - gf_j$ is determined by a map in $\mathscr{A}$ of the form $q^{\#\Lambda_j} \to X^n$. But $\operatorname{Hom}(q^{\#\Lambda_j}, -)$ commutes with μ-filtered colimits, and X^n is the μ-filtered colimit of X_i^n, $X_i \in \mathscr{I}$. For each j, we may therefore choose a map $Z' \to Z_j$ in $\mathscr{I}$ so that $h_j - gf_j$ factors through $T(q)^{\#\Lambda_j} \to Z_j \subset X$. Since there are fewer than μ objects $Z_j \in \mathscr{I}$, we may find an object $Z'' \in \mathscr{I}$ dominating them. Let $W = B(Z'', \Lambda')$; that is, form the extension $0 \to Z'' \to W \to T(q)^{\#\Lambda'} \to 0$ corresponding to the image of $\Lambda' \subset \operatorname{Ext}^1(q, (Z')^n)$ under the map $\operatorname{Ext}^1(q, (Z')^n) \to \operatorname{Ext}^1(q, (Z'')^n)$.

Because the extension classes are compatible, we may construct maps of extensions

$$\begin{tikzcd}
0 \arrow[r] & Z' \arrow[r] \arrow[d] & Y' \arrow[r, "\eta"] \arrow[d, "\rho"'] & T(q)^{\#\Lambda'} \arrow[r] \arrow[d, "1"] & 0 \\
0 \arrow[r] & Z'' \arrow[r, "\tau"] \arrow[d] & W \arrow[r, "\eta'"] \arrow[d, "\sigma"'] & T(q)^{\#\Lambda'} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & X \arrow[r, "\zeta"] & B(X,n) \arrow[r] & T(q)^{\#\operatorname{Ext}^1(x,X^n)} \arrow[r] & 0
\end{tikzcd}$$

There is no reason to expect g to be equal to $\sigma\rho$, but the difference factors through some $\varphi : T(q)^{\#\Lambda'} \to X$. Changing σ to $\sigma + \zeta\varphi\eta'$, we achieve that $g = \sigma\rho$. But now we have monomorphisms $Y_j \xrightarrow{f_j} Y' \xrightarrow{\rho} W \xrightarrow{\sigma} B(X, n)$, and we have that $h_j - gf_j = h_j - \sigma\rho f_j$ factors through a map $T(q)^{\#\Lambda_j} \xrightarrow{\theta_j} Z'' \to X$. Replacing ρf_j by $f'_j = \rho f_j + \tau\theta_j\delta_j$, we have that $\sigma f'_j = h_j$ for all $j \in J$. Thus, the monomorphisms $h_j : Y_j \to B(X, n)$ all factor through $\sigma : W \to B(X, n)$, and the subobject $\sigma : W \to B(X, n)$ belongs to $\mathcal{J}(\mathcal{I}, n)$. $\square$

Lemma 3.5. *$B(X, n)$ is the colimit of its subobjects $Y \in \mathcal{J}(\mathcal{I}, n)$.*

Proof. Let Y be an object of $\mathcal{J}(\mathcal{I}, n)$. Then we have a monomorphism of short exact sequences

$$\begin{tikzcd} 0 \arrow[r] & Y \cap X \arrow[r] \arrow[d, "f_Y"'] & Y \arrow[r] \arrow[d, "g_Y"'] & T(q^{\#\Lambda'}) \arrow[r] \arrow[d, "h_Y"'] & 0 \\ 0 \arrow[r] & X \arrow[r] & B(X, n) \arrow[r] & T(q^{\#\operatorname{Ext}^1(q, X^n)}) \arrow[r] & 0 \end{tikzcd}$$

with h_Y being the inclusion of a subcoproduct. Since the category $\mathcal{J}(\mathcal{I}, n)$ is filtered, the colimit over $\mathcal{J}(\mathcal{I}, n)$ of the top row is exact; we wish to show that the colimit of g_Y is an isomorphism, and the five lemma tells us that it suffices to prove that the colimits of f_Y and h_Y are isomorphisms. Also, since f_Y and h_Y are monomorphisms, so are their colimits. It therefore suffices to prove that the colimits of f_Y and h_Y are epi.

For f_Y, note that the category $\mathcal{I}$ embeds in the category $\mathcal{J}(\mathcal{I}, n)$; we can view a subobject $Y \subset X$ as a subobject of $B(X, n)$, where the corresponding $\Lambda' \subset \operatorname{Ext}^1(q, Y^n)$ is empty. But the colimit of $\mathcal{I}$ maps epimorphically to X, and this epimorphism will factor through the colimit of f_Y. Hence, the colimit of f_Y must be epi.

We need to show that the colimit of h_Y is epi. Take any $\lambda \in \operatorname{Ext}^1(q, X^n)$; because $\operatorname{Ext}^1(q, -)$ commutes with μ-filtered colimits and X^n is the μ-filtered colimit of X_i^n, $X_i \in \mathcal{I}$, we may choose a $Z \in \mathcal{I}$ and an element $e_\lambda \in \operatorname{Ext}^1(q, Z^n)$ mapping to λ. Form the extension $0 \to Z \to Y \to T(q) \to 0$ corresponding to e_λ. From the fact that e_λ maps to λ, we have a map of extensions

$$\begin{tikzcd} 0 \arrow[r] & Z \arrow[r] \arrow[d, "f_Y"'] & Y \arrow[r] \arrow[d, "g_Y"'] & T(q) \arrow[r] \arrow[d, "h_Y"'] & 0 \\ 0 \arrow[r] & X \arrow[r] & B(X, n) \arrow[r] & T(q^{\#\operatorname{Ext}^1(q, X^n)}) \arrow[r] & 0 \end{tikzcd}$$

where h_Y is the inclusion of the subcoproduct over the singleton $\{\lambda\}$. Thus, the image of the colimit of the h_Y must contain the coproduct over every singleton in $\operatorname{Ext}^1(q, X^n)$, and hence, it must be epi. $\square$

Lemma 3.6. *If filtered colimits of $< \mu$ objects in $\mathcal{I}$ belong to $\mathcal{I}$, then filtered colimits of $< \mu$ objects in $\mathcal{J}(\mathcal{I}, n)$ belong to $\mathcal{J}(\mathcal{I}, n)$.*

Proof. An object Y belongs to $\mathcal{J}(\mathcal{I}, n)$ if it comes with a monomorphism of short exact sequences

$$\begin{tikzcd} 0 \arrow[r] & Y \cap X \arrow[r] \arrow[d, "f_Y"'] & Y \arrow[r] \arrow[d, "g_Y"'] & T(q^{\#\Lambda'}) \arrow[r] \arrow[d, "h_Y"'] & 0 \\ 0 \arrow[r] & X \arrow[r] & B(X, n) \arrow[r] & T(q^{\#\mathrm{Ext}^1(q, X^n)}) \arrow[r] & 0 \end{tikzcd}$$

where h_Y is the inclusion of a subcoproduct and $Y \cap X \in \mathcal{I}$. A filtered colimit of objects $Y_\lambda \cap X$, $\lambda \in \Lambda$, will belong to $\mathcal{I}$ as long as $\#\Lambda < \mu$ and each $Y_\lambda \cap X \in \mathcal{I}$. Filtered colimits are exact, and hence, the filtered colimit of $< \mu$ monomorphisms of short exact sequences as above is such a monomorphism. $\square$

Construction 3.7. Recall Remark 2.15: the object $I(X) = J^\alpha(X)$ can be constructed using a single sequence. Let us now remember this sequence:

(i) $X_0 = X$.

(ii) $X_{i+1} = B(X_i, n)$ for some n depending on i. The precise relation is that if $i = \ell + m$, where ℓ is a limit ordinal and m is an integer, then $n = -m/2$ if m is even and $n = (m+1)/2$ if m is odd.

(iii) For limit ordinals j we have $X_j = \underrightarrow{\mathrm{colim}}_{i<j}\, X_i$.

Suppose we are given an α-filtered category $\mathcal{I}$ of subobjects of X, whose colimit is X. For every ordinal i, we will now form a subcategory $\mathcal{I}_i$ of subobjects of X_i. The rules are:

(i) $\mathcal{I}_0 = \mathcal{I}$.

(ii) If n is the integer for which $X_{i+1} = B(X_i, n)$, then $\mathcal{I}_{i+1} = \mathcal{J}(\mathcal{I}_i, n)$.

(iii) Let j be a limit ordinal. A subobject $Y \subset X_j$ belongs to $\mathcal{I}_j$ if and only if $Y \cap X_i$ belongs to $\mathcal{I}_i$ for all $i < j$.

Lemma 3.8. *Suppose $Y \subset X_j$ lies in $\mathcal{I}_j$ in the notation of Construction 3.7. Then in the triangle $Y \cap X \to Y \to A \to$, we have that A belongs to ${}^\perp\mathbf{K}(\mathrm{Inj}\,\mathcal{A})$.*

Proof. Consider the sequence $Y_i = Y \cap X_i$. By hypothesis, $Y_{i+1} \in \mathcal{J}(\mathcal{I}_i, n)$, and in Remark 3.2, we saw that $Y_{i+1} = B(Y_{i+1} \cap X, \Lambda') = B(Y_i, \Lambda')$ for some subset $\Lambda' \subset \mathrm{Ext}^1(q, Y_i^n)$ mapping injectively to $\mathrm{Ext}^1(q, X^n)$. And for limit ordinals ℓ, we have $Y_\ell = Y \cap X_\ell = \underrightarrow{\mathrm{colim}}_{i<\ell}(Y \cap X_i) = \underrightarrow{\mathrm{colim}}_{i<\ell} Y_i$. The lemma now follows from Lemma 2.8 and Remark 2.10. $\square$

Lemma 3.9. *Let X be an object of $\mathbf{C}(\mathscr{A})$, and let $\mathscr{I} \subset \mathbf{C}(\mathscr{B}) = \mathbf{C}(\mathscr{A}^{\mu})$ be a full subcategory of the subobjects of X. Assume $\mathscr{I}$ is μ-filtered with colimit X, and assume that filtered colimits of $< \mu$ objects in $\mathscr{I}$ belong to $\mathscr{I}$.*

Then for every ordinal $i \leq \alpha$, we have that $\mathscr{I}_i$ has the same properties: it is contained in $\mathbf{C}(\mathscr{B})$, is μ-filtered with colimit X_i, and is closed under filtered colimits of $< \mu$ objects.

Proof. For $i = 0$, we have $\mathscr{I}_0 = \mathscr{I}$ and there is nothing to prove. Suppose the result is true for i. By Lemma 3.3, we have $\mathscr{I}_{i+1} \subset \mathbf{C}(\mathscr{B})$; by Lemma 3.4, it is μ-filtered; by Lemma 3.5, the colimit is X_{i+1}; and by Lemma 3.6, it is closed under filtered colimits of $< \mu$ objects.

For the remainder of the proof, assume j is a limit ordinal and the assertions of the lemma are true for all $i < j$. Let Y be an object of $\mathscr{I}_j$. By definition, $Y \cap X_i \in \mathscr{I}_i$ for every $i < j$, and by induction, $Y \cap X_i \in \mathbf{C}(\mathscr{B})$. But $Y = \underrightarrow{\mathrm{colim}}_{i<j}(Y \cap X_i)$ is the colimit of $\leq \alpha < \mu$ objects of $\mathbf{C}(\mathscr{B})$; by Remark 1.16, $Y \in \mathbf{C}(\mathscr{B})$.

Let $Y = \underrightarrow{\mathrm{colim}}_{r\in R} Y_r$ with $Y_r \in \mathscr{I}_j$ and R be a filtered category with $< \mu$ objects. For $i < j$, we have that $Y \cap X_i = \underrightarrow{\mathrm{colim}}_{r\in R}(Y_r \cap X_i)$ belongs to $\mathscr{I}_i$ by the induction hypothesis, and hence, $Y \in \mathscr{I}_j$ by definition.

Let $\{Y_r \mid r \in R\}$ be a set of $< \mu$ objects of $\mathscr{I}_j$. By induction on $i < j$, we choose

(i) an object $Z_0 \in \mathscr{I}_0$ containing all the $Y_r \cap X_0$ and

(ii) an object $Z_{i+1} \in \mathscr{I}_{i+1}$ containing Z_i and all the objects $Y_r \cap X_{i+1}$.

(iii) For limit ordinals ℓ, define $Z_\ell = \underrightarrow{\mathrm{colim}}_{i<\ell} Z_i$. Then Z_ℓ belongs to $\mathscr{I}_\ell$ since $Z_\ell \cap X_k = \underrightarrow{\mathrm{colim}}_{i<\ell} Z_i \cap X_k$ belongs to $\mathscr{I}_k$ for all $k < \ell \leq j$.

But now $Z_j \in \mathscr{I}_j$ contains all the Y_r. Thus, $\mathscr{I}_j$ is μ-filtered.

The category $\mathscr{I}_j$ is a filtered category of subobjects of X_j, and the colimit is some subobject of X_j. But it contains the colimits of $\mathscr{I}_i \subset \mathscr{I}_j$ for all $i < j$; that is, it contains all the X_i with $i < j$. Because $X_j = \underrightarrow{\mathrm{colim}}_{i<j} X_i$, we conclude that the colimit of $\mathscr{I}_j$ is all of X_j. □

Lemma 3.10. *Let X and $\mathscr{I}$ be as in Lemma 3.9. The full subcategory of $\mathscr{I}_\alpha$ whose objects are in $\mathbf{C}(\mathrm{Inj}\,\mathscr{A})$ is cofinal.*

Proof. Let Y be an arbitrary object of $\mathscr{I}_\alpha$; we need to produce a morphism $Y \to Z$ in $\mathscr{I}_\alpha$ with $Z \in \mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. We inductively define a sequence $\{Z_i\}$ of objects in $\mathscr{I}_i$, and Z will be the colimit; the recipe is:

(i) Put $Z_0 = Y \cap X$.

(ii) Assume n is the integer for which $\mathscr{I}_{i+1} = \mathscr{J}(\mathscr{I}_i, n)$, and suppose we have defined $Z_i \in \mathscr{I}_i$. Choose an object $W_i \in \mathscr{I}_i$ containing Z_i and $Y \cap X_i$. There is a morphism $W_i \to V_i$ in $\mathscr{I}_i$ annihilating the kernel of $\mathrm{Ext}^1(q, W_i^n) \to \mathrm{Ext}^1(q, X_i^n)$; we saw its existence in the proof of Lemma 3.4. Let $\Lambda' \subset \mathrm{Ext}^1(q, V_i^n)$ be the image of the map $\mathrm{Ext}^1(q, W_i^n) \to \mathrm{Ext}^1(q, V_i^n)$; then Λ' maps injectively to $\mathrm{Ext}^1(q, X_i^n)$, and we can define $Z_{i+1} \in \mathscr{I}_{i+1}$ to be $B(V_i, \Lambda')$. That is, Z_{i+1} is given by a map of extensions

$$\begin{tikzcd} 0 \arrow[r] & V_i \arrow[r] \arrow[d, "f"'] & Z_{i+1} \arrow[r] \arrow[d, "g"'] & T(q^{\#\Lambda'}) \arrow[r] \arrow[d, "h"'] & 0 \\ 0 \arrow[r] & X \arrow[r] & B(X,n) \arrow[r] & T(q^{\#\mathrm{Ext}^1(q,X^n)}) \arrow[r] & 0 \end{tikzcd}$$

Note that the monomorphisms f and h are given, and we make a choice of a compatible g.

(iii) For limit ordinals ℓ, define $Z_\ell = \underrightarrow{\mathrm{colim}}_{i<\ell}\, Z_i$.

We have a map $Z_i \to Z_{i+1}$ that factors as $Z_i \to W_i \to V_i \to Z_{i+1}$. By construction, the map $W_i \to V_i$ kills the kernel of the map $\varphi : \mathrm{Ext}^1(q, W_i^n) \to \mathrm{Ext}^1(q, X_i^n)$ while the morphism $V_i \to Z_{i+1}$ kills the image of φ. It follows that the composite $Z_i^n \to Z_{i+1}^n$ is annihilated by $\mathrm{Ext}^1(q, -)$ for the choice of n for which $X_{i+1} = B(X_i, n)$.

The n are chosen so that, for any limit ordinal i, every integer n occurs between i and $i+\omega$. If we restrict to limit ordinals, we have that $\mathrm{Ext}^1(q, -)$ annihilates $Z_i^n \to Z_j^n$ for any integer n and any pair $i < j$ of limit ordinals. But Z_α is the α-filtered colimit of the limit ordinals $< \alpha$, and for each quotient x of the generator g, we have that $\mathrm{Ext}^1(x, -)$ commutes with α-filtered colimits. It follows that $\mathrm{Ext}^1(x, Z_\alpha^n) = 0$ for all x and all n, and hence, Z_α is a complex of injectives. And by construction, $Y = \underrightarrow{\mathrm{colim}}(Y \cap X_i)$ maps in $\mathscr{I}_\alpha$ to $Z = \underrightarrow{\mathrm{colim}}\, Z_i$. □

Corollary 3.11. *Let $Y \in \mathscr{I}_\alpha$ be in the cofinal subcategory of objects that lie in $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. Then $Y = I(Y \cap X)$.*

Proof. By Lemma 3.8, the triangle $Y \cap X \to Y \to A \to$ has $A \in {}^\perp\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. Since Y belongs to $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$, the triangle identifies Y with $I(Y \cap X)$. □

Lemma 3.12. *Let $I : \mathbf{C}(\mathscr{A}) \to \mathbf{C}(\mathrm{Inj}\,\mathscr{A})$ be the functor of Theorem 2.13. Then the objects $\{I(s) \mid s \in \mathbf{C}(\mathscr{B})\}$ generate the triangulated category $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$.*

Proof. For every nonzero object $X \in \mathbf{K}(\mathrm{Inj}\,\mathscr{A})$, we need to produce a nonzero map $I(s) \to X$ in $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ or equivalently (by the fact that I is left-adjoint to the inclusion) a nonzero map $s \to X$ in $\mathbf{K}(\mathscr{A})$. The proof is as in [Krause 2001, Lemma 2.2].

If X is not acyclic, there is a nontrivial cohomology group; without loss, we may assume $H^0(X) \neq 0$. Let $K \subset X^0$ be the kernel of $\partial : X^0 \to X^1$; we may choose a map $g \to K$ that does not factor through the image of $X^{-1} \to X^0$. But then $g \to K \to X^0$ extends to a chain map $g \to X$ that is nonzero in homology, and g is μ-presentable; that is, $g \in \mathbf{C}(\mathscr{B})$.

It remains to handle the case where X is acyclic. If X is nonzero in $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$, then it is not a contractible complex, so there must be an n for which $\mathrm{Im}(X^n \to X^{n+1})$ is not an injective object of $\mathscr{A}$. Suppose without loss that $M = \mathrm{Im}(X^{-2} \to X^{-1})$ is not injective. Then there is a quotient x of the generator g and a nonzero element of $\mathrm{Ext}^1(x, M)$. But elements of $\mathrm{Ext}^1(x, M)$ are in bijection with morphisms $x \to X$ in $\mathbf{K}(\mathscr{A})$, so we have produced a nonzero map $x \to X$ where x is μ-presentable, that is, $x \in \mathbf{C}(\mathscr{B})$. □

Theorem 3.13. *Let $\mathscr{B} \subset \mathscr{A}$ be as in Definition 1.13; by Proposition 1.18, it is precisely the category $\mathscr{B} = \mathscr{A}^\mu$ of μ-presentable objects in $\mathscr{A}$. Then the objects $\{I(s) \mid s \in \mathbf{C}(\mathscr{B})\}$ form a μ-compact generating set in the category $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. Therefore, $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ is well generated.*

Proof. The fact that these objects generate was proved in Lemma 3.12; what remains is to show that they form a μ-compact generating set, meaning that they form a μ-perfect set of μ-small objects; see [Neeman 2001, §3.3, §4.1, and §4.2]. Suppose we are given a set $\{X_\lambda \mid \lambda \in \Lambda\}$ of objects of $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$. Then the coproduct of these objects in $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ is formed by applying the functor I of Theorem 2.13 to the ordinary coproduct in $\mathbf{K}(\mathscr{A})$ or $\mathbf{C}(\mathscr{A})$. But now, in the category $\mathbf{C}(\mathscr{A})$, each X_λ is the μ-filtered colimit of its subobjects $\{s \to X_\lambda \mid s \in \mathbf{C}(\mathscr{B})\}$, and the coproduct of the X_λ satisfies

$$\coprod_{\lambda\in\Lambda} X_\lambda = \underset{\substack{\Lambda'\subset\Lambda \\ \#\Lambda'<\mu}}{\underrightarrow{\mathrm{colim}}} \coprod_{\lambda\in\Lambda'} \underset{\substack{s_\lambda\to X_\lambda \\ s_\lambda\in\mathbf{C}(\mathscr{B})}}{\underrightarrow{\mathrm{colim}}} s_\lambda .$$

Thus, we wish to apply our lemmas to the object $X = \coprod X_\lambda$ and to the category $\mathscr{I}$ consisting of subobjects $\coprod_{\lambda\in\Lambda'} s_\lambda$, where $\Lambda' \subset \Lambda$ is a set with $\#\Lambda' < \mu$ and each $s_\lambda \in \mathbf{C}(\mathscr{B})$ is a subobject of X_λ.

We have proved that $I(X) = X_\alpha$ is the colimit in $\mathbf{C}(\mathscr{A})$ of the μ-filtered category $\mathscr{I}_\alpha$, and hence, any map from the μ-presentable $s \in \mathbf{C}(\mathscr{B})$ to $I(X)$ must factor through some object $Y \in \mathscr{I}_\alpha$. By Lemma 3.8, Y fits in a triangle $Y \cap X \to Y \to A \to$ with $A \in {}^\perp\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$; if we apply the functor I, then it takes $Y \cap X \to Y$ to an isomorphism. Thus, the map $s \to I(X)$ factors as $s \to I(Y \cap X) \to I(X)$, and $Y \cap X$ is an object of $\mathscr{I}$, meaning a coproduct $\coprod_{\lambda\in\Lambda'} s_\lambda$, where $\Lambda' \subset \Lambda$ is a set with $\#\Lambda' < \mu$ and each $s_\lambda \in \mathbf{C}(\mathscr{B})$ is a subobject of X_λ. In the category $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$, we have factored the map as

$$I(s) \to I\Big(\coprod_{\lambda\in\Lambda'} s_\lambda\Big) \to I\Big(\coprod_{\lambda\in\Lambda} X_\lambda\Big).$$

Now suppose that we are given in $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ a vanishing composite

$$I(s) \xrightarrow{\theta} I\Big(\coprod_{\lambda\in\Lambda'} s_\lambda\Big) \xrightarrow{\sigma} I\Big(\coprod_{\lambda\in\Lambda} X_\lambda\Big),$$

that is, we are given a map θ so that $\sigma\theta$ is null-homotopic. Let us write this a little more compactly: we are given a morphism $\theta : I(s) \to I(W)$ so that the composite $I(s) \xrightarrow{\theta} I(W) \xrightarrow{\sigma} I(X)$ is null-homotopic with $W = \coprod_{\lambda\in\Lambda'} s_\lambda$ belonging to $\mathscr{I} = \mathscr{I}_0$. Of course, we are free to replace W by a larger subobject in $\mathscr{I}$ before proceeding any further, and Lemma 3.10 tells us that in the category $\mathscr{I}_\alpha$ the objects that belong to $\mathbf{C}(\operatorname{Inj}\mathscr{A})$ are cofinal. We may therefore produce in $\mathscr{I}_\alpha$ a map $W \to Y$ with Y in the subcategory. We have a morphism $s \to I(s) \to I(Y) = Y$ in $\mathbf{K}(\operatorname{Inj}\mathscr{A})$; choose a representative, that is, a chain map $s \to Y$ in $\mathbf{C}(\mathscr{A})$. We know that the composite $s \to Y \to I(X)$ is null-homotopic.

But $I(X)$ is the μ-filtered colimit of $\mathscr{I}_\alpha$, and s is μ-presentable. There is a map $Y \to Z$ in $\mathscr{I}_\alpha$ so that the composite $s \to Y \to Z$ is already null-homotopic. Now recalling that the maps $I(Y \cap X) \to I(Y) = Y$ and $I(Z \cap X) \to I(Z)$ are isomorphisms in $\mathbf{K}(\operatorname{Inj}\mathscr{A})$, we have proved that for some $Z \in \mathscr{I}_\alpha$ the map $s \to I(Y \cap X) \to I(Z \cap X)$ is zero in $\mathbf{K}(\operatorname{Inj}\mathscr{A})$, that is, we have proved the vanishing of some

$$I(s) \xrightarrow{\theta} I\Bigl(\coprod_{\lambda\in\Lambda'} s_\lambda\Bigr) \to I\Bigl(\coprod_{\lambda\in\Lambda''} t_\lambda\Bigr). \qquad \square$$

4. The failure of recollement

In the generality where $\mathscr{A}$ is any Grothendieck abelian category, we have natural functors

$$\mathbf{K}_{\mathrm{ac}}(\operatorname{Inj}\mathscr{A}) \xrightarrow{J} \mathbf{K}(\operatorname{Inj}\mathscr{A}) \xrightarrow{Q} \mathbf{D}(\mathscr{A})$$

that compose to zero. But the functor Q has a right adjoint Q_ρ, namely the functor taking $X \in \mathbf{D}(\mathscr{A})$ to its K-injective injective resolution. Since Q_ρ is fully faithful, the map Q must be a Verdier quotient, but J is precisely the inclusion of the kernel of Q. It therefore follows that J also has a right adjoint J_ρ.

Krause [2005] proves that, provided the category $\mathscr{A}$ is locally noetherian and $D(\mathscr{A})$ is compactly generated, then J and Q have left adjoints as well. In particular, J takes products to products: products of acyclic complexes of injectives are acyclic. What we will now produce is:

Example 4.1. There is a locally noetherian abelian category $\mathscr{A}$ for which $\mathbf{K}(\operatorname{Inj}\mathscr{A})$ is not closed under products. The category $\mathscr{A}$ will be (a special case of) the category $\mathscr{A}$ of [Neeman 2011, Construction 1.1]; the counterexample works in the generality of the $\mathscr{A}$ of [Neeman 2011, Construction 1.1], but for simplicity, we will specialize to a particular case. And the chain complex of injectives will be a minor modification of the chain complex of [Neeman 2011, proof of Theorem 1.1, pp. 830–831].

Let k be a field, and let $R_1 = R$ be the ring $k[x]/(x^2)$ of dual numbers over k. The ring R_n is $R \otimes_k R \otimes_k \cdots \otimes_k R$, the tensor product of n copies of R. The inclusions

$R_n \to R_{n+1}$ are the inclusions into the first n factors. And S is the colimit of R_n. If we write S as

$$S = \frac{k[x_1, x_2, x_3, \dots]}{(x_1^2, x_2^2, x_3^2, \dots)},$$

then $\mathscr{A}$ is the category of all S-modules M, where each element $m \in M$ is annihilated by all but finitely many of the x_i.

Let B be an injective resolution of Σk over the ring R; for definiteness, let us choose B to be the complex starting in degree -1

$$\cdots \to 0 \to 0 \to R \xrightarrow{x} R \xrightarrow{x} \cdots .$$

Let C_n be the complex $B^{\otimes n}$, that is, the tensor product of n copies of B. Then C_n is an injective resolution of $\Sigma^n k$ over the ring $R_n = k[x_1, x_2, \dots, x_n]/(x_1^2, x_2^2, \dots, x_n^2)$. Consider the chain map $k \to B$, which takes $1 \in k$ to $x \in R = B^0$, that is, the chain map

$$\begin{tikzcd}
\cdots \arrow[r] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & 0 \arrow[r] \arrow[d] & k \arrow[r] \arrow[d, "x"'] & 0 \arrow[r] \arrow[d, "x"'] & \cdots \\
\cdots \arrow[r] & 0 \arrow[r] \arrow[d, equal] & 0 \arrow[r] \arrow[d, equal] & R \arrow[r, "x"] \arrow[d, equal] & R \arrow[r, "x"] \arrow[d, equal] & R \arrow[r, "x"] \arrow[d, equal] & \cdots \\
\cdots \arrow[r] & B^{-3} \arrow[r] & B^{-2} \arrow[r] & B^{-1} \arrow[r] & B^0 \arrow[r] & B^1 \arrow[r] & \cdots
\end{tikzcd}$$

We have an induced inclusion $C_n = C_n \otimes_k k \to C_n \otimes_k B = C_{n+1}$, and we define C to be the colimit of the C_n. Then C is an acyclic complex of injective objects in $\mathscr{A}$.

Now let $c_n \in C^0$ be the cycle $x \otimes x \otimes x \otimes \cdots$, which we view as $x_1 \otimes x_2 \otimes x_3 \otimes \cdots$; the only question is which degree each $x_i = x \in R$ lives in. The rule is: for $1 \le i \le n$, we have $x_i \in B^{-1} = R$; for $n+1 \le i \le 2n$, we have $x_i \in B^1 = R$; and for $2n < i$, we put $x_i \in B^0 = R$. Note that $x_1 \otimes x_2 \otimes \cdots \otimes x_n \in C_n^{-n}$ is not a boundary; it defines the unique nonvanishing cohomology class of $C_n \cong \Sigma^n k$. Of course, $c_n \in C$ is a cycle in the acyclic complex C and hence a boundary, but it must be a boundary of some chain in C^{-1} that nontrivially involves the terms in the tensor product with $i > n$; in other words, if c_n is the boundary of a chain $b_n \in C$, then there exists an $i > n$ so that $x_i b_n \neq 0$. The product $\prod_{n=1}^{\infty} c_n$ is a product of cycles in C^0 and hence is a cycle in the complex $\prod_{n=0}^{\infty} C$. But it cannot be a boundary; if it were the boundary of $\prod_{n=1}^{\infty} b_n$, we would have infinitely many i and infinitely many n_i for which $x_i b_{n_i} \neq 0$, meaning $\prod_{n=1}^{\infty} b_n$ cannot belong to the category $\mathscr{A}$.

Remark 4.2. Since the category $\mathscr{A}$ is locally noetherian, it follows that its derived category $\mathbf{D}(\mathscr{A})$ cannot possibly be compactly generated. After all, Krause [2005] proved that when $D(\mathscr{A})$ is compactly generated then $\mathbf{K}(\mathrm{Inj}\,\mathscr{A})$ is closed under products.

References

[Bican et al. 2001] L. Bican, R. El Bashir, and E. Enochs, "All modules have flat covers", *Bull. London Math. Soc.* **33**:4 (2001), 385–390. MR 2002e:16002 Zbl 1029.16002

[Gabriel and Ulmer 1971] P. Gabriel and F. Ulmer, *Lokal präsentierbare Kategorien*, Lecture Notes in Mathematics **221**, Springer, Berlin, 1971. MR 48 #6205 Zbl 0225.18004

[Iyengar and Krause 2006] S. Iyengar and H. Krause, "Acyclicity versus total acyclicity for complexes over Noetherian rings", *Doc. Math.* **11** (2006), 207–240. MR 2007h:16013 Zbl 1119.13014

[Jørgensen 2005] P. Jørgensen, "The homotopy category of complexes of projective modules", *Adv. Math.* **193**:1 (2005), 223–232. MR 2005m:18013 Zbl 1068.18012

[Krause 2001] H. Krause, "On Neeman's well generated triangulated categories", *Doc. Math.* **6** (2001), 121–126. MR 2002c:18008 Zbl 0987.18012

[Krause 2005] H. Krause, "The stable derived category of a Noetherian scheme", *Compos. Math.* **141**:5 (2005), 1128–1162. MR 2006e:18019 Zbl 1090.18006

[Krause 2012] H. Krause, "Approximations and adjoints in homotopy categories", *Math. Ann.* **353**:3 (2012), 765–781. MR 2923949 Zbl 06050576

[Murfet 2007] D. S. Murfet, *The mock homotopy category of projectives and Grothendieck duality*, Ph.D. thesis, Australian National University, 2007, Available at http://therisingsea.org/notes/thesis.pdf.

[Neeman 2001] A. Neeman, *Triangulated categories*, Annals of Mathematics Studies **148**, Princeton University Press, 2001. MR 2001k:18010 Zbl 0974.18008

[Neeman 2008] A. Neeman, "The homotopy category of flat modules, and Grothendieck duality", *Invent. Math.* **174**:2 (2008), 255–308. MR 2009h:16008 Zbl 1184.18008

[Neeman 2010] A. Neeman, "Some adjoints in homotopy categories", *Ann. of Math.* (2) **171**:3 (2010), 2143–2155. MR 2011i:18026 Zbl 1205.18008

[Neeman 2011] A. Neeman, "Non-left-complete derived categories", *Math. Res. Lett.* **18**:5 (2011), 827–832. MR 2875857 Zbl 1244.18009

Communicated by David Benson
Received 2013-03-14 Revised 2013-08-06 Accepted 2013-09-09

amnon.neeman@anu.edu.au *Centre for Mathematics and its Applications, Mathematical Sciences Institute, Australian National University, John Dedman Building 27, Canberra ACT 0200, Australia*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.457

Essential dimension of spinor and Clifford groups

Vladimir Chernousov and Alexander Merkurjev

We conclude the computation of the essential dimension of split spinor groups, and an application to algebraic theory of quadratic forms is given. We also compute essential dimension of the split even Clifford group or, equivalently, of the class of quadratic forms with trivial discriminant and Clifford invariant.

1. Introduction

We recall briefly the definition of the essential dimension.

Let F be a field, and let $\mathcal{F} : \mathit{Fields}/F \to \mathit{Sets}$ be a functor from the category of field extensions over F to the category of sets. Let $E \in \mathit{Fields}/F$ and $K \subset E$ a subfield over F. We say that K is a *field of definition* of an element $\alpha \in \mathcal{F}(E)$ if α belongs to the image of the map $\mathcal{F}(K) \to \mathcal{F}(E)$. The *essential dimension of* α, denoted $\mathrm{ed}^{\mathcal{F}}(\alpha)$, is the least transcendence degree $\mathrm{tr.deg}_F(K)$ over all fields of definition K of α. The *essential dimension of the functor* $\mathcal{F}$ is

$$\mathrm{ed}(\mathcal{F}) = \sup\{\mathrm{ed}^{\mathcal{F}}(\alpha)\},$$

where the supremum is taken over all fields $E \in \mathit{Fields}/F$ and all $\alpha \in \mathcal{F}(E)$ (see [Berhuy and Favi 2003, Definition 1.2] or [Merkurjev 2009, §1]). Informally, the essential dimension of $\mathcal{F}$ is the smallest number of algebraically independent parameters required to define $\mathcal{F}$ and may be thought of as a measure of complexity of $\mathcal{F}$.

Let p be a prime integer. The *essential p-dimension of* $\alpha \in \mathcal{F}(E)$, denoted $\mathrm{ed}_p^{\mathcal{F}}(\alpha)$, is defined as the minimum of $\mathrm{ed}^{\mathcal{F}}(\alpha_{E'})$, where E' ranges over all finite field extensions of E of degree prime to p and $\alpha_{E'}$ is the image of α under the map $\mathcal{F}(E) \to \mathcal{F}(E')$. The *essential p-dimension of* $\mathcal{F}$ is

$$\mathrm{ed}_p(\mathcal{F}) = \sup\{\mathrm{ed}_p^{\mathcal{F}}(\alpha)\},$$

Chernousov's work has been supported in part by the Canada Research Chairs Program and an NSERC research grant. Merkurjev's work has been supported by the NSF grant DMS #1160206. The authors thank Z. Reichstein for useful comments and suggestions, and the Fields Institute for its hospitality.
MSC2010: primary 11E04, 11E57, 11E72; secondary 11E81, 14L35, 20G15.
Keywords: Linear algebraic groups, spinor groups, essential dimension, torsor, nonabelian cohomology, quadratic forms, Witt rings, the fundamental ideal.

where the supremum ranges over all fields $E \in \mathit{Fields}/F$ and all $\alpha \in \mathcal{F}(E)$. By definition, $\mathrm{ed}(\mathcal{F}) \geq \mathrm{ed}_p(\mathcal{F})$ for all p.

For convenience, we write $\mathrm{ed}_0(\mathcal{F}) = \mathrm{ed}(\mathcal{F})$, so $\mathrm{ed}_p(\mathcal{F})$ is defined for $p = 0$ and all prime p.

Let G be an algebraic group scheme over F. Write $\mathcal{F}_G$ for the functor taking a field extension E/F to the set $H^1_{\text{ét}}(E, G)$ of isomorphism classes of principal homogeneous G-spaces (G-torsors) over E. The essential (p-)dimension of $\mathcal{F}_G$ is called the *essential* (p-)*dimension of* G and is denoted by $\mathrm{ed}(G)$ and $\mathrm{ed}_p(G)$ (see [Reichstein 2000; Reichstein and Youssin 2000]). Thus, the essential dimension of G measures complexity of the class of principal homogeneous G-spaces.

In this paper, we conclude the computation of the essential dimension of the split spinor groups $\mathbf{Spin}_n$ originated in [Brosnan et al. 2010; Garibaldi 2009] and continued in [Merkurjev 2009] (Theorem 2.2). In the missing case $n = 4m \geq 16$, we prove that

$$\mathrm{ed}_2(\mathbf{Spin}_n) = \mathrm{ed}(\mathbf{Spin}_n) = 2^{(n-2)/2} + 2^m - \frac{n(n-1)}{2},$$

where 2^m is the largest power of 2 dividing n. The value of $\mathrm{ed}(\mathbf{Spin}_n)$ is surprisingly large. Recall a striking consequence of this (see [Brosnan et al. 2010, Theorem 1-1]): the Pfister number $\mathrm{Pf}(3, n)$ is at least exponential in n.

In Theorem 4.2, we give an application in algebraic theory of quadratic forms. Precisely, we determine all pairs (n, b) of natural numbers (with two possible exceptions) such that, for every field F, any quadratic form in $I^3(F)$ of dimension n contains a subform of trivial discriminant of dimension b. This result, stated entirely in terms of algebraic theory of quadratic forms, is proved using the tools of the essential dimension!

Theorem 4.2 is applied later in the paper for the computation of the essential dimension of split even Clifford group $\mathbf{\Gamma}^+_n$ or, equivalently, of the functor given by n-dimensional quadratic forms with trivial discriminant and Clifford invariant (Theorem 7.1).

We use heavily the work [Popov 1987], where the base field is assumed to be of characteristic zero. This explains the characteristic restriction in most of our results.

2. Essential dimension of $\mathbf{Spin}_n$

Let G be an algebraic group over F, and let $C \subset G$ be a normal subgroup over F. For a torsor $E \to \mathrm{Spec}(F)$ of the group $H := G/C$, consider the stack $[E/G]$ (see [Vistoli 2005]). Recall that an object of the category $[E/G](K)$ for a field extension K/F is a pair (E', φ), where E' is a G-torsor over K and $\varphi : E'/C \xrightarrow{\sim} E_K$ is an isomorphism of H-torsors over K. The essential dimension $\mathrm{ed}[E/G]$ of the stack $[E/G]$ is the essential dimension of the functor $K \mapsto$ set of isomorphism classes of objects in $[E/G](K)$.

The following was proven independently by R. Lötscher [2013, Example 3.4]:

Proposition 2.1. *Let C be a normal subgroup of an algebraic group G over F and $H = G/C$. Then*

$$\mathrm{ed}(G) \le \mathrm{ed}(H) + \max \mathrm{ed}[E/G],$$

where the maximum is taken over all field extensions L/F and all H-torsors E over L.

Proof. Let I' be a G-torsor over a field extension K/F. Then $I := I'/C$ is an H-torsor over K. There is a subextension K_0/F of K/F and an H-torsor E over K_0 such that there is an isomorphism $\varphi : I \xrightarrow{\sim} E_K$ of H-torsors and $\mathrm{tr.deg}(K_0/F) \le \mathrm{ed}(H)$.

Consider the stack $[E/G]$ over K_0. The pair (I', φ) is an object of $[E/G](K)$. There is a subextension K_1/K_0 of K/K_0 such that (I', φ) is defined over K_1 and $\mathrm{tr.deg}(K_1/K_0) \le \mathrm{ed}[E/G]$. It follows that I' is defined over the field K_1 with

$$\mathrm{tr.deg}(K_1/F) = \mathrm{tr.deg}(K_0/F) + \mathrm{tr.deg}(K_1/K_0) \le \mathrm{ed}(H) + \mathrm{ed}[E/G]. \qquad \square$$

The following theorem concludes computation of the essential dimension of the spinor groups initiated in [Brosnan et al. 2010; Garibaldi 2009] and continued in [Merkurjev 2009]. We write $\mathbf{Spin}_n$ for the split spinor group of a nondegenerate quadratic form of dimension n and maximal Witt index.

If $\mathrm{char}(F) \ne 2$, then the essential dimension of $\mathbf{Spin}_n$ has the following values for $n \le 14$ (see [Garibaldi 2009, §23]):

n	≤ 6	7	8	9	10	11	12	13	14
$\mathrm{ed}_2(\mathbf{Spin}_n) = \mathrm{ed}(\mathbf{Spin}_n)$	0	4	5	5	4	5	6	6	7

In the following theorem, we give the values of $\mathrm{ed}_p(\mathbf{Spin}_n)$ for $n \ge 15$ and $p = 0$ and 2. Note that $\mathrm{ed}_p(\mathbf{Spin}_n) = 0$ if $p \ne 0, 2$ as every $\mathbf{Spin}_n$-torsor over a field is split over an extension of degree a power of 2.

Theorem 2.2. *Let F be a field of characteristic zero. For every integer $n \ge 15$, we have*

$$\mathrm{ed}_2(\mathbf{Spin}_n) = \mathrm{ed}(\mathbf{Spin}_n) = \begin{cases} 2^{(n-1)/2} - n(n-1)/2 & \text{if } n \text{ is odd},\\ 2^{(n-2)/2} - n(n-1)/2 & \text{if } n \equiv 2 \pmod 4,\\ 2^{(n-2)/2} + 2^m - n(n-1)/2 & \text{if } n \equiv 0 \pmod 4,\end{cases}$$

where 2^m is the largest power of 2 dividing n.

Proof. The case $n \ge 15$ and n not divisible by 4 has been considered in [Brosnan et al. 2010, Theorem 3-3].

Now assume that $n > 15$ and n is divisible by 4. The inequality "$\geq$" was obtained in [Merkurjev 2009, Theorem 4.9], so we just need to prove the inequality "$\leq$". The case $n = 16$ was considered in [Merkurjev 2009, Corollary 4.10]. Assume that $n \geq 20$ and n is divisible by 4.

Consider the following diagram with exact rows:

$$\begin{tikzcd}
1 \arrow[r] & \boldsymbol{\mu}_2 \arrow[r] \arrow[d, equal] & \mathbf{Spin}_n \arrow[r] \arrow[d] & \mathbf{Spin}_n^+ \arrow[r] \arrow[d] & 1 \\
1 \arrow[r] & \boldsymbol{\mu}_2 \arrow[r] & \mathbf{O}_n^+ \arrow[r] & \mathbf{PGO}_n^+ \arrow[r] & 1
\end{tikzcd}$$

where $\mathbf{Spin}_n^+$ is the semispinor group, $\mathbf{O}_n^+$ is the split special orthogonal group and $\mathbf{PGO}_n^+$ is the split special projective orthogonal group. We see from the diagram that the image of the connecting map

$$\delta_K : H^1_{\text{ét}}(K, \mathbf{Spin}_n^+) \to H^2_{\text{ét}}(K, \boldsymbol{\mu}_2) \subset \operatorname{Br}(K)$$

is contained in the image of the other connecting map

$$H^1_{\text{ét}}(K, \mathbf{PGO}_n^+) \to H^2_{\text{ét}}(K, \boldsymbol{\mu}_2) \subset \operatorname{Br}(K)$$

for every field extension K/F. The image of the last map consists of the classes $[A]$ of all central simple K-algebras A of degree n admitting orthogonal involutions (see [Knus et al. 1998, §31]). As $\operatorname{ind}(A)$ is a power of 2 dividing n, we have $\operatorname{ind}(A) \leq 2^m$, where 2^m is the largest power of 2 dividing n.

Let E be a $\mathbf{Spin}_n^+$-torsor over K. We have shown that, if $\delta_K([E]) = [A]$ for a central simple K-algebra A, then $\operatorname{ind}(A) \leq 2^m$. It follows from [Brosnan et al. 2011, Theorem 4.1] that $\operatorname{ed}[E/\mathbf{Spin}_n] = \operatorname{ind}(A) \leq 2^m$.

It is shown in [Brosnan et al. 2010, Remark 3-10] that

$$\operatorname{ed}(\mathbf{Spin}_n^+) = 2^{(n-2)/2} - \frac{n(n-1)}{2}$$

for every integer $n \geq 20$ divisible by 4. Finally, by Proposition 2.1,

$$\operatorname{ed}(\mathbf{Spin}_n) \leq \operatorname{ed}(\mathbf{Spin}_n^+) + 2^m = 2^{(n-2)/2} + 2^m - \frac{n(n-1)}{2}. \qquad \square$$

3. The functors I_n^k

We use the following notation. Let F be a field of characteristic different from 2 and K/F a field extension. We define

$$I_n^1(K) = \boxed{\begin{array}{l}\text{Set of isomorphism classes of nondegenerate}\\ \text{quadratic forms over } K \text{ of dimension } n\end{array}}$$

and recall from [Knus et al. 1998, §29.E] the existence of a natural bijection $I_n^1(K) \simeq H^1_{\text{ét}}(K, \mathbf{O}_n)$.

Recall that the *discriminant* $\operatorname{disc}(q)$ of a form $q \in I_n^1(K)$ is equal to

$$(-1)^{n(n-1)/2}\det(q) \in K^\times/K^{\times 2}.$$

Set

$$I_n^2(K) = \{q \in I_n^1(K) : \operatorname{disc}(q) = 1\}.$$

We have a natural bijection $I_n^2(K) \simeq H^1_{\text{ét}}(K, \mathbf{O}_n^+)$ (see [Knus et al. 1998, §29.E]).

The *Clifford invariant* $c(q)$ of a form $q \in I_n^2(K)$ is the class in the Brauer group $\operatorname{Br}(K)$ of the Clifford algebra of q if n is even and the class of the even Clifford algebra if n is odd [Knus et al. 1998, §8.B]. Define

$$I_n^3(K) = \{q \in I_n^2(K) : c(q) = 0\}.$$

Remark 3.1. Our notation of the functors I_n^k for $k = 1, 2, 3$ is explained by the following property: $I_n^k(K)$ consists of all classes of quadratic forms $q \in W(K)$ of dimension n such that $q \in I(K)^k$ if n is even and $q \perp \langle -1\rangle \in I(K)^k$ if n is odd, where $I(K)$ is the fundamental ideal in the Witt ring $W(K)$ of K.

The functor I_n^3 is related to $\mathbf{Spin}_n$-torsors as follows. The short exact sequence

$$1 \to \boldsymbol{\mu}_2 \to \mathbf{Spin}_n \to \mathbf{O}_n^+ \to 1$$

yields an exact sequence

$$H^1_{\text{ét}}(K, \boldsymbol{\mu}_2) \to H^1_{\text{ét}}(K, \mathbf{Spin}_n) \to H^1_{\text{ét}}(K, \mathbf{O}_n^+) \xrightarrow{c} H^2_{\text{ét}}(K, \boldsymbol{\mu}_2), \tag{1}$$

where c is the Clifford invariant. Thus, $\operatorname{Ker}(c) = I_n^3(K)$.

The essential dimensions of I_n^1 and I_n^2 were computed in [Reichstein 2000, Theorems 10.3 and 10.4]: we have $\operatorname{ed}(I_n^1) = n$ and $\operatorname{ed}(I_n^2) = n-1$. In Section 7, we compute $\operatorname{ed}(I_n^3)$. We will need the following lemma, which was proven in [Brosnan et al. 2010, Lemma 5-1]:

Lemma 3.2. *We have* $\operatorname{ed}_p(I_n^3) \le \operatorname{ed}_p(\mathbf{Spin}_n) \le \operatorname{ed}_p(I_n^3) + 1$ *for every* $p \ge 0$.

Proof. Let K/F be a field extension. The group $H^1_{\text{ét}}(K, \boldsymbol{\mu}_2) = K^\times/K^{\times 2}$ acts transitively on the fibers of the second map in the sequence (1). It follows that the natural map $\mathbf{Spin}_n$-*Torsors* $\to I_n^3$ is a surjection with $\mathbf{G}_m$ acting surjectively on the fibers. The statement follows from [Berhuy and Favi 2003, Proposition 1.13]. □

Let $\boldsymbol{\Gamma}_n^+$ be the split even Clifford group (see [Knus et al. 1998, §23]). The commutative diagram with exact rows and columns

$$\begin{tikzcd}
 & 1 \arrow[d] & 1 \arrow[d] & & \\
1 \arrow[r] & \boldsymbol{\mu}_2 \arrow[r] \arrow[d] & \mathbf{Spin}_n \arrow[r] \arrow[d] & \mathbf{O}_n^+ \arrow[r] \arrow[d, equal] & 1 \\
1 \arrow[r] & \mathbf{G}_m \arrow[r] \arrow[d, "2"'] & \mathbf{\Gamma}_n^+ \arrow[r] \arrow[d, "\text{spinor}"', "\text{norm}"] & \mathbf{O}_n^+ \arrow[r] & 1 \\
 & \mathbf{G}_m \arrow[r, equal] \arrow[d] & \mathbf{G}_m \arrow[d] & & \\
 & 1 & 1 & &
\end{tikzcd}$$

yields a bijection $H^1_{\text{ét}}(K, \mathbf{\Gamma}_n^+) \simeq I_n^3(K)$ for any field extension K/F (see [Knus et al. 1998, §28]). In particular, $\mathrm{ed}_p(\mathbf{\Gamma}_n^+) = \mathrm{ed}_p(I_n^3)$.

4. Subforms of forms in I_n^3

In this section, we study the following problem in quadratic form theory, which will be used in Section 7 in order to compute the essential dimension of I_n^3. Note that the problem is stated entirely in terms of quadratic forms while in the solution we use the essential dimension. We don't know how to solve the problem by means of quadratic form theory.

Problem 4.1. *Given a field F, determine all integers n such that every form in $I_n^3(K)$ contains a nontrivial subform in $I^2(K)$ for any field extension K/F.*

All forms in $I_n^3(K)$ for $n \le 14$ are classified (see [Garibaldi 2009, Example 17.8, Theorems 17.13 and 21.3]). Inspection shows that for such n the problem has positive solution.

In the following theorem, we show that in the range $n \ge 15$ the problem has negative solution (with possibly two exceptions):

Theorem 4.2. *Let F be a field of characteristic zero, let $n \ge 15$ and let b be an even integer with $0 < b < n$. Then there is a field extension K/F and a form in $I_n^3(K)$ that does not contain a subform in $I_b^2(K)$ (with possible exceptions $(n, b) = (15, 8)$ or $(16, 8)$).*

Let $a := n - b$. Write $H_{a,b}$ for the image of the natural homomorphism

$$\mathbf{Spin}_a \times \mathbf{Spin}_b \to \mathbf{Spin}_n. \tag{2}$$

Note that the kernel of (2) is contained in

$$\boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2 = \mathrm{Ker}(\mathbf{Spin}_a \times \mathbf{Spin}_b \to \mathbf{O}_a^+ \times \mathbf{O}_b^+)$$

and therefore is the cyclic group of order 2 generated by $(-1,-1)$. Hence, we have an exact sequence

$$1 \to \boldsymbol{\mu}_2 \to H_{a,b} \to \mathbf{O}_a^+ \times \mathbf{O}_b^+ \to 1$$

and therefore a map

$$H^1_{\text{ét}}(R, H_{a,b}) \to H^1_{\text{ét}}(R, \mathbf{O}_a^+ \times \mathbf{O}_b^+) = H^1_{\text{ét}}(R, \mathbf{O}_a^+) \times H^1_{\text{ét}}(R, \mathbf{O}_b^+)$$

for a commutative F-algebra R.

We write $q(\eta) := (q_a, q_b)$ for the image of an element $\eta \in H^1_{\text{ét}}(R, H_{a,b})$ under this map, where $q_a \in H^1_{\text{ét}}(R, \mathbf{O}_a^+)$ and $q_b \in H^1_{\text{ét}}(R, \mathbf{O}_b^+)$.

Consider the commutative diagram with the exact rows

$$\begin{tikzcd}
1 \arrow[r] & \boldsymbol{\mu}_2 \arrow[r] \arrow[d, equal] & H_{a,b} \arrow[r] \arrow[d] & \mathbf{O}_a^+ \times \mathbf{O}_b^+ \arrow[r] \arrow[d, "\tau"'] & 1 \\
1 \arrow[r] & \boldsymbol{\mu}_2 \arrow[r] & \mathbf{Spin}_n \arrow[r] & \mathbf{O}_n^+ \arrow[r] & 1
\end{tikzcd}$$

The image of an element $\xi \in H^1_{\text{ét}}(R, \mathbf{Spin}_n)$ in $H^1_{\text{ét}}(R, \mathbf{O}_n^+)$ will be denoted by $q(\xi)$.

If $\xi \in H^1_{\text{ét}}(R, \mathbf{Spin}_n)$ is the image of an element $\eta \in H^1_{\text{ét}}(R, H_{a,b})$, then $q(\xi) = q_a \perp q_b$, the image of $(q_a, q_b) = q(\eta)$ under the map induced by τ. We can reverse this statement as follows.

Lemma 4.3. *Let $\xi \in H^1_{\acute{e}t}(R, \mathbf{Spin}_n)$ with $q(\xi) = q_a \perp q_b$, where $q_a \in H^1_{\acute{e}t}(R, \mathbf{O}_a^+)$ and $q_b \in H^1_{\acute{e}t}(R, \mathbf{O}_b^+)$. Then ξ is the image of an element η under the map $H^1_{\acute{e}t}(R, H_{a,b}) \to H^1_{\acute{e}t}(R, \mathbf{Spin}_n)$ such that $q(\eta) = (q_a, q_b)$.*

Proof. The diagram above yields a commutative diagram with the exact rows

$$\begin{tikzcd}
H^1_{\text{ét}}(R, H_{a,b}) \arrow[r] \arrow[d] & H^1_{\text{ét}}(R, \mathbf{O}_a^+) \times H^1_{\text{ét}}(R, \mathbf{O}_b^+) \arrow[r, "c'"] \arrow[d] & H^2_{\text{ét}}(R, \boldsymbol{\mu}_2) \arrow[d, equal] \\
H^1_{\text{ét}}(R, \mathbf{Spin}_n) \arrow[r] & H^1_{\text{ét}}(R, \mathbf{O}_n^+) \arrow[r, "c"] & H^2_{\text{ét}}(R, \boldsymbol{\mu}_2)
\end{tikzcd}$$

Moreover, the group $H^1_{\text{ét}}(R, \boldsymbol{\mu}_2)$ acts transitively on the fibers of the left maps in the two rows. The result follows. □

For nonnegative integers a, b and a field extension K/F, set

$$I^3_{a,b}(K) := \{(q_a, q_b) \in I^2_a(K) \times I^2_b(K) : q_a \perp q_b \in I^3_n(K)\}.$$

Corollary 4.4. *For any $\eta \in H^1_{ét}(K, H_{a,b})$, we have $q(\eta) \in I^3_{a,b}(K)$. The morphism of functors $q : H_{a,b}$-Torsors $\to I^3_{a,b}$ is surjective. In particular,* $\mathrm{ed}_p(I^3_{a,b}) \le \mathrm{ed}_p(H_{a,b})$ *for every $p \ge 0$.*

Proof. Note that the map c' in the proof of Lemma 4.3 when $R = K$ takes a pair (q_a, q_b) to the Clifford invariant of $q_a \perp q_b$ in $\mathrm{Br}(K)$. The pair $(q_a, q_b) \in I^2_a(K) \times I^2_b(K)$ comes from $H^1_{\text{ét}}(K, H_{a,b})$ if and only if the Clifford invariant of $q_a \perp q_b$ is split, i.e., $q_a \perp q_b \in I^3_n(K)$. □

Lemma 4.5. *For an even a and any b,*

$$\mathrm{ed}_p(I^3_{a,b}) \le \mathrm{ed}_p(I^3_{a-1,b}) + 1$$

for every $p \ge 0$.

Proof. Consider the morphism of functors

$$\alpha : \mathbf{G}_m \times I^3_{a-1,b} \to I^3_{a,b}, \quad (\lambda; f, g) \mapsto (\lambda(f \perp \langle -1 \rangle), g).$$

Every form h in $I^2_a(K)$ can be written in the form $h = \lambda(f \perp \langle -1 \rangle)$ for a value λ of h and a form $f \in I^2_{a-1}(K)$; i.e., α is a surjection, whence the result. □

Write V_n and W_n for the (semi)spinor and regular representations, respectively, of the group $\mathbf{Spin}_n$. We have

$$\dim(V_n) = \begin{cases} 2^{(n-1)/2} & \text{if } n \text{ is odd,} \\ 2^{(n-2)/2} & \text{if } n \text{ is even} \end{cases}$$

and $\dim(W_n) = n$. We consider the tensor product $V_{a,b} := V_a \otimes V_b$ as the representation of the group $H_{a,b}$. We also view W_a and W_b as $H_{a,b}$-representations via the natural homomorphisms $H_{a,b} \to \mathbf{O}^+_a$ and $H_{a,b} \to \mathbf{O}^+_b$, respectively.

A representation V of an algebraic group H is *generically free* if the stabilizer of a generic vector in V is trivial. In this case, by [Reichstein and Youssin 2000],

$$\mathrm{ed}(H) \le \dim(V) - \dim(H).$$

Lemma 4.6. *Let a be odd and b even. Suppose that $V_{a,b}$ is a generically free representation of the image of the homomorphism $H_{a,b} \to \mathbf{GL}(V_{a,b})$. Then $V_{a,b} \oplus W_b$ is a generically free representation of $H_{a,b}$. In particular,*

$$\mathrm{ed}(H_{a,b}) \le \dim(V_{a,b}) + \dim(W_b) - \dim(H_{a,b}).$$

Proof. Write C_n for the kernel of $\mathbf{Spin}_n \to \mathbf{PGO}^+_n$ and C'_n for the kernel of $\mathbf{Spin}_n \to \mathbf{O}^+_n$, so $C'_n = \{\pm 1\} \subset C_n$. By assumption, the generic stabilizer H of the action of $\mathbf{Spin}_a \times \mathbf{Spin}_b$ on $V_{a,b}$ is contained in the center $C_a \times C_b$. Since $C_b / C'_b = \boldsymbol{\mu}_2$ acts on W_b by multiplication by -1, we have $H \subset C_a \times C'_b \simeq \boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$. Note that $\boldsymbol{\mu}_2 \times 1$ and $1 \times \boldsymbol{\mu}_2$ act by multiplication by -1 on $V_{a,b}$; hence, H is generated by $(-1, -1)$. It follows that $H_{a,b} = (\mathbf{Spin}_a \times \mathbf{Spin}_b)/H$ acts generically freely on $V_{a,b} \oplus W_b$. □

Proposition 4.7. *Let* $\mathrm{char}(F) = 0$. *If* $n = a + b \geq 15$ *with* $a \leq b$, *then* $V_{a,b}$ *is a generically free representation of the image of* $H_{a,b} \to \mathbf{GL}(V_{a,b})$ *if and only if* $(a, b) \neq (3, 12), (4, 11), (4, 12), (6, 10)$ *and* $(8, 8)$.

Proof. All the cases of infinite generic stabilizers H are listed in [Èlašvili 1972, §3, Row 7 of Table 6]: H is infinite if and only if $(a, b) = (3, 12)$ and $(4, 12)$.

If H is finite, by [Popov 1987, Theorem 1, Rows 1, 12 and 13 of Table 1], H is nontrivial if and only if $(a, b) = (4, 11)$, $(6, 10)$ and $(8, 8)$. □

Proof of Theorem 4.2. Note that the case (n, b) with n even implies the case $(n-1, b)$. Indeed, suppose that every form in I^3_{n-1} for an even n contains a subform from I^2_b. Take any form $q \in I^3_n(K)$ for a field extension K/F, and write $q = \lambda(f \perp \langle -1 \rangle)$ for a $\lambda \in K^\times$ and $f \in I^3_{n-1}(K)$. If f contains a subform $h \in I^2_b(K)$, then q contains λh.

We need to show that the natural morphism of functors $I^3_{a,b} \to I^3_n$ is not surjective. It suffices to prove that $\mathrm{ed}(I^3_{a,b}) < \mathrm{ed}(I^3_n)$. We may assume that n (and hence also a) is even. Moreover, we may assume that $a \leq b$.

Suppose that $n \geq 18$. By Proposition 4.7, Lemmas 4.5 and 4.6 and Corollary 4.4,

$$\begin{aligned}
\mathrm{ed}(I^3_{a,b}) &\leq \mathrm{ed}(I^3_{a-1,b}) + 1 \\
&\leq \mathrm{ed}(H_{a-1,b}) + 1 \\
&\leq \dim(V_{a-1,b}) + \dim(W_b) - \dim(H_{a-1,b}) + 1 \\
&= 2^{n/2-2} + b - (a-1)(a-2)/2 - b(b-1)/2 + 1 \\
&= 2^{n/2-2} - (a^2 + b^2 - 3a - 3b)/2 \\
&\leq 2^{n/2-2} - (n^2 - 6n)/4
\end{aligned}$$

as $a^2 + b^2 \geq n^2/2$. The last integer is strictly less than

$$2^{n/2-1} - n(n-1)/2 - 1 \leq \mathrm{ed}(\mathbf{Spin}_n) - 1 \leq \mathrm{ed}(I^3_n)$$

by Theorem 2.2 and Lemma 3.2.

It remains to consider the case $n = 16$. Note that, by Theorem 2.2 and Lemma 3.2,

$$\mathrm{ed}(I^3_{16}) \geq \mathrm{ed}(\mathbf{Spin}_{16}) - 1 = 23. \tag{3}$$

We shall prove that $\mathrm{ed}(I^3_{a,b}) < 23$. All possible values of b are 8, 10, 12 and 14.

***Case* $(n, b) = (16, 10)$.** Consider the representation $V := W_6 \oplus V_{6,10} \oplus W_{10}$ of $H_{6,10}$. We claim that V is generically free. The stabilizer in $\mathbf{Spin}_6$ of a point in general position in W_6 is $\mathbf{Spin}_5$. Hence, the stabilizer in $H_{6,10}$ of a point in general position in W_6 is $H_{5,10}$. Note that the restriction of $V_{6,10}$ to $H_{5,10}$ is isomorphic to $V_{5,10}$. Finally, the $H_{5,10}$-representation $V_{5,10} \oplus W_{10}$ is generically free by Proposition 4.7.

It follows from (3) and Corollary 4.4 that

$$\operatorname{ed}(I^3_{6,10}) \le \operatorname{ed}(H_{6,10}) \le \dim(V) - \dim(H_{6,10}) = 80 - 60 = 20.$$

***Case* $(n, b) = (16, 12)$.** Consider the representation $V := W_3 \oplus W_3 \oplus V_{3,12} \oplus W_{12}$ of $H_{3,12}$. We claim that V is generically free as the representation of $H_{3,12}$. Indeed, the stabilizer in $H_{3,12}$ of a generic vector in W_{12} is $H_{3,11}$. We are reduced to showing that $W_3 \oplus W_3 \oplus V_{3,11}$ is a generically free representation of $H_{3,11}$. By [Popov 1987, §5, p. 246], the generic stabilizer S of $H_{3,11}$ in $V_{3,11}$ is finite (isomorphic to $\boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$), and the restriction to S of the natural projection $H_{3,11} \to \mathbf{O}_3^+$ is injective. It remains to notice that the representation $W_3 \oplus W_3$ of $\mathbf{O}_3^+ = \mathbf{PGL}_2$ is generically free.

It follows from Lemmas 4.5 and 4.6 and Corollary 4.4 that

$$\begin{aligned}\operatorname{ed}(I^3_{4,12}) \le \operatorname{ed}(I^3_{3,12}) + 1 &\le \operatorname{ed}(H_{3,12}) + 1 \\ &\le \dim(V) - \dim(H_{3,12}) + 1 = 82 - 69 + 1 = 14.\end{aligned}$$

***Case* $(n, b) = (16, 14)$.** As every form in I_2^3 is hyperbolic, we have $I^3_{2,14} = I^3_{14}$ and $\operatorname{ed}(I^3_{14}) = 7$ by Theorem 2.2. □

5. Unramified principal homogeneous spaces

Let G be an algebraic group over F, and let K/F be a field extension with a discrete valuation v trivial on F. Write O for the valuation ring of v. It is a local F-algebra. We say that a class $\xi \in H^1_{\text{ét}}(K, G)$ is *unramified* (with respect to v) if ξ belongs to the image of the map $H^1_{\text{ét}}(O, G) \to H^1_{\text{ét}}(K, G)$.

Let $\bar{K}$ be the residue field of v. The ring homomorphism $O \to \bar{K}$ yields a map $H^1_{\text{ét}}(O, G) \to H^1_{\text{ét}}(\bar{K}, G)$. This map is a bijection if K is complete (see [SGA 3 1970, Exposé XXIV, Proposition 8.1]). Hence, we have the map

$$H^1_{\text{ét}}(\bar{K}, G) \xrightarrow{\sim} H^1_{\text{ét}}(O, G) \to H^1_{\text{ét}}(K, G). \tag{4}$$

Example 5.1. Let $\operatorname{char}(F) \ne 2$ and $G = \mathbf{O}_n$. Then $H^1_{\text{ét}}(K, G)$ is the set of isomorphism classes of nondegenerate quadratic forms of dimension n over K. A quadratic form q over a field K with a discrete valuation is unramified if and only if $q \simeq \langle a_1, a_2, \dots, a_n \rangle$, where a_i are units in the valuation ring O in K. In general, every q can be written $q = q_1 \perp \pi q_2 \perp h$, where π is a prime element, q_1 and q_2 are unramified anisotropic quadratic forms and h is a hyperbolic form. The form q is unramified if and only if $q_2 = 0$. It follows that, if two forms q and πq are both unramified, then q is hyperbolic. If K is complete, then the map (4) takes $f = \langle \bar{a}_1, \bar{a}_2, \dots, \bar{a}_n \rangle$ over $\bar{K}$, where a_i are units in O, to $f_K := \langle a_1, a_2, \dots, a_n \rangle$.

6. Essential dimension of PI_n^3

Two quadratic forms f and g over a field K are called *similar* if $f = \lambda g$ for some $\lambda \in K^\times$. If n is even, we write $PI_n^3(K)$ for the set of similarity classes of forms in $I_n^3(K)$. The group $K^\times$ acts transitively on the fibers of the natural surjective map $I_n^3(K) \to PI_n^3(K)$. Hence,

$$\mathrm{ed}_p(PI_n^3) \le \mathrm{ed}_p(I_n^3) \le \mathrm{ed}_p(PI_n^3) + 1$$

for any $p \ge 0$ by [Berhuy and Favi 2003, Proposition 1.13].

Proposition 6.1. *Let* $\mathrm{char}(F) \ne 2$. *For an even* $n \ge 8$, *and* $p = 0$ *or* 2, *we have*

$$\mathrm{ed}_p(PI_n^3) = \mathrm{ed}_p(I_n^3) - 1.$$

Proof. Let K/F be a field extension, and let $q \in I_n^3(K)$ be a nonhyperbolic form. Consider the form tq over the field $K((t))$. It suffices to show that

$$\mathrm{ed}_p^{I_n^3}(tq) \ge \mathrm{ed}_p^{PI_n^3}(q) + 1.$$

Let $M/K((t))$ be a finite field extension of degree prime to p (i.e., $M = K((t))$ if $p = 0$ and $[M : K((t))]$ is odd if $p = 2$), let L/F be a subextension of M/F and let $f \in I_n^3(L)$ be such that $\mathrm{tr.deg}(L/F) = \mathrm{ed}_p^{I_n^3}(tq)$ and $tq_M \simeq f_M$.

Let v be the (unique) extension on M of the discrete valuation of $K((t))$, and let w be the restriction of v on L. The residue field $\overline{M}$ is a finite extension of K of degree prime to p. As the form q is not hyperbolic, q_M is not hyperbolic, and therefore, the form $tq_M \simeq f_M$ is ramified by Example 5.1. It follows that w is nontrivial, i.e., w is a discrete valuation on L.

Let $\hat{L}$ be the completion of L. Note that, as M is complete, we can identify $\hat{L}$ with a subfield of M. Write $f_{\hat{L}} \simeq (f_1)_{\hat{L}} \perp \pi(f_2)_{\hat{L}}$, where f_1 and f_2 are quadratic forms over the residue field $\bar{L}$ and $\pi \in L$ is a prime element (see Example 5.1). Note that $f_1, f_2 \in I^2(\bar{L})$ by [Elman et al. 2008, Lemma 19.4]. If the ramification index e of M/L is even, then π is a unit in the valuation ring O of M modulo squares in $M^\times$; hence, f_M is unramified, a contradiction. It follows that e is odd. Writing $\pi = ut^e$ with a unit $u \in O^\times$, we have

$$tq_M \simeq f_M \simeq (f_1)_M \perp \pi(f_2)_M \simeq (f_1)_M \perp ut(f_2)_M;$$

hence, $(f_1)_M = 0$ and $q_M = u(f_2)_M$ in $W(M)$. It follows that $(f_1)_{\overline{M}} = 0$ and $q_{\overline{M}} = \bar{u}(f_2)_{\overline{M}}$ in $W(\overline{M})$, and therefore,

$$q_{\overline{M}} = \bar{u}(f_2)_{\overline{M}} = \bar{u}g_{\overline{M}}, \tag{5}$$

where $g := f_1 \perp f_2$ is the form over $\bar{L}$ of dimension n. Note that $f_{\hat{L}} - g_{\hat{L}} = \langle \pi, -1\rangle (f_2)_{\hat{L}} \in I^3(\hat{L})$; hence, $g_{\hat{L}} \in I^3(\hat{L})$ and $g \in I^3(\bar{L})$.

It follows from (5) that $q_{\overline{M}}$ is similar to $g_{\overline{M}}$, i.e., the form q is p-defined over $\bar{L}$ for the functor PI_n^3 (see [Merkurjev 2009, §1.1]), and therefore,

$$\operatorname{ed}_p^{I_n^3}(tq) = \operatorname{tr.deg}(L/F) \geq \operatorname{tr.deg}(\bar{L}/F) + 1 \geq \operatorname{ed}_p^{PI_n^3}(q) + 1. \qquad \square$$

7. Essential dimension of $\mathbf{\Gamma}_n^+$

In this section, we compute the essential dimension of $\mathbf{\Gamma}_n^+$ and I_n^3.

Theorem 7.1. *Let F be a field of characteristic zero. Then for every integer $n \geq 15$ and $p = 0$ or 2, we have*

$$\operatorname{ed}_p(\mathbf{\Gamma}_n^+) = \operatorname{ed}_p(I_n^3) = \begin{cases} 2^{(n-1)/2} - 1 - n(n-1)/2 & \text{if } n \text{ is odd,} \\ 2^{(n-2)/2} - n(n-1)/2 & \text{if } n \equiv 2 \pmod 4, \\ 2^{(n-2)/2} + 2^m - 1 - n(n-1)/2 & \text{if } n \equiv 0 \pmod 4, \end{cases}$$

where 2^m is the largest power of 2 dividing n.

If $\operatorname{char}(F) \neq 2$, then the essential dimension of I_n^3 has the following values for $n \leq 14$:

n	≤ 6	7	8	9	10	11	12	13	14
$\operatorname{ed}_2(I_n^3) = \operatorname{ed}(I_n^3)$	0	3	4	4	4	5	6	6	7

Proof. We will prove the theorem case by case.

Case $n \equiv 2 \pmod 4$ and $n \geq 10$. The exact sequence

$$1 \to \boldsymbol{\mu}_4 \to \mathbf{Spin}_n \to \mathbf{PGO}_n^+ \to 1$$

yields a surjective map $\mathbf{Spin}_n$-*Torsors*$(K) \to PI_n^3(K)$ for any K/F, with the group $K^\times$ acting transitively on the fibers of this map. It follows from Theorem 2.2, Proposition 6.1 and Lemma 3.2 that

$$\operatorname{ed}_2(I_n^3) = \operatorname{ed}_2(PI_n^3) + 1 \geq \operatorname{ed}_2(\mathbf{Spin}_n) = \operatorname{ed}(\mathbf{Spin}_n) \geq \operatorname{ed}(I_n^3) \geq \operatorname{ed}_2(I_n^3).$$

Hence, $\operatorname{ed}_2(I_n^3) = \operatorname{ed}(I_n^3) = \operatorname{ed}(\mathbf{Spin}_n)$. The latter value is known by Theorem 2.2.

Case $n \not\equiv 2 \pmod 4$ and $n \geq 15$. Let $n = a + b$ with even $b \neq 2$. Let Z be the trivial group if $b = 0$ and the image of the center C_b of $\mathbf{Spin}_b$ in $H_{a,b}$ if $b \geq 4$. Then Z is central in $H_{a,b}$; hence, the group $H^1_{\text{ét}}(K, Z)$ acts on $H^1_{\text{ét}}(K, H_{a,b})$.

Lemma 7.2. *Let $\xi, \eta \in H^1_{\acute{e}t}(K, H_{a,b})$ with even $b \neq 2$. Suppose that $q(\xi) = q_a \perp q_b$ and $q(\eta) = q_a \perp \lambda q_b$ with the forms $q_a \in I_a^2(K)$ and $q_b \in I_b^2(K)$ and $\lambda \in K^\times$. Then $\eta = \alpha\xi$ for some $\alpha \in H^1_{\acute{e}t}(K, Z)$.*

Proof. The statement is trivial if $b = 0$, so assume that $b \geq 4$. The restriction of the natural homomorphism $H_{a,b} \to \mathbf{O}_b^+$ to the subgroup Z yields a surjection

$\varphi: Z \to \boldsymbol{\mu}_2 = \mathrm{Center}(\mathbf{O}_b^+)$. The kernel of φ coincides with the kernel C of the canonical homomorphism $H_{a,b} \to \mathbf{O}_a^+ \times \mathbf{O}_b^+$.

As Z is isomorphic to $\boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$ or $\boldsymbol{\mu}_4$, the homomorphism $\varphi^*: H^1_{\text{ét}}(K, Z) \to H^1_{\text{ét}}(K, \boldsymbol{\mu}_2) = K^\times/K^{\times 2}$ is surjective. Let $\gamma \in H^1_{\text{ét}}(K, Z)$ be such that $\varphi^*(\gamma) = \lambda K^{\times 2}$. Then $q(\gamma\xi) = q_a \perp \lambda q_b = q(\eta)$. Then there is $\beta \in H^1_{\text{ét}}(K, C)$ such that $\eta = \beta(\gamma\xi)$. Hence, $\eta = \alpha\xi$, where $\alpha = \beta'\gamma$ with β' the image of β under the map $H^1_{\text{ét}}(K, C) \to H^1_{\text{ét}}(K, Z)$ induced by the inclusion of C into Z. $\square$

Let $\xi \in H^1_{\text{ét}}(K, \mathbf{Spin}_n)$ be such that the form $q = q(\xi) \in I^3_n(K)$ is generic for the functor I^3_n (see [Merkurjev 2009, §2.2]). In particular, $\mathrm{ed}^{I^3_n}(q) = \mathrm{ed}(I^3_n)$. Note that q is anisotropic.

Identifying $\boldsymbol{\mu}_2$ with the kernel of $\mathbf{Spin}_n \to \mathbf{O}_n^+$, we have an action of $H^1_{\text{ét}}(E, \boldsymbol{\mu}_2) = E^\times/E^{\times 2}$ on $H^1_{\text{ét}}(E, \mathbf{Spin}_n)$, where $E = K((t))$. Consider the element $t\xi_E \in H^1_{\text{ét}}(E, \mathbf{Spin}_n)$ over E. We claim that $t\xi_E$ is ramified. Suppose not, i.e., $t\xi_E$ comes from an element $\rho \in H^1_{\text{ét}}(O, \mathbf{Spin}_n)$, where $O = K[[t]]$. Let $q' \in H^1_{\text{ét}}(O, \mathbf{O}_n^+)$ be the image of ρ viewed as a quadratic form over O. We have

$$q'_E = q(t\xi_E) = q(\xi_E) = q_E;$$

hence, $q' = q_O$. Then ρ and ξ_O belong to the same fiber of the map

$$H^1_{\text{ét}}(O, \mathbf{Spin}_n) \to H^1_{\text{ét}}(O, \mathbf{O}_n^+).$$

As the group $H^1_{\text{ét}}(O, \boldsymbol{\mu}_2) = O^\times/O^{\times 2}$ acts transitively on the fiber, there is a unit $u \in O^\times$ satisfying $t\xi_E = u\xi_E$. It follows from [Knus et al. 1998, Proposition 28.11] that tu^{-1} is in the image spinor norm map

$$\mathbf{O}^+(q_E) \to H^1_{\text{ét}}(E, \boldsymbol{\mu}_2) = E^\times/E^{\times 2}$$

for the form q_E; hence, q is isotropic by [Elman et al. 2008, Theorem 18.3], a contradiction. The claim is proven.

Let L/F be a subextension of E/F, and let $\eta \in H^1_{\text{ét}}(L, \mathbf{Spin}_n)$ be such that $\mathrm{tr.deg}(L/F) = \mathrm{ed}^{\mathbf{Spin}_n}(t\xi)$ and $\eta_E \simeq t\xi_E$. We have $q(\eta)_E = q(t\xi) = q(\xi_E) = q_E$; hence, the form $q(\eta)_E$ is anisotropic.

Let v be the restriction on L of the discrete valuation of E. As $t\xi$ is ramified, v is nontrivial; hence, v is a discrete valuation. Let $\pi \in L$ be a prime element.

Consider the completion $\hat{L}$ of L. As E is complete, we can view $\hat{L}$ as a subfield of E. Write $q(\eta_{\hat{L}}) = (q_a)_{\hat{L}} \perp \pi(q_b)_{\hat{L}}$, where q_a and q_b are anisotropic quadratic forms over the residue field $\bar{L}$ of dimension a and b, respectively. As $q(\eta) \in I^3(\hat{L})$, we have $q_b \in I^2(\bar{L})$, and therefore, b is even and $b \neq 2$. By Lemma 4.3, there is $\eta' \in H^1_{\text{ét}}(\hat{L}, H_{a,b})$ that maps to η with $q(\eta') = ((q_a)_{\hat{L}}, \pi(q_b)_{\hat{L}})$.

We claim that the ramification index e of the extension $E/\hat{L}$ is odd. Suppose e is even. Note that $q_a \perp q_b \in I^3_n(\bar{L})$. Lemma 4.3 allows us to choose an unramified element $\nu \in H^1_{\text{ét}}(\hat{L}, H_{a,b})$ with $q(\nu) = ((q_a)_{\hat{L}}, (q_b)_{\hat{L}})$. By Lemma 7.2, there is $\alpha \in H^1_{\text{ét}}(\hat{L}, Z)$ such that $\eta' = \alpha\nu$. If b is divisible by 4, we have $Z \simeq \boldsymbol{\mu}_2 \times \boldsymbol{\mu}_2$. As e is even, α is unramified over E; hence, η'_E is unramified. It follows that $\eta_E \simeq t\xi$ is also unramified, a contradiction.

Suppose that $b \equiv 2 \pmod 4$. Note that $0 < b < n$ since $n \not\equiv 2 \pmod 4$. Write $\pi = ut^k$ with a unit $u \in O^\times$ and even k. Then

$$(q_a \perp uq_b)_E \simeq (q_a \perp \pi q_b)_E \simeq q(\eta_E) \simeq q(t\xi_E) = q(\xi_E) = q_E.$$

It follows that $q \simeq (q_a)_K \perp (\bar{u}q_b)_K$, i.e., q contains the subform $(\bar{u}q_b)_K$ in $I^2(K)$ of dimension b. This contradicts Theorem 4.2. The claim is proven.

Thus, e is odd. We have

$$(q_a \perp utq_b)_E \simeq (q_a \perp \pi q_b)_E \simeq q(\eta_E) \simeq q(t\xi_E) = q(\xi_E) = q_E.$$

It follows that $(q_b)_K$ is hyperbolic and hence $(q_a \perp q_b)_K = (q_a)_K = q$ in $W(K)$, i.e., $(q_a \perp q_b)_K \simeq q$.

Note that $(q_a)_{\hat{L}} = (q_a)_{\hat{L}} + \pi(q_b)_{\hat{L}} = q(\eta_{\hat{L}}) \in I^3(\hat{L})$; hence, $q_a \in I^3(\bar{L})$ and $q_a \perp q_b \in I^3_n(\bar{L})$. Therefore, q is defined over $\bar{L}$ for the functor I^3_n; hence,

$$\mathrm{ed}^{\mathbf{Spin}_n}(t\xi) = \mathrm{tr.deg}(L/F) \geq \mathrm{tr.deg}(\bar{L}/F) + 1 \geq \mathrm{ed}^{I^3_n}(q) + 1 = \mathrm{ed}(I^3_n) + 1.$$

It follows that $\mathrm{ed}(\mathbf{Spin}_n) \geq \mathrm{ed}(I^3_n) + 1$; hence, $\mathrm{ed}(I^3_n) = \mathrm{ed}(\mathbf{Spin}_n) - 1$ by Lemma 3.2. The value of $\mathrm{ed}(\mathbf{Spin}_n)$ is given in Theorem 2.2.

In what follows, we use the following observation (see [Berhuy and Favi 2003]): if a functor $\mathcal{F}$ admits a nontrivial cohomological invariant of degree d with values in $\mathbb{Z}/2\mathbb{Z}$, then $\mathrm{ed}_2(\mathcal{F}) \geq d$.

Case n = 7. Every form q in $I^3_7(K)$ is the pure subform of a 3-fold Pfister form $\langle\langle a, b, c\rangle\rangle$; hence, $\mathrm{ed}(I^3_7) \leq 3$. On the other hand, the Arason invariant $e_3(q \perp \langle -1\rangle) = (a) \cup (b) \cup (c) \in H^3(K, \mathbb{Z}/2\mathbb{Z})$ is nontrivial (see [Garibaldi 2009, §18.6]); hence, $\mathrm{ed}_2(I^3_7) \geq 3$.

Case n = 8. Every form q in $I^3_8(K)$ is a multiple $e\langle\langle a, b, c\rangle\rangle$ of a 3-fold Pfister form; hence, $\mathrm{ed}(I^3_8) \leq 4$. The invariant $a_4(q) = (e) \cup (a) \cup (b) \cup (c) \in H^4(K, \mathbb{Z}/2\mathbb{Z})$ is nontrivial; hence, $\mathrm{ed}_2(I^3_8) \geq 4$.

Case n = 9 and 10. Every form q in $I^3_9(K)$ or $I^3_{10}(K)$ is equal to $f \perp \langle 1\rangle$ or $f \perp \langle 1, -1\rangle$, respectively, where f is a multiple of a 3-fold Pfister form over K, by [Lam 2005, XII.2.8]. Hence, $I^3_8 \simeq I^3_9 \simeq I^3_{10}$.

Case n = 11. The degree-5 cohomological invariant a_5 of $\mathbf{Spin}_{11}$ defined in [Garibaldi 2009, §20.8] factors through a nontrivial invariant of I^3_{11}; hence $\mathrm{ed}_2(I^3_{11}) \geq 5$. On the other hand, $\mathrm{ed}(I^3_{11}) \leq \mathrm{ed}(\mathbf{Spin}_{11}) = 5$.

Case $n = 12$. The degree-6 cohomological invariant a_6 of $\mathbf{Spin}_{12}$ defined in [Garibaldi 2009, §20.13] factors through a nontrivial invariant of I^3_{12}, so $\mathrm{ed}_2(I^3_{12}) \geq 6$. On the other hand, $\mathrm{ed}(I^3_{12}) \leq \mathrm{ed}(\mathbf{Spin}_{12}) = 6$.

Case $n = 13$ and 14. We know from the beginning of the proof (case $n \equiv 2 \pmod 4$ and $n \geq 10$) and from Theorem 2.2 that $\mathrm{ed}_2(I^3_{14}) = \mathrm{ed}(I^3_{14}) = \mathrm{ed}(\mathbf{Spin}_{14}) = 7$. By Lemma 4.5, $\mathrm{ed}_2(I^3_{13}) = \mathrm{ed}_2(I^3_{13,0}) \geq \mathrm{ed}_2(I^3_{14,0}) - 1 = 6$. On the other hand, $\mathrm{ed}(I^3_{13}) \leq \mathrm{ed}(\mathbf{Spin}_{13}) = 6$. □

References

[Berhuy and Favi 2003] G. Berhuy and G. Favi, "Essential dimension: a functorial point of view (after A. Merkurjev)", *Doc. Math.* **8** (2003), 279–330. MR 2004m:11056 Zbl 1101.14324

[Brosnan et al. 2010] P. Brosnan, Z. Reichstein, and A. Vistoli, "Essential dimension, spinor groups, and quadratic forms", *Ann. of Math.* (2) **171**:1 (2010), 533–544. MR 2011f:11053 Zbl 1252.11034

[Brosnan et al. 2011] P. Brosnan, Z. Reichstein, and A. Vistoli, "Essential dimension of moduli of curves and other algebraic stacks", *J. Eur. Math. Soc.* **13**:4 (2011), 1079–1112. MR 2012g:14012 Zbl 1234.14003

[Èlašvili 1972] A. G. Èlašvili, "Stationary subalgebras of points of the common state for irreducible linear Lie groups", *Funkts. Anal. Prilozh.* **6**:2 (1972), 65–78. In Russian; translated in *Funct. Anal. Appl.* **6**:2 (1972), 139–148. MR 46 #3690 Zbl 0252.22016

[Elman et al. 2008] R. Elman, N. Karpenko, and A. Merkurjev, *The algebraic and geometric theory of quadratic forms*, American Mathematical Society Colloquium Publications **56**, American Mathematical Society, Providence, RI, 2008. MR 2009d:11062 Zbl 1165.11042

[Garibaldi 2009] S. Garibaldi, "Cohomological invariants: exceptional groups and spin groups", *Mem. Amer. Math. Soc.* **200**:937 (2009). MR 2010g:20079 Zbl 1191.11009

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, American Mathematical Society Colloquium Publications **44**, American Mathematical Society, Providence, RI, 1998. MR 2000a:16031 Zbl 0955.16001

[Lam 2005] T. Y. Lam, *Introduction to quadratic forms over fields*, Graduate Studies in Mathematics **67**, American Mathematical Society, Providence, RI, 2005. MR 2005h:11075 Zbl 1068.11023

[Lötscher 2013] R. Lötscher, "A fiber dimension theorem for essential and canonical dimension", *Compos. Math.* **149**:1 (2013), 148–174. MR 3011881 Zbl 1260.14059

[Merkurjev 2009] A. S. Merkurjev, "Essential dimension", pp. 299–325 in *Quadratic forms — algebra, arithmetic, and geometry*, edited by R. Baeza et al., Contemp. Math. **493**, Amer. Math. Soc., Providence, RI, 2009. MR 2010i:14014 Zbl 1188.14006

[Popov 1987] A. M. Popov, "Finite isotropy subgroups in general position of irreducible semisimple linear Lie groups", *Tr. Mosk. Mat. Obs.* **50** (1987), 209–248. In Russian; translated in *Trans. Mosc. Math. Soc.* **1988** (1988), 205–249. MR 89a:20049 Zbl 0661.22009

[Reichstein 2000] Z. Reichstein, "On the notion of essential dimension for algebraic groups", *Transform. Groups* **5**:3 (2000), 265–304. MR 2001j:20073 Zbl 0981.20033

[Reichstein and Youssin 2000] Z. Reichstein and B. Youssin, "Essential dimensions of algebraic groups and a resolution theorem for G-varieties", *Canad. J. Math.* **52**:5 (2000), 1018–1056. MR 2001k:14088 Zbl 1044.14023

[SGA 3 1970] M. Demazure and A. Grothendieck (editors), *Schémas en groupes, III: Structure des schémas en groupes réductifs* (Séminaire de Géométrie Algébrique du Bois Marie, 1962–1964), Lecture Notes in Mathematics **153**, Springer, Berlin, 1970. MR 43 #223c Zbl 0212.52810

[Vistoli 2005] A. Vistoli, “Grothendieck topologies, fibered categories and descent theory”, pp. 1–104 in *Fundamental algebraic geometry*, Math. Surveys Monogr. **123**, Amer. Math. Soc., Providence, RI, 2005. MR 2223406 Zbl 1085.14001

Communicated by Raman Parimala
Received 2013-03-27 Revised 2013-05-25 Accepted 2013-06-24

chernous@math.ualberta.ca *Department of Mathematical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada*

merkurev@math.ucla.edu *Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095-1555, United States*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.473

On Deligne's category $\underline{\mathrm{Rep}}^{ab}(S_d)$

Jonathan Comes and Victor Ostrik

Dedicated to the memory of Andrei Zelevinsky

We prove a universal property of Deligne's category $\underline{\mathrm{Rep}}^{ab}(S_d)$. Along the way, we classify tensor ideals in the category $\underline{\mathrm{Rep}}(S_d)$.

1. Introduction

1A. Let F be a field of characteristic zero and let I be a finite set. Let S_I be the symmetric group of the permutations of I and let $\mathrm{Rep}(S_I)$ be the category of finite-dimensional F-linear representations of S_I considered as a symmetric tensor category. Let $X_I \in \mathrm{Rep}(S_I)$ be the space of F-valued functions on I with an obvious action of S_I. The object X_I with pointwise operations has a natural structure of associative commutative algebra with unit 1_{X_I} in the category $\mathrm{Rep}(S_I)$. We have a morphism $\mathrm{Tr} : X_I \to F$ defined as a trace of the operator of left multiplication; clearly the map $X_I \otimes X_I \to F$ given by $x \otimes y \mapsto \mathrm{Tr}(xy)$ is a nondegenerate pairing. Finally, $\mathrm{Tr}(1_{X_I}) = \dim(X_I) = |I|$ where $|I| \geq 0$ is the cardinality of I.

Now let G be a finite group acting on d-dimensional associative commutative unital algebra T over F such that the pairing $\mathrm{Tr}(xy)$ is nondegenerate. It is easy to see[1] that there exists a finite set I with $|I| = d$ and an essentially unique tensor functor $F : \mathrm{Rep}(S_I) \to \mathrm{Rep}(G)$ such that $F(X_I) \simeq T$ (isomorphism of G-algebras); in this sense the category $\mathrm{Rep}(S_I)$ is a universal category (in the realm of representation categories of finite groups) with object X_I as above.

1B. For an arbitrary symmetric tensor category $\mathcal{T}$ one can consider objects $T \in \mathcal{T}$ satisfying the following:

(a) T has a structure of associative commutative algebra (given by the multiplication map $\mu_T : T \otimes T \to T$) with unit (given by the map $1_T : \mathbf{1} \to T$).

MSC2010: primary 18D10; secondary 19D23.
Keywords: tensor categories, symmetric group.

[1]Set I to be the set of F-algebra homomorphisms $T \to \bar{F}$ where $\bar{F}$ is an algebraic closure of F and use an obvious homomorphism $G \to S_I$.

(b) The object T is rigid. Moreover, if we define the map $\mathrm{Tr} : T \to \mathbf{1}$ as the composition

$$T \xrightarrow{\mathrm{id}_T \otimes \mathrm{coev}_T} T \otimes T \otimes T^* \xrightarrow{\mu_T \otimes \mathrm{id}_{T^*}} T \otimes T^* \simeq T^* \otimes T \xrightarrow{\mathrm{ev}_T} \mathbf{1},$$

then the pairing $T \otimes T \xrightarrow{\mu_T} T \xrightarrow{\mathrm{Tr}} \mathbf{1}$ is nondegenerate, that is, it corresponds to an isomorphism $T \simeq T^*$ under the identification $\mathrm{Hom}(T \otimes T, \mathbf{1}) = \mathrm{Hom}(T, T^*)$.

(c) We have $\dim(T) = t \in F$ (equivalently, $\mathrm{Tr}(1_T) = t$).

For an arbitrary $t \in F$, Deligne [2007] defined a symmetric tensor category $\underline{\mathrm{Rep}}(S_t)$ with a distinguished object X which is universal in the following sense:

Proposition 1.1 [Deligne 2007, Proposition 8.3]. *Let $\mathcal{T}$ be a Karoubian symmetric tensor category over F. The functor $\mathcal{F} \mapsto \mathcal{F}(X)$ is an equivalence of the category of braided tensor functors $\underline{\mathrm{Rep}}(S_t) \to \mathcal{T}$ with the category of objects $T \in \mathcal{T}$ satisfying* (a), (b), (c) *above and their isomorphisms.*

Note that for $t = d \in \mathbb{Z}_{\geq 0}$, Proposition 1.1 applied to $T = X_I$ (with $|I| = d$) produces a canonical functor $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_d)$ (where $S_d := S_I$). It is known (see [Deligne 2007, Théorème 6.2]) that this functor is surjective on Hom's. Moreover, the morphisms sent to zero by this functor are precisely the so-called negligible morphisms (see [Deligne 2007, §6.1]).

1C. The category $\underline{\mathrm{Rep}}(S_t)$ is a Karoubian category; it is not abelian for $t = d \in \mathbb{Z}_{\geq 0}$. Remarkably, in [2007, Proposition 8.19] Deligne defined an *abelian* symmetric tensor category $\underline{\mathrm{Rep}}^{ab}(S_d)$ and a fully faithful braided tensor functor $\underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}^{ab}(S_d)$.[2] The main goal of this paper is to prove a certain universal property of the category $\underline{\mathrm{Rep}}^{ab}(S_d)$ conjectured in [Deligne 2007, Conjecture 8.21].

To state this property we need to use the language of algebraic geometry *within* an abelian symmetric tensor category $\mathcal{T}$ (see [Deligne 1990]). Namely, for an object $T \in \mathcal{T}$ satisfying (a), (b), (c) above we can talk about the (affine) $\mathcal{T}$-scheme $\mathbf{I} := \mathrm{Spec}(T)$ and the affine group scheme $S_{\mathbf{I}}$ of its automorphisms; see [Deligne 2007, §8.10]. Furthermore, assume that the category $\mathcal{T}$ is *pre-Tannakian* (see Section 2A below), that is, it satisfies the finiteness conditions in [Deligne 1990, 2.12.1]. Recall that in this case a *fundamental group* of $\mathcal{T}$ is defined in [Deligne 1990, §8.13]. This is an affine group scheme $\pi \in \mathcal{T}$ which acts functorially on any object of $\mathcal{T}$ and this action is compatible with a formation of tensor products. In particular, the action of π on T gives a homomorphism $\varepsilon : \pi \to S_{\mathbf{I}}$. Let $\mathrm{Rep}(S_{\mathbf{I}})$ be the category of representations of $S_{\mathbf{I}}$ (see [Deligne 2007, §8.10]) and let $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ be the full subcategory of $\mathrm{Rep}(S_{\mathbf{I}})$ consisting of such representations $\rho : S_{\mathbf{I}} \to \mathrm{GL}(V)$ that the action $\rho \circ \varepsilon$ of π on V coincides with the canonical action (see [Deligne 2007,

[2]We refer the reader to [Deligne 2007, §5.8] for an example of Karoubian symmetric tensor category which admits no braided tensor functor to an abelian symmetric tensor category.

§8.20]). $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ is an abelian symmetric tensor category and T is one of its objects. It follows that the functor $\mathcal{F} : \underline{\mathrm{Rep}}(S_t) \to \mathcal{T}$ constructed in Proposition 1.1 factorizes as $\underline{\mathrm{Rep}}(S_t) \xrightarrow{\mathcal{F}_T} \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon) \to \mathcal{T}$, where the functor $\mathcal{F}_T$ is constructed by applying Proposition 1.1 to $T \in \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ and $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon) \to \mathcal{T}$ is the forgetful functor. Here is the main result of this paper:

Theorem 1.2 (compare [Deligne 2007, 8.21.2]). *Let $\mathcal{T}$ be a pre-Tannakian category and $T \in \mathcal{T}$ be an object satisfying* (a), (b), (c) *from Section 1B with $t = d \in \mathbb{Z}_{\geq 0} \subset F$. Then the category* $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ *endowed with the functor* $\mathcal{F}_T : \underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ *is equivalent to one of the following*:

(a) $\mathrm{Rep}(S_d)$ *together with the functor* $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_d)$ *from Section 1B.*

(b) $\underline{\mathrm{Rep}}^{ab}(S_d)$ *together with the fully faithful functor* $\underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}^{ab}(S_d)$ *above.*

Remark 1.3. A similar (and easier) statement holds true for $t \notin \mathbb{Z}_{\geq 0}$; see [Deligne 2007, Corollary B2].

1D. The forgetful functor $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon) \to \mathcal{T}$ above is an exact braided tensor functor. Thus Theorem 1.2 implies that for a pre-Tannakian category $\mathcal{T}$ a braided tensor functor $\mathcal{F} : \underline{\mathrm{Rep}}(S_d) \to \mathcal{T}$ either factorizes through $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_d)$ or extends to an exact tensor functor $\underline{\mathrm{Rep}}^{ab}(S_d) \to \mathcal{T}$. A crucial step in our proof of Theorem 1.2 is a construction of the pre-Tannakian category $\mathcal{K}_d^0$ and fully faithful embedding $\underline{\mathrm{Rep}}(S_d) \subset \mathcal{K}_d^0$ such that we have the following *extension property*: a tensor (not necessarily braided) functor $\underline{\mathrm{Rep}}(S_d) \to \mathcal{T}$ either factorizes through $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_d)$ or extends to an exact tensor functor $\mathcal{K}_d^0 \to \mathcal{T}$; see Section 5A. Then we use general properties of the fundamental groups from [Deligne 1990, §8] in order to prove that $\mathcal{K}_d^0$ satisfies the universal property as in Theorem 1.2 and, in fact, is equivalent to $\underline{\mathrm{Rep}}^{ab}(S_d)$.

The following analogy plays a significant role in the proof of Theorem 1.2. Let $TL(q)$ be the *Temperley–Lieb category*; see, for example, [Freedman 2003, §A1]. Assume that q is a nontrivial root of unity. It is well known that the category $TL(q)$ is tensor equivalent to the category of *tilting modules* over quantum SL(2); see, for example, [Ostrik 2008, proof of Theorem 2.4]. Thus $TL(q)$ is a Karoubian tensor category (braided but not symmetric) endowed with a fully faithful functor to the abelian tensor category $\mathcal{C}_q$ of finite-dimensional representations of quantum SL(2). On the other hand there exists a well known semisimple tensor category $\bar{\mathcal{C}}_q$ and a full tensor functor $TL(q) \twoheadrightarrow \bar{\mathcal{C}}_q$; see, e.g., [Andersen 1992, §4]. We consider the diagram $\bar{\mathcal{C}}_q \twoheadleftarrow TL(q) \subset \mathcal{C}_q$ as a counterpart of the diagram $\mathrm{Rep}(S_d) \twoheadleftarrow \underline{\mathrm{Rep}}(S_d) \subset \underline{\mathrm{Rep}}^{ab}(S_d)$.

The main technical result of [Ostrik 2008] states that tensor functors $TL(q) \to \mathcal{D}$ to certain abelian tensor categories $\mathcal{D}$ factorize either through $TL(q) \to \bar{\mathcal{C}}_q$ or through $TL(q) \subset \mathcal{C}_q$ (see [Ostrik 2008, §2.6]) which is reminiscent of the extension property of the category $\mathcal{K}_d^0$ above; see also [Ostrik 2008, Remark 2.10]. Thus in the construction of $\mathcal{K}_d^0$ we follow the strategy from [Ostrik 2008] with crucial use of information from [Comes and Ostrik 2011]. Namely, we find $\mathcal{K}_d^0$ inside the homotopy category of $\underline{\mathrm{Rep}}(S_d)$ as a heart of a suitable *t-structure* (see Section 4B). The definition of the t-structure is based on Lemma 3.11 (due to P. Deligne) and almost immediately implies the extension property of the category $\mathcal{K}_d^0$ mentioned above. However, the verification of the axioms of a t-structure is quite nontrivial. To do this we use a decomposition of the category $\underline{\mathrm{Rep}}(S_d)$ into *blocks* described in [Comes and Ostrik 2011, Theorem 5.3]. We provide a blockwise description of the t-structure above in Section 4C2. We then observe that the description above coincides with the description of a well known t-structure on the blocks of the Temperley–Lieb category.

2. Preliminaries

2A. *Tensor categories terminology.* To us a *tensor* (*or monoidal*) *category* is a category with a tensor product functor endowed with an associativity constraint and a unit object $\mathbf{1}$; see, for example, [Bakalov and Kirillov 2001, Definition 1.1.7]. Recall that a tensor category is called *rigid* if any object admits both a left and right dual; see [ibid., Definition 2.1.1]. A *braided tensor category* is a tensor category equipped with a braiding; see [ibid., Definition 1.2.3]. A *symmetric tensor category* is a braided tensor category such that the square of the braiding is the identity.

Recall that F is a fixed field of characteristic zero. All categories and functors considered in this paper are going to be F-linear. So, an *F-linear tensor category* (or *tensor category over F*) is a tensor category which is F-linear (but not necessarily additive) and such that the tensor product functor is F-bilinear. A *Karoubian tensor category* over F is an F-linear tensor category which is Karoubian as an F-linear category (i.e., it is additive and every idempotent endomorphism is a projection to a direct summand). A *tensor ideal* $\mathcal{I}$ in a tensor category $\mathcal{T}$ consists of subspaces $\mathcal{I}(X,Y) \subset \mathrm{Hom}_{\mathcal{T}}(X,Y)$ for every $X, Y \in \mathcal{T}$ such that (i) $h \circ g \circ f \in \mathcal{I}(X,W)$ whenever $f \in \mathrm{Hom}_{\mathcal{T}}(X,Y)$, $g \in \mathcal{I}(Y,Z)$, $h \in \mathrm{Hom}_{\mathcal{T}}(Z,W)$, and (ii) $f \otimes \mathrm{id}_Z \in \mathcal{I}(X \otimes Z, Y \otimes Z)$ whenever $f \in \mathcal{I}(X,Y)$. For example, if the category $\mathcal{T}$ has a well defined trace the collection of *negligible morphisms*[3] forms a tensor ideal; see [Freedman 2003, §A1.3].

Finally we say that an F-linear symmetric tensor category $\mathcal{T}$ is *pre-Tannakian* if the following conditions are satisfied:

(a) All Hom's are finite-dimensional vector spaces over F and $\mathrm{End}(\mathbf{1}) = F$.

(b) $\mathcal{T}$ is an abelian category and all objects have finite length.

(c) $\mathcal{T}$ is rigid.

[3] Recall that a morphism $f \in \mathrm{Hom}_{\mathcal{T}}(X,Y)$ is negligible if $\mathrm{Tr}(fg) = 0$ for any $g \in \mathrm{Hom}_{\mathcal{T}}(Y,X)$. We will call an object negligible if its identity morphism is negligible.

Remark 2.1. In the terminology of [Deligne 1990] a pre-Tannakian category is the same as a "catégorie tensorielle" (see [ibid., §2.1]) satisfying a finiteness assumption [ibid., 2.12.1]. This is precisely the class of tensor categories over F for which a *fundamental group* (see [ibid., §8]) is defined.

2B. *The category* $\underline{\mathrm{Rep}}(S_t)$. We recall here briefly the construction of the category $\underline{\mathrm{Rep}}(S_t)$ following [Comes and Ostrik 2011, §2]. We refer the reader to *loc. cit.* and [Deligne 2007, §8] for much more detailed exposition.

2B1. *The category* $\underline{\mathrm{Rep}}_0(S_t)$. Let A be a finite set. A *partition* π of A is a collection of nonempty subsets $\pi_i \subset A$ such that $A = \bigsqcup_i \pi_i$ (disjoint union); the subsets π_i are called *parts* of the partition π. We say that partition π is *finer* than partition μ of the same set if any part of π is a subset of some part of μ. For three finite sets A, B, C and the partitions π of $A \sqcup B$ and μ of $B \sqcup C$ we define the partition $\mu \star \pi$ of $A \sqcup B \sqcup C$ as the finest partition such that parts of π and μ are subsets of its parts. The partition $\mu \star \pi$ induces a partition $\mu \cdot \pi$ of $A \sqcup C$ such that parts of $\mu \cdot \pi$ are nonempty intersections of parts of $\mu \star \pi$ with $A \sqcup C \subset A \sqcup B \sqcup C$; we also define an integer $\ell(\mu, \pi)$ which is the number of parts of $\mu \star \pi$ contained in B.

Definition 2.2. Given $t \in F$, we define the F-linear symmetric tensor category $\underline{\mathrm{Rep}}_0(S_t)$ as follows:

- Objects: finite sets; object corresponding to a finite set A is denoted $[A]$.
- Morphisms: $\mathrm{Hom}([A], [B])$ is the F-linear span of partitions of $A \sqcup B$; composition of morphisms represented by partitions $\pi \in \mathrm{Hom}([A], [B])$ and $\mu \in \mathrm{Hom}([B], [C])$ is $t^{\ell(\mu,\pi)}\mu \cdot \pi \in \mathrm{Hom}([A], [C])$.
- Tensor product: disjoint union (see [Comes and Ostrik 2011, Definition 2.15]); unit object is $[\varnothing]$; tensor product of morphisms, associativity and commutativity constraints are the obvious ones (see [ibid., §2.2]).

The category $\underline{\mathrm{Rep}}_0(S_t)$ has a distinguished object [pt] where pt is a one-element set. The object [pt] has a natural structure of commutative associative algebra in $\underline{\mathrm{Rep}}_0(S_t)$ where the multiplication (resp. unit) map is given by the partition of $\mathrm{pt} \sqcup \mathrm{pt} \sqcup \mathrm{pt}$ (resp. pt) consisting of one part. It is immediate to check that the object [pt] satisfies conditions (a), (b), (c) from Section 1B. Moreover, we have the following universal property:

Proposition 2.3. *Let $\mathcal{T}$ be an F-linear symmetric tensor category. The functor from the category of braided tensor functors $\mathcal{F} : \underline{\mathrm{Rep}}_0(S_t) \to \mathcal{T}$ to the category of objects $T \in \mathcal{T}$ satisfying* (a), (b), (c) *from Section 1B and their isomorphisms, which sends $\mathcal{F} \mapsto \mathcal{F}([\mathrm{pt}])$ and sends natural transformations $(\eta : \mathcal{F} \to \mathcal{F}') \mapsto \eta_{[\mathrm{pt}]}$, is an equivalence of categories.*

Sketch of proof. We restrict ourselves to a description of the inverse functor on objects; for more details, see [Deligne 2007, §8]. So assume that $T \in \mathcal{T}$ satisfies (a), (b), (c) from Section 1B. We define $\mathcal{F}([A]) = T^{\otimes A}$ (here $T^{\otimes A}$ is a tensor product of copies of T labeled by elements of A; since the category $\mathcal{T}$ is symmetric this is well defined). The tensor structure on the functor $\mathcal{F}$ will be given by the obvious isomorphisms $T^{\otimes A \sqcup B} = T^{\otimes A} \otimes T^{\otimes B}$. It remains to define $\mathcal{F}$ on the morphisms. Observe that a morphism from $\mathrm{Hom}([A], [B])$ represented by a partition π of $A \sqcup B$ is a tensor product of morphisms corresponding to partitions with precisely one part $\pi = \bigotimes_i \pi_i$. Thus it is sufficient to define $\mathcal{F}(\pi)$ only for π consisting of one part $A \sqcup B$. In this case we set $\mathcal{F}(\pi) = T^{\otimes A} \to T \to T^{\otimes B}$ where the first map is the multiplication morphism $T^{\otimes A} \to T$ and the second one is the dual to the multiplication morphism $T^{\otimes B} \to T$, where T and T^* are identified via (b) from Section 1B. One verifies that the assumptions (a), (b), (c) from Section 1B ensure that the tensor functor $\mathcal{F}$ is well defined. □

2B2. *The categories* $\underline{\mathrm{Rep}}(S_t)$ *and* $\underline{\mathrm{Rep}}^{ab}(S_d)$.

Definition 2.4 (compare [Deligne 2007, Définition 2.17] or [Comes and Ostrik 2011, Definition 2.19]). The category $\underline{\mathrm{Rep}}(S_t)$ is the Karoubian (or pseudoabelian) envelope[4] of the category $\underline{\mathrm{Rep}}_0(S_t)$.

It follows immediately from Proposition 2.3 that the category $\underline{\mathrm{Rep}}(S_t)$ has universal property from Proposition 1.1. We now use this universal property to construct Deligne's category $\underline{\mathrm{Rep}}^{ab}(S_d)$ from the introduction.

It is known (see [Deligne 2007, Théorème 2.18] or [Comes and Ostrik 2011, Corollary 5.21]) that the category $\underline{\mathrm{Rep}}(S_t)$ is semisimple (and hence pre-Tannakian) for $t \notin \mathbb{Z}_{\geq 0}$. In particular, the category $\underline{\mathrm{Rep}}(S_{-1})$ is pre-Tannakian, so its fundamental group π is defined. For any $d \in \mathbb{Z}_{\geq 0}$ we can consider the commutative associative algebra with nondegenerate trace pairing $T_d \in \underline{\mathrm{Rep}}(S_{-1})$ which is a direct sum of [pt] and $d+1$ copies of the algebra $\mathbf{1} = [\varnothing]$. Clearly, $\dim(T_d) = d$, so we can use Proposition 1.1 to construct a symmetric tensor functor $\underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}(S_{-1})$. Using the general properties of the fundamental group we get a factorization of this functor as $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon) \to \underline{\mathrm{Rep}}(S_{-1})$ (here $\mathbf{I} = \mathrm{Spec}(T_d)$ and $\varepsilon : \pi \to S_{\mathbf{I}}$ is the canonical homomorphism). It is clear that the category $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ is pre-Tannakian; it is proved in [Deligne 2007, Proposition 8.19] that the functor $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ is fully faithful. We set $\underline{\mathrm{Rep}}^{ab}(S_d) := \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$; as explained above this is a pre-Tannakian category and we have a fully faithful braided tensor functor $\underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}^{ab}(S_d)$.

Remark 2.5. The existence of the embedding $\underline{\mathrm{Rep}}(S_t) \subset \underline{\mathrm{Rep}}^{ab}(S_t)$ implies that $Y_1 \otimes Y_2 \neq 0$ for nonzero objects $Y_1, Y_2 \in \underline{\mathrm{Rep}}(S_t)$ (this is true in any abelian rigid

[4]We refer the reader to [Deligne 2007, §1.7-1.8] for the discussion of this notion.

tensor category with simple unit object). The same result can be proved directly as follows. Given finite sets A and B, it follows from the definition of tensor products that the obvious map $\mathrm{End}([A]) \otimes \mathrm{End}([B]) \to \mathrm{End}([A] \otimes [B]) = \mathrm{End}([A \sqcup B])$ is injective. Since any indecomposable object of $\underline{\mathrm{Rep}}(S_t)$ is the image of a primitive idempotent $e \in \mathrm{End}([A])$ for some finite set A (see, e.g., [Comes and Ostrik 2011, Proposition 2.20]), it follows that the tensor product of two nonzero morphisms in $\underline{\mathrm{Rep}}(S_t)$ is nonzero. The statement for objects follows by considering their identity morphisms.

2B3. *Indecomposable objects of the category* $\underline{\mathrm{Rep}}(S_t)$. The indecomposable objects of the category $\underline{\mathrm{Rep}}(S_t)$ are classified up to isomorphism in [Comes and Ostrik 2011, Theorem 3.3]. The isomorphism classes are labeled by the Young diagrams of all sizes in the following way. Let λ be a Young diagram of size $n = |\lambda|$ and let y_λ be the corresponding primitive idempotent in FS_n, the group algebra of the symmetric group.[5] The symmetric braiding gives rise to an action of S_n on $[\mathrm{pt}]^{\otimes n}$; let $[\mathrm{pt}]^\lambda$ denote the image of $y_\lambda \in \mathrm{End}([\mathrm{pt}]^{\otimes n})$. For any Young diagram λ of size $|\lambda|$ there is a unique indecomposable object $L(\lambda) \in \underline{\mathrm{Rep}}(S_t)$ characterized by the following properties:

(a) $L(\lambda)$ is not a direct summand of $[\mathrm{pt}]^{\otimes k}$ for $k < |\lambda|$.

(b) $L(\lambda)$ is a direct summand (with multiplicity 1) of $[\mathrm{pt}]^\lambda$.

It is proved in [Comes and Ostrik 2011, Theorem 3.3] that the indecomposable objects $L(\lambda)$ are well defined up to isomorphism, and any indecomposable object of $\underline{\mathrm{Rep}}(S_t)$ is isomorphic to precisely one $L(\lambda)$.

2B4. *Blocks of the category* $\underline{\mathrm{Rep}}(S_t)$. Let $\mathscr{A}$ be a Karoubian category such that any object decomposes into a finite direct sum of indecomposable objects. The set of isomorphism classes of indecomposable objects of $\mathscr{A}$ splits into *blocks* which are equivalence classes of the weakest equivalence relation for which two indecomposable objects are equivalent whenever there exists a nonzero morphism between them. We will also use the term block to refer to a full subcategory of $\mathscr{A}$ generated by the indecomposable objects in a single block.

The main result of [Comes and Ostrik 2011] is the description of blocks of the category $\underline{\mathrm{Rep}}(S_t)$. We describe the results of *loc. cit.* here. We will represent a Young diagram λ as an infinite nonincreasing sequence $(\lambda_1, \lambda_2, \ldots)$ of nonnegative integers such that $\lambda_k = 0$ for some $k > 0$; see [Comes and Ostrik 2011, §1.1]. For a Young diagram λ and $t \in F$ we define a sequence $\mu_\lambda(t) = (t - |\lambda|, \lambda_1 - 1, \lambda_2 - 2, \ldots)$.

Theorem 2.6 [Comes and Ostrik 2011, Theorem 5.3]. *The objects $L(\lambda)$ and $L(\lambda')$ of $\underline{\mathrm{Rep}}(S_t)$ are in the same block if and only if $\mu_\lambda(t)$ is a permutation of $\mu_{\lambda'}(t)$.*

[5] Here y_λ is a scalar multiple of the so-called Young symmetrizer (see, for instance, [Fulton and Harris 1991]).

Let $\mathcal{B}$ be the set of blocks of the category $\underline{\mathrm{Rep}}(S_t)$; for any $\mathsf{b} \in \mathcal{B}$ let us denote by $\underline{\mathrm{Rep}}_{\mathsf{b}}(S_t)$ the corresponding subcategory of $\underline{\mathrm{Rep}}(S_t)$; we have a decomposition $\underline{\mathrm{Rep}}(S_t) = \bigoplus_{\mathsf{b}\in\mathcal{B}} \underline{\mathrm{Rep}}_{\mathsf{b}}(S_t)$.

Proposition 2.7. *Let* $\mathsf{b} \in \mathcal{B}$. *One of the following holds*:

(i) b *is semisimple* (*or trivial*): *the category* $\underline{\mathrm{Rep}}_{\mathsf{b}}(S_t)$ *is equivalent to the category* Vec_F *of finite-dimensional F-vector spaces as an additive category. We will denote by $L = L(\mathsf{b})$ the unique indecomposable object of this block. Then* $\dim(L) = 0$, *or*, *equivalently*, id_L *is negligible.*

(ii) b *is nonsemisimple* (*or infinite*): *in this case the additive category* $\underline{\mathrm{Rep}}_{\mathsf{b}}(S_t)$ *is described in [Comes and Ostrik 2011, §6]* (*in particular*, *it does not depend on a choice of nonsemisimple block* b). *There is a natural labeling of indecomposable objects of the category* $\underline{\mathrm{Rep}}_{\mathsf{b}}(S_t)$ *by nonnegative integers*; *we will denote these objects by* $L_0, L_1, \ldots.$ *Then* $\dim(L_i) = 0$ *for* $i > 0$ *and* $\dim(L_0) \neq 0$, *that is*, id_{L_i} *is negligible if and only if* $i > 0$.

Further, it is shown in [Comes and Ostrik 2011] that for any $t \in F$ there are infinitely many semisimple blocks and finitely many (precisely the number of Young diagrams of size t) nonsemisimple blocks. In particular, for $t \notin \mathbb{Z}_{\geq 0}$ all blocks are semisimple (hence the category $\underline{\mathrm{Rep}}(S_t)$ is semisimple).

2C. ***Temperley–Lieb category.*** The results on the category $\underline{\mathrm{Rep}}(S_t)$ in many respects are parallel to the results on the Temperley–Lieb category $TL(q)$. We recall the definition and some properties of this category here.

Definition 2.8 (see, for example, [Freedman 2003, §A1.2]). Let q be a nonzero element of an algebraic closure of F such that $q+q^{-1} \in F$. We define the F-linear tensor category $TL_0(q)$ as follows:

- Objects: finite subsets of $\mathbb{R}$ considered up to isotopy; we will denote the object corresponding to the set A by $\langle A\rangle$.
- Morphisms: $\mathrm{Hom}(\langle A\rangle, \langle B\rangle)$ is the F-linear span of one-dimensional submanifolds of $\mathbb{R} \times [0, 1]$ with boundary $A \sqcup B$ where $A \subset \mathbb{R} \times 0$ and $B \subset \mathbb{R} \times 1$ (such submanifolds are called embedded unoriented *bordisms* from A to B) modulo the relation $[\text{bordism} \sqcup \text{circle}] = (q+q^{-1})[\text{bordism}]$; composition is given by juxtaposition.
- Tensor product: disjoint union (write $\mathbb{R} = \mathbb{R}_{<0} \sqcup 0 \sqcup \mathbb{R}_{>0}$ and identify $\mathbb{R}_{<0}$ and $\mathbb{R}_{>0}$ with $\mathbb{R}$); the unit object is $\langle\varnothing\rangle$; tensor product of morphisms and associativity constraint are the obvious ones.

Next we define the category $TL(q)$ as the Karoubian envelope of the category $TL_0(q)$. The category $TL(q)$ has a universal property (see, e.g., [Ostrik 2008, Theorem 2.4]) but we don't need it here. The indecomposable objects of the category $TL(q)$ are labeled by nonnegative integers: for any $i \in \mathbb{Z}_{\geq 0}$ there is a unique indecomposable object V_i which is a direct summand (with multiplicity 1) of $\langle \mathrm{pt}\rangle^{\otimes i}$ but is not a direct summand of $\langle \mathrm{pt}\rangle^{\otimes k}$ whenever $k < i$.

The category $TL(q)$ is semisimple for generic values of q; more precisely the category $TL(q)$ is not semisimple precisely when there exists a positive integer l such that $1 + q^2 + \cdots + q^{2l} = 0$ (we will denote the smallest such integer by l_q). Assume that the category $TL(q)$ is not semisimple. Then we have a full tensor functor $TL(q) \to \bar{\mathscr{C}}_q$ and a fully faithful tensor functor $TL(q) \to \mathscr{C}_q$, where $\bar{\mathscr{C}}_q$ is a semisimple tensor category (sometimes called the "Verlinde category") and $\mathscr{C}_q$ is the abelian tensor category of finite-dimensional representations of quantum SL(2); see, e.g., [Ostrik 2008, Theorem 2.4].

The blocks of the category $TL(q)$ are well known. Similarly to the case of the category $\underline{\mathrm{Rep}}(S_d)$ there are infinitely many semisimple blocks (which are equivalent to the category Vec_F as an additive category) and finitely many (precisely l_q) nonsemisimple blocks. The following observation is very important for this paper:

Proposition 2.9 [Comes and Ostrik 2011, Remark 6.5]. *All nonsemisimple blocks of the category $TL(q)$ are equivalent as additive categories. Moreover, they are equivalent to the category $\underline{\mathrm{Rep}}_{\mathsf{b}}(S_d)$, where b is any nonsemisimple block of the category $\underline{\mathrm{Rep}}(S_d)$.*

Remark 2.10. We can transport a labeling of indecomposable objects of $\underline{\mathrm{Rep}}_{\mathsf{b}}(S_d)$ (see Proposition 2.7(ii)) to a nonsemisimple block of the category $TL(q)$ via the equivalence of Proposition 2.9 (it is easy to see that the resulting labeling does not depend on a choice of the equivalence).

Recall that the category $TL(q)$ has a natural spherical structure and so the dimensions $\dim_{TL(q)}(Y)$ of objects $Y \in TL(q)$ are defined; see, e.g., [Freedman 2003, §A1.3]. The following result is well known; see, e.g., [Andersen 1992, (1.6) and Proposition 3.5]:

Lemma 2.11. *Let L be a unique indecomposable object in a semisimple block of $TL(q)$. Then $\dim_{TL(q)}(L) = 0$. For a nonsemisimple block we have $\dim_{TL(q)}(L_i) = 0$ for $i > 0$ and $\dim_{TL(q)}(L_0) \neq 0$, where L_i are indecomposable objects in this block labeled as in Remark 2.10.* □

3. Tensor ideals and the object $\Delta \in \underline{\mathrm{Rep}}(S_d)$

In this section we define objects $\Delta_n \in \underline{\mathrm{Rep}}(S_t)$ for $n \in \mathbb{Z}_{\geq 0}$ and $t \in F$. We then give Δ_n the structure of a commutative associative algebra in $\underline{\mathrm{Rep}}(S_t)$ and study many

Δ_n-modules. Finally, using our results on the objects Δ_n, we classify tensor ideals in $\underline{\mathrm{Rep}}(S_d)$ when d is a nonnegative integer. Before defining the objects Δ_n we prove the following easy observation which will be used throughout this section.

Proposition 3.1. *Suppose $A_0, \dots, A_n$ and $B_0, \dots, B_m$ are finite sets with $A_0 = B_0$ and $A_n = B_m$. Suppose further that f_i (resp. g_i) is an F-linear combination of partitions of $A_{i-1} \sqcup A_i$ (resp. $B_{i-1} \sqcup B_i$) whose coefficients do not depend on t for all $1 \le i \le n$ (resp. $1 \le i \le m$). If $f_n \cdots f_1 = g_m \cdots g_1$ in $\underline{\mathrm{Rep}}_0(S_t)$ for infinitely many values of $t \in F$, then $f_n \cdots f_1 = g_m \cdots g_1$ in $\underline{\mathrm{Rep}}_0(S_t)$ for all $t \in F$.*

Proof. For each $t \in F$ and partition π of $A_0 \sqcup A_n = B_0 \sqcup B_m$, let $a_\pi(t) \in F$ (resp. $b_\pi(t) \in F$) be such that $f_n \cdots f_1 = \sum_\pi a_\pi(t)\pi$ (resp. $g_m \cdots g_1 = \sum_\pi b_\pi(t)\pi$) in $\underline{\mathrm{Rep}}_0(S_t)$ where the sum is taken over all partitions π of $A_0 \sqcup A_n = B_0 \sqcup B_m$. Then $f_n \cdots f_1 = g_m \cdots g_1$ in $\underline{\mathrm{Rep}}_0(S_t)$ if and only if $a_\pi(t) = b_\pi(t)$ for all π. By the definition of composition in $\underline{\mathrm{Rep}}_0(S_t)$, both $a_\pi(t)$ and $b_\pi(t)$ are polynomials in t for each π. The result follows since a polynomial in t is determined by finitely many values of t. □

3A. *The objects $\Delta_n \in \underline{\mathrm{Rep}}(S_t)$.* Suppose n is a nonnegative integer and let $A_n = \{i \mid 1 \le i \le n\}$. Consider the endomorphism $x_n = x_{\mathrm{id}_n} : [A_n] \to [A_n]$ in $\underline{\mathrm{Rep}}_0(S_t)$ (see [Comes and Ostrik 2011, Equation (2.1)]).

Proposition 3.2. *x_n is an idempotent which is equal to its dual for all $n \ge 0$.*

Proof. The fact that $x_n^* = x_n$ follows from the definition of x_n. By Proposition 3.1, it suffices to show x_n is an idempotent in $\underline{\mathrm{Rep}}_0(S_t)$ for infinitely many values of t. It follows from [Comes and Ostrik 2011, Theorem 2.6 and Equation (2.2)] that x_n is an idempotent in $\underline{\mathrm{Rep}}_0(S_t)$ whenever t is an integer greater than $2n$. □

Since $\underline{\mathrm{Rep}}(S_t)$ is a Karoubian category (i.e., $\underline{\mathrm{Rep}}(S_t)$ contains images of idempotents) the following definition is valid.

Definition 3.3. Let $\Delta_n \in \underline{\mathrm{Rep}}(S_t)$ denote the image of the idempotent x_n.[6]

Note that the commutative associative algebra structure on [pt] extends in an obvious way to a commutative associative algebra structure on $[A_n] \cong [\mathrm{pt}]^{\otimes n}$. Let $\mu_n : [A_n] \otimes [A_n] \to [A_n]$ be the multiplication map and $1_n : \mathbf{1} \to [A_n]$ the unit map.

Proposition 3.4. *The multiplication map $x_n \mu_n (x_n \otimes x_n) : \Delta_n \otimes \Delta_n \to \Delta_n$ gives Δ_n the structure of a commutative associative algebra in $\underline{\mathrm{Rep}}(S_t)$ with unit given by $x_n 1_n : \mathbf{1} \to \Delta_n$.*

[6]In the notation of [Comes and Ostrik 2011], $\Delta_n = ([n], x_n)$.

Proof. We are required to show the following equalities hold in $\underline{\mathrm{Rep}}_0(S_t)$:

$$x_n\mu_n(x_n\mu_n(x_n\otimes x_n)\otimes x_n)=x_n\mu_n(x_n\otimes(x_n\mu_n(x_n\otimes x_n))), \tag{3-1}$$

$$x_n\mu_n(x_n 1_n\otimes x_n)=x_n=x_n\mu_n(x_n\otimes x_n 1_n), \tag{3-2}$$

$$x_n\mu_n(x_n\otimes x_n)\beta_{n,n}(x_n\otimes x_n)=x_n\mu_n(x_n\otimes x_n), \tag{3-3}$$

where $\beta_{n,n}: A_n\otimes A_n\to A_n\otimes A_n$ is the braiding morphism (see for example [Comes and Ostrik 2011, §2.2]). By Proposition 3.1, it suffices to show (3-1), (3-2) and (3-3) hold for infinitely many values of t.[7] Using Theorem 2.6 and Equation (2.2) of the same reference it is easy to show (3-1), (3-2) and (3-3) hold whenever t is a sufficiently large integer. □

By Proposition 3.4 we can consider the category Δ_n-mod of all left Δ_n-modules.

3B. ***Some Δ_n-modules.*** Suppose j is a nonnegative integer with $1\le j\le n$. Given a finite set X, let $\Theta^j_X:\mathrm{Hom}_{\underline{\mathrm{Rep}}(S_t)}(A_n,X)\to\mathrm{Hom}_{\underline{\mathrm{Rep}}(S_t)}(A_{n+1},X)$ and $\Theta^X_j:\mathrm{Hom}_{\underline{\mathrm{Rep}}(S_t)}(X,A_n)\to\mathrm{Hom}_{\underline{\mathrm{Rep}}(S_t)}(X,A_{n+1})$ be the F-linear maps defined on partitions as follows: if π is a partition of $X\sqcup A_n$, then $\Theta^j_X(\pi)=\Theta^X_j(\pi)$ is the unique partition of $X\sqcup A_{n+1}$ which restricts to π and has j and $n+1$ in the same part. Now let $\Theta_j:\mathrm{End}_{\underline{\mathrm{Rep}}(S_t)}(A_n)\to\mathrm{End}_{\underline{\mathrm{Rep}}(S_t)}(A_{n+1})$ be the F-linear map $\Theta_j=\Theta^j_{A_n}\circ\Theta^{A_n}_j$. It is easy to check that Θ_j is an injective (nonunital) F-algebra homomorphism for each $1\le j\le n$. In particular, by Proposition 3.2, $x_{n,j}:=\Theta_j(x_n)$ is an idempotent for each j.

Definition 3.5. Let $\Delta_n(j)\in\underline{\mathrm{Rep}}(S_t)$ denote the image of $x_{n,j}$.

Next we give $\Delta_n(j)$ the structure of a Δ_n-module. Let

$$\alpha=x_{n,j}\Theta^{A_n}_j(x_n)x_n:\Delta_n\to\Delta_n(j),$$
$$\beta=x_{n,j}\Theta_j(\mu_n)(x_{n,j}\otimes x_{n,j}):\Delta_n(j)\otimes\Delta_n(j)\to\Delta_n(j).$$

Finally, let $\phi=\beta(\alpha\otimes x_{n,j}):\Delta_n\otimes\Delta_n(j)\to\Delta_n(j)$.

Proposition 3.6. (1) *The map ϕ gives $\Delta_n(j)$ the structure of a Δ_n-module.*

(2) *The map $x_{n,j}\Theta^{A_n}_j(\mathrm{id}_{A_n})x_n:\Delta_n\to\Delta_n(j)$ is an isomorphism of Δ_n-modules with inverse $x_n\Theta^j_{A_n}(\mathrm{id}_{A_n})x_{n,j}$.*

Proof. For part (1) we are required to show the following equation holds in $\underline{\mathrm{Rep}}_0(S_t)$:

$$\begin{aligned}&x_{n,j}\Theta_j(\mu_n)\big((x_{n,j}\Theta^{A_n}_j(x_n)x_n\mu_n(x_n\otimes x_n))\otimes x_{n,j}\big)\\&\quad=x_{n,j}\Theta_j(\mu_n)\big(x_{n,j}\Theta^{A_n}_j(x_n)x_n\otimes(x_{n,j}\Theta_j(\mu_n)((x_{n,j}\Theta^{A_n}_j(x_n)x_n)\otimes x_{n,j}))\big).\end{aligned}\tag{3-4}$$

[7] In fact, (3-1), (3-2) and (3-3) do not depend on t, so we only need to verify they hold for some t.

For part (2) we are required to show the following equations hold in $\underline{\mathrm{Rep}}_0(S_t)$:

$$x_{n,j}\Theta_j^{A_n}(\mathrm{id}_{A_n})x_n\Theta_{A_n}^j(\mathrm{id}_{A_n})x_{n,j} = x_{n,j}, \tag{3-5}$$

$$x_n\Theta_{A_n}^j(\mathrm{id}_{A_n})x_{n,j}\Theta_j^{A_n}(\mathrm{id}_{A_n})x_n = x_n. \tag{3-6}$$

Now use Proposition 3.1 and [Comes and Ostrik 2011, Theorem 2.6 and Equation (2.2)]. □

Next, we give the object Δ_{n+1} the structure of a Δ_n-module. To do so, set $\psi = x_{n+1}(\mu_n \otimes \mathrm{id}_{[\mathrm{pt}]})(x_n \otimes x_{n+1}) : \Delta_n \otimes \Delta_{n+1} \to \Delta_{n+1}$.

Proposition 3.7. *The map ψ gives Δ_{n+1} the structure of a Δ_n-module.*

Proof. We are required to show the following equation holds in $\underline{\mathrm{Rep}}_0(S_t)$:

$$\begin{aligned} &x_{n+1}(\mu_n \otimes \mathrm{id}_{[\mathrm{pt}]})\big((x_n\mu_n(x_n \otimes x_n)) \otimes x_{n+1}\big) \\ &\qquad = x_{n+1}(\mu_n \otimes \mathrm{id}_{[\mathrm{pt}]})\big(x_n \otimes (x_{n+1}(\mu_n \otimes \mathrm{id}_{[\mathrm{pt}]})(x_n \otimes x_{n+1}))\big). \end{aligned} \tag{3-7}$$

Now use Proposition 3.1 and [Comes and Ostrik 2011, Theorem 2.6 and Equation (2.6)]. □

The following lemma will be important for us later.

Lemma 3.8. $\Delta_n \otimes [\mathrm{pt}] \cong \Delta_{n+1} \oplus \Delta_n(1) \oplus \cdots \oplus \Delta_n(n)$ *in the category Δ_n-mod.*

Proof. First, using Proposition 3.1 and [Comes and Ostrik 2011, Theorem 2.6 and Equation (2.6)] it is easy to show that the following identities hold in $\underline{\mathrm{Rep}}_0(S_t)$:

$$x_n \otimes \mathrm{id}_{[\mathrm{pt}]} = x_{n+1} + \sum_{1\le j\le n} x_{n,j}, \tag{3-8}$$

$$x_{n,j}x_{n+1} = 0 = x_{n+1}x_{n,j} \qquad (1 \le j \le n), \tag{3-9}$$

$$x_{n,j}x_{n,k} = \delta_{j,k}x_{n,j} \qquad (1 \le j, k \le n). \tag{3-10}$$

Next, define $\Psi : \Delta_n \otimes [\mathrm{pt}] \to \Delta_{n+1} \oplus \Delta_n(1) \oplus \cdots \oplus \Delta_n(n)$ by

$$\Psi = \begin{bmatrix} x_{n+1}(x_n \otimes \mathrm{id}_{[\mathrm{pt}]}) \\ x_{n,1}(x_n \otimes \mathrm{id}_{[\mathrm{pt}]}) \\ \vdots \\ x_{n,n}(x_n \otimes \mathrm{id}_{[\mathrm{pt}]}) \end{bmatrix}.$$

Using (3-8) it is easy to check that Ψ is an isomorphism in $\underline{\mathrm{Rep}}(S_t)$ with inverse

$$\Psi^{-1} = \big[(x_n \otimes \mathrm{id}_{[\mathrm{pt}]})x_{n+1} \;\; (x_n \otimes \mathrm{id}_{[\mathrm{pt}]})x_{n,1} \;\; \cdots \;\; (x_n \otimes \mathrm{id}_{[\mathrm{pt}]})x_{n,n}\big].$$

It remains to show that Ψ and Ψ^{-1} are morphisms in the category Δ_n-mod. Showing that Ψ is a morphism in Δ_n-mod amounts to showing the following equations hold in $\underline{\mathrm{Rep}}_0(S_t)$:

$$\begin{aligned}
&x_{n+1}(\mu_n \otimes \mathrm{id}_{[\mathrm{pt}]})(x_n \otimes (x_{n+1}(x_n \otimes \mathrm{id}_{[\mathrm{pt}]})))\\
&\qquad\qquad\qquad = x_{n+1}((x_n\mu_n(x_n \otimes x_n)) \otimes \mathrm{id}_{[\mathrm{pt}]}),\\
&x_{n,j}\Theta_j(\mu_n)((x_{n,j}\Theta_j^{A_n}(x_n)x_n) \otimes (x_{n,j}(x_n \otimes \mathrm{id}_{[\mathrm{pt}]})))\\
&\qquad\qquad\qquad = x_{n,j}((x_n\mu_n(x_n \otimes x_n)) \otimes \mathrm{id}_{[\mathrm{pt}]}) \qquad (1 \le j \le n).
\end{aligned} \tag{3-11}$$

To show the equations in (3-11) hold, use Proposition 3.1 and [Comes and Ostrik 2011, Theorem 2.6 and Equation (2.6)]. The proof for Ψ^{-1} is similar. □

3C. *The category* $\underline{\mathbf{Rep}}^{\Delta_n}(S_t)$. Let Δ_n-mod$_0$ denote the full subcategory of Δ_n-mod such that a Δ_n-module M is in Δ_n-mod$_0$ if and only if $M \cong \Delta_n \otimes Y$ in Δ_n-mod for some $Y \in \underline{\mathrm{Rep}}(S_t)$. Let $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ denote the Karoubian envelope of Δ_n-mod$_0$. The advantage of working in $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ rather than in the category Δ_n-mod is that we can give $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ the structure of a tensor category with relative ease. Indeed, given $M, M' \in \Delta_n$-mod$_0$ we know $M \cong \Delta_n \otimes Y$ and $M' \cong \Delta_n \otimes Y'$ as Δ_n-modules for some $Y, Y' \in \underline{\mathrm{Rep}}(S_t)$. Set $M \otimes_{\Delta_n} M' := \Delta_n \otimes Y \otimes Y'$. Given $N, N' \in \Delta_n$-mod$_0$ with $N \cong \Delta_n \otimes Z$ and $N' \cong \Delta_n \otimes Z'$ and morphisms $f \in \mathrm{Hom}_{\Delta_n\text{-mod}_0}(M, N)$ and $g \in \mathrm{Hom}_{\Delta_n\text{-mod}_0}(M', N')$, write

$$\tilde{f} : \Delta_n \otimes Y \xrightarrow{\cong} M \xrightarrow{f} N \xrightarrow{\cong} \Delta_n \otimes Z,$$
$$\tilde{g} : \Delta_n \otimes Y' \xrightarrow{\cong} M' \xrightarrow{g} N' \xrightarrow{\cong} \Delta_n \otimes Z'.$$

Define $f \otimes_{\Delta_n} g : M \otimes_{\Delta_n} M' \to N \otimes_{\Delta_n} N'$ to be the composition

$$\begin{aligned}
M \otimes_{\Delta_n} M' = \Delta_n \otimes Y \otimes Y' &\xrightarrow{\tilde{f}\otimes \mathrm{id}_{Y'}} \Delta_n \otimes Z \otimes Y' \xrightarrow{\sim} \Delta_n \otimes Y' \otimes Z\\
&\xrightarrow{\tilde{g}\otimes \mathrm{id}_{Z}} \Delta_n \otimes Z' \otimes Z \xrightarrow{\sim} \Delta_n \otimes Z \otimes Z' = N \otimes_{\Delta_n} N'.
\end{aligned}$$

It is easy to check that $\otimes_{\Delta_n} : \Delta_n\text{-mod}_0 \times \Delta_n\text{-mod}_0 \to \Delta_n\text{-mod}_0$ is a bifunctor which (with the obvious choice of constraints) makes Δ_n-mod$_0$ into a rigid symmetric tensor category. The tensor structure on Δ_n-mod$_0$ extends in an obvious way to make $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ a rigid symmetric tensor category too.

Notice that Δ_{n+1} is an object in $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$. Indeed, by Lemma 3.8, the Δ_n-module Δ_{n+1} is the image of an idempotent of the form $\Delta_n \otimes [\mathrm{pt}] \to \Delta_n \otimes [\mathrm{pt}]$. This idempotent is an element of $\mathrm{End}_{\Delta_n\text{-mod}_0}(\Delta_n \otimes [\mathrm{pt}])$; hence its image is an object in the Karoubian category $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$. The next two propositions concern the structure of $\Delta_{n+1} \in \underline{\mathrm{Rep}}^{\Delta_n}(S_t)$. We start by computing its dimension:

Proposition 3.9. $\qquad \dim_{\underline{\mathrm{Rep}}^{\Delta_n}(S_t)}(\Delta_{n+1}) = t - n.$

Proof. First, by Lemma 3.8 and Proposition 3.6(2),

$$\dim_{\underline{\mathrm{Rep}}^{\Delta_n}(S_t)}(\Delta_{n+1}) = \dim_{\underline{\mathrm{Rep}}^{\Delta_n}(S_t)}(\Delta_n \otimes [\mathrm{pt}]) - n \dim_{\underline{\mathrm{Rep}}^{\Delta_n}(S_t)}(\Delta_n).$$

Now, consider the tensor functor $\Delta_n \otimes - : \underline{\mathrm{Rep}}(S_t) \to \underline{\mathrm{Rep}}^{\Delta_n}(S_t)$. Since tensor functors preserve dimension, $\dim_{\underline{\mathrm{Rep}}^{\Delta_n}(S_t)}(\Delta_n) = \dim_{\underline{\mathrm{Rep}}(S_t)}([\varnothing]) = 1$ and $\dim_{\underline{\mathrm{Rep}}^{\Delta_n}(S_t)}(\Delta_n \otimes [\mathrm{pt}]) = \dim_{\underline{\mathrm{Rep}}(S_t)}([\mathrm{pt}]) = t$. $\square$

Our next aim is to show $\Delta_{n+1} \in \underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ satisfies (a) and (b) from Section 1B. To do so, let inc : $\Delta_{n+1} \to \Delta_n \otimes [\mathrm{pt}]$ and proj : $\Delta_n \otimes [\mathrm{pt}] \to \Delta_{n+1}$ denote the morphisms in $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ determined by Lemma 3.8. Moreover, let

$$m : (\Delta_n \otimes [\mathrm{pt}]) \otimes_{\Delta_n} (\Delta_n \otimes [\mathrm{pt}]) \to \Delta_n \otimes [\mathrm{pt}]$$

denote the morphism $\Delta_n \otimes [\mathrm{pt}] \otimes [\mathrm{pt}] \xrightarrow{\mathrm{id}_{\Delta_n} \otimes \mu_1} \Delta_n \otimes [\mathrm{pt}]$. Now consider the morphisms

$$\Delta_{n+1} \otimes_{\Delta_n} \Delta_{n+1} \xrightarrow{\mathrm{inc} \otimes_{\Delta_n} \mathrm{inc}} (\Delta_n \otimes [\mathrm{pt}]) \otimes_{\Delta_n} (\Delta_n \otimes [\mathrm{pt}]) \xrightarrow{m} \Delta_n \otimes [\mathrm{pt}] \xrightarrow{\mathrm{proj}} \Delta_{n+1} \tag{3-12}$$

and

$$\Delta_n \xrightarrow{\mathrm{id}_{\Delta_n} \otimes 1_1} \Delta_n \otimes [\mathrm{pt}] \xrightarrow{\mathrm{proj}} \Delta_{n+1}. \tag{3-13}$$

Proposition 3.10. *With the multiplication and unit maps given by* (3-12) *and* (3-13) *respectively*, $\Delta_{n+1} \in \underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ *satisfies* (a) *and* (b) *from Section 1B.*

Proof. Write $\mu_{\Delta_{n+1}}$ and $1_{\Delta_{n+1}}$ for the morphisms given by (3-12) and (3-13) respectively. First, it is easy to see that m (resp. $\mathrm{id}_{\Delta_n} \otimes 1_1$) is a morphism of Δ_n-modules. Hence, $\mu_{\Delta_{n+1}}$ (resp. $1_{\Delta_{n+1}}$) is a morphism of Δ_n-modules too. Now, to show Δ_{n+1} satisfies (a) from Section 1B we must show the following equations hold in $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$:

$$\begin{aligned} \mu_{\Delta_{n+1}}(\mu_{\Delta_{n+1}} \otimes_{\Delta_n} \mathrm{id}_{\Delta_{n+1}}) &= \mu_{\Delta_{n+1}}(\mathrm{id}_{\Delta_{n+1}} \otimes_{\Delta_n} \mu_{\Delta_{n+1}}), \\ \mu_{\Delta_{n+1}}(1_{\Delta_{n+1}} \otimes_{\Delta_n} \mathrm{id}_{\Delta_{n+1}}) &= \mathrm{id}_{\Delta_{n+1}} = \mu_{\Delta_{n+1}}(\mathrm{id}_{\Delta_{n+1}} \otimes_{\Delta_n} 1_{\Delta_{n+1}}), \\ \mu_{\Delta_{n+1}} \beta_{\Delta_{n+1}, \Delta_{n+1}} &= \mu_{\Delta_{n+1}}, \end{aligned} \tag{3-14}$$

where $\beta_{\Delta_{n+1},\Delta_{n+1}} : \Delta_{n+1} \otimes_{\Delta_n} \Delta_{n+1} \to \Delta_{n+1} \otimes_{\Delta_n} \Delta_{n+1}$ denotes the braiding morphism. To do so, first notice that by (3-8) the morphisms proj, inc, and $\mathrm{id}_{\Delta_{n+1}}$ are all given by x_{n+1}. Let τ (resp. ν) denote the identity morphism on $\Delta_{n+1} \otimes_{\Delta_n} \Delta_{n+1}$ (resp. $\Delta_{n+1} \otimes_{\Delta_n} \Delta_{n+1} \otimes_{\Delta_n} \Delta_{n+1}$). Then, by the definition of $\otimes_{\Delta_n}$, we have the following realizations of τ and ν as morphisms in $\underline{\mathrm{Rep}}_0(S_t)$:

$$\begin{aligned} \tau &= (x_n \otimes \beta_{1,1})(x_{n+1} \otimes \mathrm{id}_{[\mathrm{pt}]})(x_n \otimes \beta_{1,1})(x_{n+1} \otimes \mathrm{id}_{[\mathrm{pt}]}), \\ \nu &= (x_n \otimes \beta_{1,2})(x_{n+1} \otimes \mathrm{id}_{[\mathrm{pt}]\otimes[\mathrm{pt}]})(x_n \otimes \beta_{2,1})(\tau \otimes \mathrm{id}_{[\mathrm{pt}]}), \end{aligned} \tag{3-15}$$

where $\beta_{n,m} : A_n \otimes A_m \to A_m \otimes A_n$ denotes the braiding morphism in $\underline{\mathrm{Rep}}_0(S_t)$ for each $n, m \geq 0$. Moreover,

$$\begin{aligned} 1_{\Delta_{n+1}} = x_{n+1}(x_n \otimes 1_1), \quad \mu_{\Delta_{n+1}} = x_{n+1}(x_n \otimes \mu_1)\tau, \\ \beta_{\Delta_{n+1},\Delta_{n+1}} = \tau(x_n \otimes \beta_{1,1})\tau. \end{aligned} \tag{3-16}$$

Thus, showing the equations in (3-14) hold in $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ amounts to showing the following equations hold in $\underline{\mathrm{Rep}}_0(S_t)$:

$$
\begin{aligned}
&x_{n+1}(x_n\otimes\mu_1)\tau(x_n\otimes\beta_{1,1})(x_{n+1}\otimes\mathrm{id}_{[\mathrm{pt}]})(x_n\otimes\beta_{1,1})\big((x_{n+1}(x_n\otimes\mu_1)\tau)\otimes\mathrm{id}_{[\mathrm{pt}]}\big)\nu\\
&=x_{n+1}(x_n\otimes\mu_1)\tau(x_n\otimes\beta_{1,1})\big((x_{n+1}(x_n\otimes\mu_1)\tau)\otimes\mathrm{id}_{[\mathrm{pt}]}\big)(x_n\otimes\beta_{1,2})(x_{n+1}\otimes\mathrm{id}_{[\mathrm{pt}]\otimes[\mathrm{pt}]})\nu,\\
&x_{n+1}(x_n\otimes\mu_1)\tau(x_n\otimes\beta_{1,1})(x_{n+1}\otimes\mathrm{id}_{[\mathrm{pt}]})(x_n\otimes\beta_{1,1})(x_{n+1}(x_n\otimes 1_1)\otimes\mathrm{id}_{[\mathrm{pt}]})\\
&\qquad\qquad=x_{n+1}=x_{n+1}(x_n\otimes\mu_1)\tau(x_{n+1}\otimes\beta_{1,1})\big((x_{n+1}(x_n\otimes 1_1))\otimes\mathrm{id}_{[\mathrm{pt}]}\big)x_{n+1},\\
&\qquad\qquad\qquad x_{n+1}(x_n\otimes\mu_1)\tau(x_n\otimes\beta_{1,1})\tau=x_{n+1}(x_n\otimes\mu_1)\tau.
\end{aligned}
$$

All equations above are straightforward to check using Proposition 3.1 and [Comes and Ostrik 2011, Theorem 2.6 and Equation (2.6)]. Thus Δ_{n+1} satisfies part (a) from Section 1B.

To show Δ_{n+1} satisfies Section 1B(b), first notice that $\Delta_{n+1}\in\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$ is self-dual (because the morphism x_{n+1} is self-dual). Hence, we are required to show that the following morphism is invertible in $\underline{\mathrm{Rep}}^{\Delta_n}(S_t)$:

$$\big((\mathrm{Tr}\ \mu_{\Delta_{n+1}})\otimes_{\Delta_n}\mathrm{id}_{\Delta_{n+1}}\big)(\mathrm{id}_{\Delta_{n+1}}\otimes_{\Delta_n}\mathrm{coev}_{\Delta_{n+1}}):\Delta_{n+1}\to\Delta_{n+1},\tag{3-17}$$

where the morphism $\mathrm{Tr}:\Delta_{n+1}\to\Delta_n$ is defined in Section 1B(b). In fact, we claim the morphism in (3-17) is equal to the identity morphism $\mathrm{id}_{\Delta_{n+1}}$. To prove this claim, first notice that

$$\mathrm{Tr}=\mathrm{ev}_{\Delta_{n+1}}\beta_{\Delta_{n+1},\Delta_{n+1}}(\mu_{\Delta_{n+1}}\otimes_{\Delta_n}\mathrm{id}_{\Delta_{n+1}})(\mathrm{id}_{\Delta_{n+1}}\otimes_{\Delta_n}\mathrm{coev}_{\Delta_{n+1}}).\tag{3-18}$$

Also, $\mathrm{ev}_{\Delta_{n+1}}=x_n(x_n\otimes\mathrm{ev}_{[\mathrm{pt}]})\tau$ and $\mathrm{coev}_{\Delta_{n+1}}=\tau(x_n\otimes\mathrm{coev}_{[\mathrm{pt}]})x_n$. Hence, using (3-15), (3-16), and the definition of $\otimes_{\Delta_n}$, we can realize the morphism in (3-17) as a morphism in $\underline{\mathrm{Rep}}_0(S_t)$. Now use Proposition 3.1 and [Comes and Ostrik 2011, Theorem 2.6 and Equation (2.6)] to show that this morphism is equal to x_{n+1}. □

3D. ***Deligne's lemma.*** Fix an integer $d\geq 0$. Set $\Delta=\Delta_{d+1}\in\underline{\mathrm{Rep}}(S_d)$ and $\Delta^+=\Delta_{d+2}\in\underline{\mathrm{Rep}}^{\Delta}(S_d)$. By Proposition 3.9, $\dim_{\underline{\mathrm{Rep}}^{\Delta}(S_d)}(\Delta^+)=-1$. Hence, by Propositions 1.1 and 3.10, there exists a tensor functor $\mathcal{F}_\Delta:\underline{\mathrm{Rep}}(S_{-1})\to\underline{\mathrm{Rep}}^{\Delta}(S_d)$ with $\mathcal{F}_\Delta([\mathrm{pt}])=\Delta^+$. Let $\underline{\mathrm{Res}}^{S_d}_{S_{-1}}$ denote the tensor functor $\underline{\mathrm{Rep}}(S_d)\to\underline{\mathrm{Rep}}(S_{-1})$ described in Definition 2.4, i.e., the functor prescribed by Proposition 1.1 with $\underline{\mathrm{Res}}^{S_d}_{S_{-1}}([\mathrm{pt}])=[\mathrm{pt}]\oplus[\varnothing]^{\oplus d+1}$. Then we have the following:

Lemma 3.11. *The functor $\Delta\otimes-:\underline{\mathrm{Rep}}(S_d)\to\underline{\mathrm{Rep}}^{\Delta}(S_d)$ is isomorphic to the composition $\mathcal{F}_\Delta\circ\underline{\mathrm{Res}}^{S_d}_{S_{-1}}$.*

Proof. Both $\Delta\otimes-$ and $\mathcal{F}_\Delta\circ\underline{\mathrm{Res}}^{S_d}_{S_{-1}}$ are tensor functors which map $[\mathrm{pt}]\in\underline{\mathrm{Rep}}(S_d)$ to an object isomorphic to $\Delta^+\oplus\Delta^{\oplus d+1}\in\underline{\mathrm{Rep}}^{\Delta}(S_d)$ (see Propositions 3.6(2) and 3.8). Hence, by Proposition 1.1, they are isomorphic. □

The following corollary to Deligne's lemma will be used in the next section to classify tensor ideals in $\underline{\mathrm{Rep}}(S_d)$.

Corollary 3.12. *Every nonzero tensor ideal in $\underline{\mathrm{Rep}}(S_d)$ contains a nonzero identity morphism.*

Proof. Suppose $\mathcal{I}$ is a nonzero tensor ideal in $\underline{\mathrm{Rep}}(S_d)$. Since tensor ideals are closed under composition, it suffices to show that $\mathcal{I}$ contains a morphism which has a nonzero isomorphism as a direct summand. Let f be a nonzero morphism in $\mathcal{I}$. Then, by Remark 2.5, $\mathrm{id}_\Delta \otimes f$ is also a nonzero morphism in $\mathcal{I}$. By Lemma 3.11, we have $\mathrm{id}_\Delta \otimes f = \mathcal{F}_\Delta(f')$ for some nonzero morphism f' in $\underline{\mathrm{Rep}}(S_{-1})$. Since $\underline{\mathrm{Rep}}(S_{-1})$ is semisimple (see [Deligne 2007, Théorème 2.18] or [Comes and Ostrik 2011, Corollary 5.21]) it follows that f' (and therefore $\mathcal{F}_\Delta(f')$) is the direct sum of isomorphisms and zero morphisms. □

3E. ***Tensor ideals in*** $\underline{\mathbf{Rep}}(S_d)$**.** In this section we use results from [Comes and Ostrik 2011] along with Corollary 3.12 to classify tensor ideals in $\underline{\mathrm{Rep}}(S_d)$ for arbitrary $d \in \mathbb{Z}_{\geq 0}$.[8] We begin by introducing an equivalence class on Young diagrams:

Definition 3.13. Consider the weakest equivalence relation on the set of all Young diagrams such that λ and μ are equivalent whenever the indecomposable object $L(\lambda)$ is a direct summand of $L(\mu) \otimes [\mathrm{pt}]$ in $\underline{\mathrm{Rep}}(S_d)$. When λ and μ are in the same equivalence class we write $\lambda \overset{d}{\approx} \mu$.

The following proposition contains enough information on the equivalence relation $\overset{d}{\approx}$ for us to classify tensor ideals in $\underline{\mathrm{Rep}}(S_d)$.

Proposition 3.14. *Assume d is a nonnegative integer and λ, μ are Young diagrams.*

(1) *A nonzero morphism of the form $L(\lambda) \to L(\mu)$ is a negligible morphism in $\underline{\mathrm{Rep}}(S_d)$ if and only if $L(\lambda)$ or $L(\mu)$ is not the minimal indecomposable object in an infinite block of $\underline{\mathrm{Rep}}(S_d)$.*

(2) *$\lambda \overset{d}{\approx} \mu$ whenever $L(\lambda)$ and $L(\mu)$ are in trivial blocks of $\underline{\mathrm{Rep}}(S_d)$.*

(3) *$\lambda \overset{d}{\approx} \mu$ whenever $L(\lambda)$ is a nonminimal indecomposable object in an infinite block and $L(\mu)$ is in a trivial block of $\underline{\mathrm{Rep}}(S_d)$.*

(4) *$\lambda \overset{d}{\approx} \mu$ whenever neither $L(\lambda)$ nor $L(\mu)$ is a minimal indecomposable object in an infinite block of $\underline{\mathrm{Rep}}(S_d)$.*

(5) *Suppose $\lambda \overset{d}{\approx} \mu$ and $\mathcal{I}$ is a tensor ideal in $\underline{\mathrm{Rep}}(S_d)$ containing $\mathrm{id}_{L(\lambda)}$. Then $\mathrm{id}_{L(\mu)}$ is also in $\mathcal{I}$.*

[8]If $t \notin \mathbb{Z}_{\geq 0}$, then $\underline{\mathrm{Rep}}(S_t)$ is semisimple (see [Deligne 2007, Théorème 2.18] or [Comes and Ostrik 2011, Corollary 5.21]). Hence, there are no nonzero proper tensor ideals in $\underline{\mathrm{Rep}}(S_t)$ when $t \notin \mathbb{Z}_{\geq 0}$.

Proof. Part (1) follows from [Comes and Ostrik 2011, Proposition 3.25, Corollary 5.9, and Theorem 6.10]. Part (2) is easy to check using [Comes and Ostrik 2011, Propositions 3.12, 5.15 and Lemma 5.20(1)]. Part (4) follows from parts (2) and (3). Part (5) is easy to check. Hence, it suffices to prove part (3). To do so, let b denote the infinite block of $\underline{\mathrm{Rep}}(S_d)$ containing $L(\lambda)$. We will proceed by induction on b with respect to $\prec$ (see [Comes and Ostrik 2011, Definition 5.12]).

If b is the minimal with respect to $\prec$, then using [Comes and Ostrik 2011, Proposition 3.12 and Lemmas 5.18(1) and 5.20(1)] we can find a Young diagram ρ with $L(\rho)$ in a trivial block of $\underline{\mathrm{Rep}}(S_d)$ such that $\lambda \overset{d}{\approx} \rho$. By part (2), $\rho \overset{d}{\approx} \mu$ and we are done. Now suppose b is not minimal with respect to $\prec$. Then, using [Comes and Ostrik 2011, Proposition 3.12 and Lemmas 5.18(2) and 5.20(2)], we can find a Young diagram ρ' with $\lambda \overset{d}{\approx} \rho'$ such that $L(\rho')$ is in an infinite block b' of $\underline{\mathrm{Rep}}(S_d)$ with $\mathsf{b}' \precneqq \mathsf{b}$. By induction $\rho' \overset{d}{\approx} \mu$ and we are done. □

We are now ready to classify tensor ideals in $\underline{\mathrm{Rep}}(S_d)$.

Theorem 3.15. *If d is a nonnegative integer, then the only nonzero proper tensor ideal in $\underline{\mathrm{Rep}}(S_d)$ is the ideal of negligible morphisms.*

Proof. Assume $\mathcal{I}$ is a nonzero proper tensor ideal of $\underline{\mathrm{Rep}}(S_d)$. Then $\mathcal{I}$ is contained in the ideal of negligible morphisms (see [Freedman 2003, Proposition 3.1]), hence we must show that $\mathcal{I}$ contains all negligible morphisms. Suppose λ is a Young diagram such that $L(\lambda)$ is not the minimal indecomposable object in an infinite block of $\underline{\mathrm{Rep}}(S_d)$. By Proposition 3.14(1), it suffices to show $\mathrm{id}_{L(\lambda)}$ is contained in $\mathcal{I}$. By Corollary 3.12, there exists a nonzero identity morphism in $\mathcal{I}$. It follows that $\mathcal{I}$ contains $\mathrm{id}_{L(\mu)}$ for some Young diagram μ. In particular, $\mathrm{id}_{L(\mu)}$ is a negligible morphism. Hence, by Proposition 3.14(1), $L(\mu)$ is not the minimal indecomposable object in an infinite block of $\underline{\mathrm{Rep}}(S_d)$. Thus, by Proposition 3.14(4), $\lambda \overset{d}{\approx} \mu$. Finally, by Proposition 3.14(5), $\mathrm{id}_{L(\lambda)}$ is contained in $\mathcal{I}$. □

Corollary 3.16. *The tensor ideal in $\underline{\mathrm{Rep}}(S_d)$ generated by id_Δ is the ideal of all negligible morphisms.*

Proof. $\mathrm{id}_\Delta = x_{d+1}$ is a nonzero negligible morphism in $\underline{\mathrm{Rep}}(S_d)$ (see [Comes and Ostrik 2011, Remark 3.22]). Hence, the result follows from Theorem 3.15. □

4. The t-structure on $K^b(\underline{\mathrm{Rep}}(S_d))$

4A. ***Homotopy category.*** Let $\mathcal{A}$ be an additive category. Let $K^b(\mathcal{A})$ be the bounded homotopy category of $\mathcal{A}$; see, e.g., [Kashiwara and Schapira 2006, §11]. Thus the objects of $K^b(\mathcal{A})$ are finite complexes of objects in $\mathcal{A}$ and the morphisms are morphisms of complexes up to homotopy. The category $K^b(\mathcal{A})$ has a natural structure of a triangulated category; see *loc. cit.* In particular, for each integer n we have a translation functor $[n] : K^b(\mathcal{A}) \to K^b(\mathcal{A})$.

Any object $A \in \mathcal{A}$ can be considered as a complex $A[0]$ concentrated in degree 0 or, more generally, as a complex $A[n]$ concentrated in degree $-n$. Thus we have a fully faithful functor $\mathcal{A} \to K^b(\mathcal{A})$, $A \mapsto A[0]$. We will say that an object $K \in K^b(\mathcal{A})$ is *split* if it is isomorphic to an object of the form $\bigoplus_i A_i[n_i]$ with $A_i \in \mathcal{A}$, $n_i \in \mathbb{Z}$.

Now assume that $\mathcal{A}$ is an additive tensor category. The category $K^b(\mathcal{A})$ has a natural structure of an additive tensor category. If the category $\mathcal{A}$ is braided or symmetric then so is the category $K^b(\mathcal{A})$. The functor $\mathcal{A} \to K^b(\mathcal{A})$, $A \mapsto A[0]$ has an obvious structure of a (braided) tensor functor. If the category $\mathcal{A}$ is rigid so is the category $K^b(\mathcal{A})$.

4B. *Definition of t-structure.* We can apply the construction from Section 4A to the case $\mathcal{A} = \underline{\mathrm{Rep}}(S_d)$. We obtain a triangulated tensor category $\mathcal{K}_d := K^b(\underline{\mathrm{Rep}}(S_d))$.

Proposition 4.1. *For any $K \in \mathcal{K}_d$ the object $\Delta \otimes K$ is split.*

Proof. By Lemma 3.11, the functor $\Delta \otimes - : \underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}(S_d)$ is naturally isomorphic to a composition $\underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}(S_{-1}) \to \underline{\mathrm{Rep}}(S_d)$. The category $\underline{\mathrm{Rep}}(S_{-1})$ is semisimple ([Deligne 2007, Théorème 2.18] or [Comes and Ostrik 2011, Corollary 5.21]), so every object of $K^b(\underline{\mathrm{Rep}}(S_{-1}))$ is split. The result follows. □

We define $\mathcal{K}_d^{\le 0}$ as the full subcategory of $\mathcal{K}_d$ consisting of objects K such that $\Delta \otimes K$ is concentrated in nonpositive degrees (that is, isomorphic to $\bigoplus_i A_i[n_i]$ with $A_i \in \mathcal{A}$ and $n_i \in \mathbb{Z}_{\ge 0}$). Similarly, we define $\mathcal{K}_d^{\ge 0}$ as the full subcategory of $\mathcal{K}_d$ consisting of objects K such that $\Delta \otimes K$ is concentrated in nonnegative degrees. The following result will be proved in Section 4C.

Theorem 4.2. *The pair $(\mathcal{K}_d^{\le 0}, \mathcal{K}_d^{\ge 0})$ is a t-structure* (see [Beĭlinson et al. 1982, Définition 1.3.1]) *on the category $\mathcal{K}_d$.*

Recall that the core of this t-structure is the subcategory $\mathcal{K}_d^0 = \mathcal{K}_d^{\le 0} \cap \mathcal{K}_d^{\ge 0}$. By definition this means that $K \in \mathcal{K}_d^0$ if and only if $\Delta \otimes K$ is concentrated in degree zero. In particular, for any $A \in \underline{\mathrm{Rep}}(S_d)$ the object $A[0]$ is in $\mathcal{K}_d^0$.

Corollary 4.3. (a) *The category $\mathcal{K}_d^0$ is abelian.*

(b) *The category $\mathcal{K}_d^0$ is a tensor subcategory of $\mathcal{K}_d$.*

Proof. Part (a) follows from Theorem 4.2 and [Beĭlinson et al. 1982, Théorème 1.3.6]. For (b) we need to check that for $K, K' \in \mathcal{K}_d^0$ we have $K \otimes K' \in \mathcal{K}_d^0$. Assume this is not the case. This means that the split complex $\Delta \otimes K \otimes K'$ is not concentrated in degree zero. Since $\Delta \otimes X \neq 0$ for any $0 \neq X \in \underline{\mathrm{Rep}}(S_d)$ (see Remark 2.5) we get that $\Delta \otimes \Delta \otimes K \otimes K'$ is split and not concentrated in degree zero. But this is not the case since $\Delta \otimes \Delta \otimes K \otimes K' \simeq (\Delta \otimes K) \otimes (\Delta \otimes K')$ and both $\Delta \otimes K$ and $\Delta \otimes K'$ are split and concentrated in degree zero. □

We will show in Section 4C that the category $\mathcal{K}_d^0$ is actually pre-Tannakian. Thus we constructed a fully faithful tensor functor $\underline{\mathrm{Rep}}(S_d) \to \mathcal{K}_d^0$, where $\mathcal{K}_d^0$ is a pre-Tannakian category. Of course a priori this might be quite different from Deligne's functor $\underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}^{ab}(S_d)$.

4C. ***Verification of t-structure axioms.*** The main goal of this section is to prove Theorem 4.2.

4C1. We start by reformulating the definition of $\mathcal{K}_d^{\leq 0}$ and $\mathcal{K}_d^{\geq 0}$ in terms of negligible objects, i.e., objects whose identity morphisms are negligible.

Proposition 4.4. *Let $K \in \mathcal{K}_d$. Then $K \in \mathcal{K}_d^{\leq 0}$ if and only if* $\mathrm{Hom}(K, A[n]) = 0$ *for any negligible $A \in \underline{\mathrm{Rep}}(S_d)$ and $n \in \mathbb{Z}_{<0}$. Similarly, $K \in \mathcal{K}_d^{\geq 0}$ if and only if* $\mathrm{Hom}(K, A[n]) = 0$ *for any negligible $A \in \underline{\mathrm{Rep}}(S_d)$ and $n \in \mathbb{Z}_{>0}$.*

Proof. We prove only the characterization of $\mathcal{K}_d^{\leq 0}$ (the case of $\mathcal{K}_d^{\geq 0}$ is similar). Assume first that $\mathrm{Hom}(K, A[n]) = 0$ for any negligible A and $n \in \mathbb{Z}_{<0}$. By Proposition 3.2 $\Delta^* = \Delta$, thus by Corollary 3.16 $\Delta \otimes B = \Delta^* \otimes B$ is negligible for all $B \in \underline{\mathrm{Rep}}(S_d)$. Hence, $\mathrm{Hom}(\Delta \otimes K, B[n]) = \mathrm{Hom}(K, \Delta^* \otimes B[n]) = 0$ for any $B \in \underline{\mathrm{Rep}}(S_d)$ and $n \in \mathbb{Z}_{<0}$. Since by Proposition 4.1 the object $\Delta \otimes K \in \mathcal{K}_d$ is split we get immediately that $K \in \mathcal{K}_d^{\leq 0}$.

Conversely, assume that $K \in \mathcal{K}_d^{\leq 0}$. Then by definition $\mathrm{Hom}(\Delta \otimes K, B[n]) = 0$ for any $B \in \underline{\mathrm{Rep}}(S_d)$ and $n \in \mathbb{Z}_{<0}$. Hence $\mathrm{Hom}(K, \Delta^* \otimes B[n]) = 0$. Since, by Corollary 3.16, any negligible object is a direct summand of an object of the form $\Delta \otimes B = \Delta^* \otimes B$ we are done. ☐

4C2. *Blockwise description of* $(\mathcal{K}_d^{\leq 0}, \mathcal{K}_d^{\geq 0})$. Recall that the category $\underline{\mathrm{Rep}}(S_d)$ decomposes into a direct sum of blocks $\underline{\mathrm{Rep}}(S_d) = \bigoplus_{\mathsf{b}} \underline{\mathrm{Rep}}_{\mathsf{b}}(S_d)$; see Section 2B4. Similarly, we have a decomposition $\mathcal{K}_d = \bigoplus_{\mathsf{b}} (\mathcal{K}_d)_{\mathsf{b}}$ (in other words, for any object $K \in \mathcal{K}_d$ we have a canonical decomposition $K = \bigoplus_{\mathsf{b}} K_{\mathsf{b}}$ where all the terms of the complex $K_{\mathsf{b}} \in (\mathcal{K}_d)_{\mathsf{b}}$ are in the block $\underline{\mathrm{Rep}}_{\mathsf{b}}(S_d)$). Since $\Delta \otimes \big(\bigoplus_{\mathsf{b}} K_{\mathsf{b}}\big) = \bigoplus_{\mathsf{b}} \Delta \otimes K_{\mathsf{b}}$ we see that $K = \bigoplus_{\mathsf{b}} K_{\mathsf{b}} \in \mathcal{K}_d^{\leq 0}$ if and only if $K_{\mathsf{b}} \in \mathcal{K}_d^{\leq 0}$ for any b (and similarly for $\mathcal{K}_d^{\geq 0}$). In other words $\mathcal{K}_d^{\leq 0} = \bigoplus_{\mathsf{b}} (\mathcal{K}_d^{\leq 0})_{\mathsf{b}}$ where $(\mathcal{K}_d^{\leq 0})_{\mathsf{b}} = \mathcal{K}_d^{\leq 0} \cap (\mathcal{K}_d)_{\mathsf{b}}$, that is, the subcategory $\mathcal{K}_d^{\leq 0}$ is compatible with the block decomposition (and similarly for $\mathcal{K}_d^{\geq 0} = \bigoplus_{\mathsf{b}} (\mathcal{K}_d^{\geq 0})_{\mathsf{b}}$). Thus in order to verify that $(\mathcal{K}_d^{\leq 0}, \mathcal{K}_d^{\geq 0})$ is a t-structure on $\mathcal{K}_d$ it is sufficient to verify that $((\mathcal{K}_d^{\leq 0})_{\mathsf{b}}, (\mathcal{K}_d^{\geq 0})_{\mathsf{b}})$ is a t-structure on $(\mathcal{K}_d)_{\mathsf{b}}$ for every block b. Fortunately, Proposition 4.4 gives rise to an easy description of $(\mathcal{K}_d^{\leq 0})_{\mathsf{b}}$ and $(\mathcal{K}_d^{\geq 0})_{\mathsf{b}}$.

Proposition 4.5. *Let $K \in (\mathcal{K}_d)_{\mathsf{b}}$.*

(a) *Assume that* b *is a semisimple block and let L be a unique indecomposable object in* b. *Then $K \in (\mathcal{K}_d^{\leq 0})_{\mathsf{b}}$ (resp. $K \in (\mathcal{K}_d^{\geq 0})_{\mathsf{b}}$) if and only if $K \in (\mathcal{K}_d)_{\mathsf{b}}$ and* $\mathrm{Hom}(K, L[n]) = 0$ *for any $n \in \mathbb{Z}_{<0}$ (resp. for $n \in \mathbb{Z}_{>0}$).*

(b) *Assume that* b *is a nonsemisimple block with indecomposable objects* L_i *for* $i \in \mathbb{Z}_{\geq 0}$ *labeled as in Proposition 2.7(ii). Then* $K \in (\mathcal{K}_d^{\leq 0})_{\mathsf{b}}$ (*resp.* $K \in (\mathcal{K}_d^{\geq 0})_{\mathsf{b}}$) *if and only if* $K \in (\mathcal{K}_d)_{\mathsf{b}}$ *and* $\mathrm{Hom}(K, L_i[n]) = 0$ *for all* $i > 0$ *and any* $n \in \mathbb{Z}_{<0}$ (*resp. for* $n \in \mathbb{Z}_{>0}$).

Proof. Combine Proposition 4.4 and Proposition 2.7. □

4C3. *Analogy with Temperley–Lieb category.* The definition of the t-structure in Section 4B was motivated by the following analogy. Pick a nontrivial root of unity q such that $q + q^{-1} \in F$ and recall the Temperley–Lieb category $TL(q)$ from Section 2C. Consider the category $K^b(TL(q))$. It is well known (see, e.g., [Ostrik 2008, Proposition 2.7] that the embedding $TL(q) \subset \mathcal{C}_q$ induces an equivalence of triangulated categories $K^b(TL(q)) \simeq D^b(\mathcal{C}_q)$, where $D^b(\mathcal{C}_q)$ is the derived category of the abelian category $\mathcal{C}_q$. In particular the category $\mathcal{D}_q := K^b(TL(q))$ inherits a natural t-structure $(\mathcal{D}_q^{\leq 0}, \mathcal{D}_q^{\geq 0})$ from the category $D^b(\mathcal{C}_q)$; see, e.g., [Beĭlinson et al. 1982, Exemple 1.3.2(i)].[9] This t-structure can be characterized as follows.

Let $\mathrm{St} := V_{l-1} \in TL(q)$ be the so-called *Steinberg module*. It is known (see [Andersen et al. 1991, Theorem 9.8]) that St is a projective object of the category $\mathcal{C}_q$. Thus $\mathrm{St} \otimes Y$ is a projective object of $\mathcal{C}_q$ for any $Y \in \mathcal{C}_q$; see [Andersen et al. 1991, Lemma 9.10]. In particular, for any $K \in \mathcal{D}_q$ the object $\mathrm{St} \otimes K \in \mathcal{D}_q$ is isomorphic to its cohomology (as a finite complex consisting of projective modules and with projective cohomology). It is well known that each projective object of $\mathcal{C}_q$ is contained in $TL(q) \subset \mathcal{C}_q$; see [Andersen 1992, (5.7)]. Thus, in the language of Section 4A, for any $K \in K^b(TL(q))$ the complex $\mathrm{St} \otimes K$ is split (analogous to Proposition 4.1). It is clear that $K \in \mathcal{D}_q^{\leq 0}$ if and only if $\mathrm{St} \otimes K$ is concentrated in nonpositive degrees and similarly for $\mathcal{D}_q^{\geq 0}$. This is a counterpart of the definition of the t-structure $(\mathcal{K}_d^{\leq 0}, \mathcal{K}_d^{\geq 0})$.

Furthermore, it is known that each direct summand of $\mathrm{St} \otimes Y$ for $Y \in TL(q)$ is negligible (see [Andersen 1992, Proposition 3.5 and Lemma 3.6]) and that each negligible object of $TL(q)$ is a direct summand of $\mathrm{St} \otimes Y$ with $Y \in TL(q)$ (see [Andersen 1992, p. 158]). Thus we have the following counterpart of Proposition 4.4 (with a similar proof):

(a) *Let* $K \in \mathcal{D}_q$. *Then* $K \in \mathcal{D}_q^{\leq 0}$ (*resp.* $K \in \mathcal{D}_q^{\geq 0}$) *if and only if* $\mathrm{Hom}(K, A[n]) = 0$ *for any negligible* $A \in TL(q)$ *and* $n \in \mathbb{Z}_{<0}$ (*resp.* $n \in \mathbb{Z}_{>0}$).

Hence, following Section 4C2, we can give a blockwise description of the t-structure $(\mathcal{D}_q^{\leq 0}, \mathcal{D}_q^{\geq 0})$. For a block b let $(\mathcal{D}_q)_{\mathsf{b}}$ denote the full subcategory of $\mathcal{D}_q = K^b(TL(q))$ consisting of complexes with all terms from the block b. Using Lemma 2.11 we obtain the following counterpart of Proposition 4.5:

[9] Thus the category $\mathcal{D}_q^{\leq 0}$ consists of objects of $D^b(\mathcal{C}_q)$ with nontrivial cohomology only in nonpositive degrees and similarly for $\mathcal{D}_q^{\geq 0}$.

(b) *Let* b *be a nonsemisimple block of* $TL(q)$ *with indecomposable objects* L_i *for* $i \in \mathbb{Z}_{\geq 0}$ *labeled as in Remark 2.10. Let* $K \in (\mathcal{D}_q)_{\mathsf{b}}$. *Then* $K \in \mathcal{D}_q^{\leq 0}$ (*resp.* $K \in \mathcal{D}_q^{\geq 0}$) *if and only if* $\mathrm{Hom}(K, L_i[n]) = 0$ *for all* $i > 0$ *and any* $n \in \mathbb{Z}_{<0}$ (*resp. for* $n \in \mathbb{Z}_{>0}$).

From this description it is clear that the pair $(\mathcal{D}_q^{\leq 0} \cap (\mathcal{D}_q)_{\mathsf{b}}, \mathcal{D}_q^{\geq 0} \cap (\mathcal{D}_q)_{\mathsf{b}})$ of subcategories of $(\mathcal{D}_q)_{\mathsf{b}}$ corresponds to the pair $((\mathcal{K}_d^{\leq 0})_{\mathsf{b}'}, (\mathcal{K}_d^{\geq 0})_{\mathsf{b}'})$ under the equivalence $(\mathcal{D}_q)_{\mathsf{b}} \simeq (\mathcal{K}_d)_{\mathsf{b}'}$ induced by the equivalence of blocks from Proposition 2.9. Since $(\mathcal{D}_q^{\leq 0} \cap (\mathcal{D}_q)_{\mathsf{b}}, \mathcal{D}_q^{\geq 0} \cap (\mathcal{D}_q)_{\mathsf{b}})$ is a t-structure on the category $(\mathcal{D}_q)_{\mathsf{b}}$ we have the following:

Corollary 4.6. *Let* b *be a nonsemisimple block of the category* $TL(q)$ *and let* b' *be an equivalent block in the category* $\underline{\mathrm{Rep}}(S_d)$ *as in Proposition 2.9. Then* $((\mathcal{K}_d^{\leq 0})_{\mathsf{b}'}, (\mathcal{K}_d^{\geq 0})_{\mathsf{b}'})$ *is a* t*-structure on the category* $(\mathcal{K}_d)_{\mathsf{b}'}$. □

4C4. *Proof of Theorem 4.2.* It suffices to show $((\mathcal{K}_d^{\leq 0})_{\mathsf{b}}, (\mathcal{K}_d^{\geq 0})_{\mathsf{b}})$ is a t-structure on $(\mathcal{K}_d)_{\mathsf{b}}$ for every block b. If the block b is semisimple then the category $(\mathcal{K}_d)_{\mathsf{b}}$ can be identified with $K^b(\mathrm{Vec}_F)$ and Proposition 4.5(a) shows that $((\mathcal{K}_d^{\leq 0})_{\mathsf{b}}, (\mathcal{K}_d^{\geq 0})_{\mathsf{b}})$ is the standard t-structure on $K^b(\mathrm{Vec}_F)$.

It remains to consider the case when b is a nonsemisimple block. Choose a nontrivial root of unity q such that $q + q^{-1} \in F$ (for example, a primitive cubic root of unity ζ will work for any F since $\zeta + \zeta^{-1} = -1 \in F$). Then there is a nonsemisimple block in $TL(q)$ which is equivalent to b (Proposition 2.9). Hence, by Corollary 4.6, $((\mathcal{K}_d^{\leq 0})_{\mathsf{b}}, (\mathcal{K}_d^{\geq 0})_{\mathsf{b}})$ is a t-structure on $(\mathcal{K}_d)_{\mathsf{b}}$. □

4C5. *Complements.* The proof in Section 4C4 implies the following:

Corollary 4.7. (a) *The category* $\mathcal{K}_d^0$ *is pre-Tannakian.*

(b) *Any object of the category* $\mathcal{K}_d^0$ *is isomorphic to a subquotient of a direct sum of tensor powers of* [pt].

Proof. We already know that the category $\mathcal{K}_d^0$ is an abelian tensor category (see Corollary 4.3). It is obvious that Hom's are finite-dimensional and $\mathrm{End}(\mathbf{1}) = F$ since this is true in the category $\mathcal{K}_d$. The category $\mathcal{K}_d^0$ is rigid: if $\Delta \otimes K$ is concentrated in degree zero then the same is true for $\Delta \otimes K^* \simeq (\Delta \otimes K)^*$. It remains to check that any object of $\mathcal{K}_d^0$ has finite length. It is clear that we can verify this block by block. The result is clear for semisimple blocks since by Proposition 4.5(a) the core of the corresponding t-structure identifies with Vec_F. This is also clear for nonsemisimple blocks since the corresponding t-structure (described in Proposition 4.5) identifies with the t-structure on a block of the Temperley–Lieb category and the corresponding core has all objects of finite length since this is true for the category $\mathcal{C}_q$. This proves (a).

For (b) we use the same argument as above: it is sufficient to verify the statement block by block. Here the result is trivial for semisimple blocks and is known for nonsemisimple ones since it is known to hold for the category $\mathcal{C}_q$. □

Remark 4.8. Using similar techniques of importing known results about the category $\mathscr{C}_q$ to the category $\mathcal{K}_d^0$ we can obtain detailed information about this category. In particular, we see that the category $\mathcal{K}_d^0$ has enough projective objects; all indecomposable projective objects are direct summands of tensor powers of [pt] (but powers of [pt] are not projective in general; for example $[\mathrm{pt}]^{\otimes 0}=\mathbf{1}$ is not projective). Thus Corollary 4.7(b) can be improved: any object of the category $\mathcal{K}_d^0$ is isomorphic to a quotient of a direct sum of tensor powers of [pt].

5. Universal property

5A. *Extension property of the category* $\mathcal{K}_d^0$.

Proposition 5.1. *Let $\mathcal{T}$ be a pre-Tannakian category and let $\mathcal{F}:\underline{\mathrm{Rep}}(S_d)\to\mathcal{T}$ be a tensor functor. Assume that $\mathcal{F}(\Delta)\neq 0$. Then the functor $\mathcal{F}$ (uniquely) factorizes as $\underline{\mathrm{Rep}}(S_d)\to\mathcal{K}_d^0\to\mathcal{T}$, where $\mathcal{K}_d^0\to\mathcal{T}$ is an exact tensor functor.*

Proof. Let $K\in\mathcal{K}_d^0$. We can consider $\mathcal{F}(K)\in K^b(\mathcal{T})$. Since the category $\mathcal{T}$ is abelian we can talk about cohomology of $\mathcal{F}(K)$.

Lemma 5.2. $H^i(\mathcal{F}(K))=0$ *for* $i\neq 0$.

Proof. Notice that for any $0\neq X\in\mathcal{T}$ we have $X\otimes\mathcal{F}(\Delta)\neq 0$. Since the endofunctor $-\otimes\mathcal{F}(\Delta)$ of the category $\mathcal{T}$ is exact (see, e.g., [Bakalov and Kirillov 2001, Proposition 2.1.8]) we see that $H^i(\mathcal{F}(K\otimes\Delta))=H^i(\mathcal{F}(K)\otimes\mathcal{F}(\Delta))=H^i(\mathcal{F}(K))\otimes\mathcal{F}(\Delta)$. By the definition of $\mathcal{K}_d^0$ the cohomology of $\mathcal{F}(K\otimes\Delta)$ is concentrated in degree zero and we are done. □

We now define the functor $\mathcal{K}_d^0\to\mathcal{T}$ as $K\mapsto H^0(\mathcal{F}(K))$ with the tensor structure induced by the one on $\mathcal{F}$ (or rather its extension to $K^b(\underline{\mathrm{Rep}}(S_d))\to\mathcal{K}^b(\mathcal{T})$). □

Remark 5.3. Here is an example of a tensor functor between abelian rigid tensor categories which is not exact. Let k be a field of characteristic 2 and consider the category $\mathrm{Rep}_k(\mathbb{Z}/2\mathbb{Z})$ of finite-dimensional k-representations of $\mathbb{Z}/2\mathbb{Z}$. This category has precisely 2 indecomposable objects: one is simple and 1-dimensional; the other is projective and has categorical dimension 0. Thus the quotient of $\mathrm{Rep}_k(\mathbb{Z}/2\mathbb{Z})$ by the negligible morphisms is equivalent to the category Vec_k of finite-dimensional vector spaces over k. Clearly the quotient functor $\mathrm{Rep}_k(\mathbb{Z}/2\mathbb{Z})\to\mathrm{Vec}_k$ is not exact since it sends the projective object to zero. One can also construct a similar example over a field of characteristic zero using the representation category of the additive supergroup of a 1-dimensional odd space.

5B. *Fundamental groups of* $\mathcal{K}_d^0$ *and* $\mathbf{Rep}(S_d)$.

Let π be the fundamental group of the pre-Tannakian category $\mathcal{K}_d^0$. The action of π on $[\mathrm{pt}]\in\underline{\mathrm{Rep}}(S_d)\subset\mathcal{K}_d^0$ defines a homomorphism $\pi\to S_{\mathbf{I}}$ where $\mathbf{I}=\mathrm{Spec}([\mathrm{pt}])$.

Proposition 5.4. *The homomorphism $\varepsilon:\pi\to S_{\mathbf{I}}$ is in fact an isomorphism.*

Proof. Since the object [pt] generates $\mathcal{K}_d^0$ (see Corollary 4.7(b)) the homomorphism $\varepsilon : \pi \to S_{\mathbf{I}}$ is an embedding.

Consider the category $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$. It is shown in (the proof of) [Deligne 2007, Proposition B1] that its fundamental group is precisely the group $S_{\mathbf{I}}^{\varepsilon} = \mathrm{Aut}_{\mathrm{Rep}(S_{\mathbf{I}},\varepsilon)}(\mathbf{I})$. We have an obvious tensor functor $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$; by Proposition 5.1 it extends to a tensor functor $\mathcal{F} : \mathcal{K}_d^0 \to \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$. Thus we have a homomorphism $S_{\mathbf{I}}^{\varepsilon} \to \mathcal{F}(\pi)$. It is clear that the composition $S_{\mathbf{I}}^{\varepsilon} \to \mathcal{F}(\pi) \subset \mathcal{F}(S_{\mathbf{I}}) = S_{\mathbf{I}}^{\varepsilon}$ is the identity map. The result follows. □

5C. *Proof of Theorem 1.2.* We start with the following result:

Proposition 5.5 [Deligne 1990, 8.14(ii)]. *The fundamental group of the category* $\mathrm{Rep}(S_d)$ *is the group* S_d *acting on itself by conjugation.* □

Remark 5.6. It is explained in *loc. cit.* that we can replace S_d with any affine algebraic group G in the statement of the proposition.

Theorem 5.7. *Let* $\mathcal{T}$ *be a pre-Tannakian category and let* $\mathcal{F} : \underline{\mathrm{Rep}}(S_d) \to \mathcal{T}$ *be a tensor functor with* $T = \mathcal{F}([\mathrm{pt}])$.

(a) *If* $\mathcal{F}(\Delta) = 0$, *then the category* $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ *endowed with the functor* $\mathcal{F}_T : \underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ *is equivalent to* $\mathrm{Rep}(S_d)$ *equipped with the functor* $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_d)$.

(b) *If* $\mathcal{F}(\Delta) \neq 0$, *then the category* $\mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ *endowed with the functor* $\mathcal{F}_T : \underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_{\mathbf{I}}, \varepsilon)$ *is equivalent to* $\mathcal{K}_d^0$ *equipped with the functor* $\underline{\mathrm{Rep}}(S_d) \to \mathcal{K}_d^0$.

Proof. (a) In this case $\mathcal{F}$ factorizes as $\underline{\mathrm{Rep}}(S_d) \to \mathrm{Rep}(S_d) \to \mathcal{T}$ (see Corollary 3.16). The result follows from [Deligne 1990, Théorème 8.17] and Proposition 5.5.

(b) In this case $\mathcal{F}$ extends to a functor $\underline{\mathrm{Rep}}(S_d) \to \mathcal{K}_d^0 \to \mathcal{T}$ by Proposition 5.1. The result follows from [Deligne 1990, Théorème 8.17] and Proposition 5.4. □

If we apply Theorem 5.7(b) to the category $\mathcal{T} = \underline{\mathrm{Rep}}(S_{-1})$ and the functor $\underline{\mathrm{Res}}_{S_{-1}}^{S_d} : \underline{\mathrm{Rep}}(S_d) \to \mathcal{T}$ described in Definition 2.4 and Section 3D we obtain the following:

Corollary 5.8. *The category* $\underline{\mathrm{Rep}}^{ab}(S_d)$ *endowed with the functor* $\underline{\mathrm{Rep}}(S_d) \to \underline{\mathrm{Rep}}^{ab}(S_d)$ *is equivalent to the category* $\mathcal{K}_d^0$ *with the functor* $\underline{\mathrm{Rep}}(S_d) \to \mathcal{K}_d^0$. □

Clearly Theorem 5.7 and Corollary 5.8 together imply Theorem 1.2.

Acknowledgments

This paper owes its existence to Pierre Deligne, who explained to the second author a proof of Lemma 3.11, which is crucial to this paper, during a visit to the Institute for Advanced Study. Both authors are happy to express their deep gratitude to him and to the Institute for Advanced Study, which made this interaction possible. The authors are also very grateful to Alexander Kleshchev, who initiated this project. We also thank Michael Finkelberg and Friedrich Knop for their interest in this work and Darij Grinberg for his detailed comments. The second author's work was partially supported by NSF grant DMS-0602263.

References

[Andersen 1992] H. H. Andersen, "Tensor products of quantized tilting modules", *Comm. Math. Phys.* **149**:1 (1992), 149–159. MR 94b:17015 Zbl 0760.17004

[Andersen et al. 1991] H. H. Andersen, P. Polo, and K. X. Wen, "Representations of quantum algebras", *Invent. Math.* **104**:1 (1991), 1–59. MR 92e:17011 Zbl 0724.17012

[Bakalov and Kirillov 2001] B. Bakalov and A. Kirillov, Jr., *Lectures on tensor categories and modular functors*, University Lecture Series **21**, Amer. Math. Soc., Providence, RI, 2001. MR 2002d:18003 Zbl 0965.18002

[Beĭlinson et al. 1982] A. A. Beĭlinson, J. Bernstein, and P. Deligne, "Faisceaux pervers", pp. 5–171 in *Analysis and topology on singular spaces, I* (Luminy, 1981)), Astérisque **100**, Soc. Math. France, Paris, 1982. MR 86g:32015 Zbl 0536.14011

[Comes and Ostrik 2011] J. Comes and V. Ostrik, "On blocks of Deligne's category $\underline{\mathrm{Rep}}(S_t)$", *Adv. Math.* **226**:2 (2011), 1331–1377. MR 2012b:20020 Zbl 1225.18005

[Deligne 1990] P. Deligne, "Catégories tannakiennes", pp. 111–195 in *The Grothendieck Festschrift, II*, edited by P. Cartier et al., Progr. Math. **87**, Birkhäuser, Boston, 1990. MR 92d:14002 Zbl 0727.14010

[Deligne 2007] P. Deligne, "La catégorie des représentations du groupe symétrique S_t, lorsque t n'est pas un entier naturel", pp. 209–273 in *Algebraic groups and homogeneous spaces*, edited by V. B. Mehta, Tata Inst. Fund. Res., Mumbai, 2007. MR 2009b:20021 Zbl 1165.20300

[Freedman 2003] M. H. Freedman, "A magnetic model with a possible Chern–Simons phase", *Comm. Math. Phys.* **234**:1 (2003), 129–183. MR 2004b:81193 Zbl 1060.81054

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory: A first course*, Graduate Texts in Mathematics **129**, Springer, New York, 1991. MR 93a:20069 Zbl 0744.22001

[Kashiwara and Schapira 2006] M. Kashiwara and P. Schapira, *Categories and sheaves*, Grundlehren der Mathematischen Wissenschaften **332**, Springer, Berlin, 2006. MR 2006k:18001 Zbl 1118.18001

[Ostrik 2008] V. Ostrik, "Module categories over representations of $\mathrm{SL}_q(2)$ in the non-semisimple case", *Geom. Funct. Anal.* **17**:6 (2008), 2005–2017. MR 2009a:20082 Zbl 1139.18007

Communicated by Susan Montgomery
Received 2013-04-11 Accepted 2013-08-31

jcomes@math.uoregon.edu *Department of Mathematics, University of Oregon, Eugene, OR 97403, United States*

vostrik@math.uoregon.edu *Department of Mathematics, University of Oregon, Eugene, OR 97403, United States*

ALGEBRA AND NUMBER THEORY 8:2 (2014)
dx.doi.org/10.2140/ant.2014.8.497

Algebraicity of the zeta function associated to a matrix over a free group algebra

Christian Kassel and Christophe Reutenauer

Following and generalizing a construction by Kontsevich, we associate a zeta function to any matrix with entries in a ring of noncommutative Laurent polynomials with integer coefficients. We show that such a zeta function is an algebraic function.

1. Introduction

Fix a commutative ring K. Let F be a free group on a finite number of generators $X_1, \ldots, X_n$ and

$$KF = K\langle X_1, X_1^{-1}, \ldots, X_n, X_n^{-1}\rangle$$

be the corresponding group algebra: equivalently, it is the algebra of noncommutative Laurent polynomials with coefficients in K. Any element $a \in KF$ can be uniquely written as a finite sum of the form

$$a = \sum_{g\in F} (a, g) g,$$

where $(a, g) \in K$.

Let M be a $d \times d$ matrix with coefficients in KF. For any $n \geq 1$, we may consider the n-th power M^n of M and its trace $\mathrm{Tr}(M^n)$, which is an element of KF. We define the integer $a_n(M)$ as the coefficient of 1 in the trace of M^n:

$$a_n(M) = (\mathrm{Tr}(M^n), 1). \tag{1-1}$$

Let g_M and P_M be the formal power series

$$g_M = \sum_{n\geq 1} a_n(M) t^n \quad \text{and} \quad P_M = \exp\biggl(\sum_{n\geq 1} a_n(M) \frac{t^n}{n}\biggr). \tag{1-2}$$

They are related by

$$g_M = t\,\frac{d \log(P_M)}{dt}.$$

MSC2010: primary 05A15, 68Q70, 68R15; secondary 05E15, 14H05, 14G10.
Keywords: noncommutative formal power series, language, zeta function, algebraic function.

We call P_M the *zeta function* of the matrix M by analogy with the zeta function of a noncommutative formal power series (see next section); the two concepts will be related in Proposition 4.1.

The motivation for the definition of P_M comes from the well-known identity expressing the inverse of the reciprocal polynomial of the characteristic polynomial of a matrix M with entries in a commutative ring

$$\frac{1}{\det(1-tM)} = \exp\biggl(\sum_{n\geq 1} \mathrm{Tr}(M^n)\frac{t^n}{n}\biggr).$$

Note that, for any scalar $\lambda \in K$, the corresponding series for the matrix λM become

$$g_{\lambda M}(t) = g_M(\lambda t) \quad \text{and} \quad P_{\lambda M}(t) = P_M(\lambda t). \tag{1-3}$$

Our main result is the following; it was inspired by Theorem 1 of [Kontsevich 2011]:

Theorem 1.1. *For each matrix $M \in M_d(KF)$ where $K = \mathbb{Q}$ is the ring of rational numbers, the formal power series P_M is algebraic.*

The special case $d = 1$ is due to Kontsevich [2011]. A combinatorial proof in the case $d = 1$ and F is a free group on one generator appears in [Reutenauer and Robado 2012].

Observe that by the rescaling equalities (1-3) it suffices to prove the theorem when $K = \mathbb{Z}$ is the ring of integers.

It is crucial for the veracity of Theorem 1.1 that the variables do not commute: for instance, if $a = x + y + x^{-1} + y^{-1} \in \mathbb{Z}[x, x^{-1}, y, y^{-1}]$, where x and y are commuting variables, then $\exp(\sum_{n\geq 1}(a^n, 1)t^n/n)$ is a formal power series with integer coefficients but not an algebraic function (this follows from Example 3 in [Bousquet-Mélou 2005, §1]).

The paper is organized as follows. In Section 2, we define the zeta function ζ_S of a noncommutative formal power series S and show that it can be expanded as an infinite product under a cyclicity condition that is satisfied by the characteristic series of cyclic languages.

In Section 3, we recall the notion of algebraic noncommutative formal power series and some of their properties.

In Section 4, we reformulate the zeta function of a matrix as the zeta function of a noncommutative formal power series before giving the proof of Theorem 1.1; the latter follows the steps sketched in [Kontsevich 2011] and relies on the results of the previous sections as well as on an algebraicity result by André [2004] elaborating on an idea of D. and G. Chudnovsky.

We concentrate on two specific matrices in Section 5. We give a closed formula for the zeta function of the first matrix; its nonzero coefficients count the planar rooted bicubic maps as well as Chapoton's "new intervals" in a Tamari lattice (see [Chapoton 2006; Tutte 1963]).

2. Cyclic formal power series

General definitions. As usual, if A is a set, we denote by A^* the free monoid on A: it consists of all words on the alphabet A, including the empty word 1.

Let $A^+ = A - \{1\}$. Recall that $w \in A^+$ is *primitive* if it cannot be written as u^r for any integer $r \geq 2$ and any $u \in A^+$. Two elements $w, w' \in A^+$ are *conjugate* if $w = uv$ and $w' = vu$ for some $u, v \in A^*$.

Given a set A and a commutative ring K, let $K\langle\langle A\rangle\rangle$ be the algebra of noncommutative formal power series on the alphabet A. For any element $S \in K\langle\langle A\rangle\rangle$ and any $w \in A^*$, we define the coefficient $(S, w) \in K$ by

$$S = \sum_{w \in A^*} (S, w) w.$$

As an example of such noncommutative formal power series, take the characteristic series $\sum_{w \in L} w$ of a language $L \subseteq A^*$. In the sequel, we shall identify a language with its characteristic series.

The *generating series* g_S of an element $S \in K\langle\langle A\rangle\rangle$ is the image of S under the algebra map $\varepsilon : K\langle\langle A\rangle\rangle \to K[[t]]$ sending each $a \in A$ to the variable t. We have

$$g_S - (S, 1) = \sum_{w \in A^+} (S, w) t^{|w|} = \sum_{n \geq 1} \Biggl(\sum_{|w|=n} (S, w) \Biggr) t^n, \tag{2-1}$$

where $|w|$ is the length of w.

The *zeta function* ζ_S of $S \in K\langle\langle A\rangle\rangle$ is defined by

$$\zeta_S = \exp\Biggl(\sum_{w \in A^+} (S, w) \frac{t^{|w|}}{|w|} \Biggr) = \exp\Biggl(\sum_{n \geq 1} \Biggl(\sum_{|w|=n} (S, w) \Biggr) \frac{t^n}{n} \Biggr). \tag{2-2}$$

The formal power series g_S and ζ_S are related by

$$t \frac{d \log(\zeta_S)}{dt} = t \frac{\zeta_S'}{\zeta_S} = g_S - (S, 1), \tag{2-3}$$

where ζ_S' is the derivative of ζ_S with respect to the variable t.

Cyclicity.

Definition 2.1. An element $S \in K\langle\langle A\rangle\rangle$ is cyclic if

(i) $\forall u, v \in A^*, \ (S, uv) = (S, vu)$ and

(ii) $\forall w \in A^+, \ \forall r \geq 2, \ (S, w^r) = (S, w)^r$.

Cyclic languages provide examples of cyclic formal power series. Recall from [Berstel and Reutenauer 1990, §2] that a language $L \subseteq A^*$ is *cyclic* if

(1) $\forall u, v \in A^*, \quad uv \in L \Longleftrightarrow vu \in L$ and

(2) $\forall w \in A^+, \forall r \geq 2, \quad w^r \in L \Longleftrightarrow w \in L.$

The characteristic series of a cyclic language is a cyclic formal power series in the above sense.

Let L be any set of representatives of conjugacy classes of primitive elements of A^+.

Proposition 2.2. *If $S \in K\langle\langle A\rangle\rangle$ is a cyclic formal power series, then*

$$\zeta_S = \prod_{\ell \in L} \frac{1}{1-(S, \ell)t^{|\ell|}}.$$

Proof. Since both sides of the equation have the same constant term 1, it suffices to prove that they have the same logarithmic derivative. The logarithmic derivative of the right-hand side multiplied by t is equal to

$$\sum_{\ell \in L} \frac{|\ell|(S, \ell)t^{|\ell|}}{1-(S, \ell)t^{|\ell|}},$$

which in turn is equal to

$$\sum_{\ell \in L,\ k \geq 1} |\ell|(S, \ell)^k t^{k|\ell|}.$$

In view of (2-1) and (2-3), it is enough to check that, for all $n \geq 1$,

$$\sum_{|w|=n} (S, w) = \sum_{\ell \in L,\ k \geq 1,\ k|\ell|=n} |\ell|(S, \ell)^k. \tag{2-4}$$

Now any word $w = u^k$ is the k-th power of a unique primitive word u, which is the conjugate of a unique element $\ell \in L$. Moreover, w has exactly $|\ell|$ conjugates and, since S is cyclic, we have

$$(S, w) = (S, u^k) = (S, u)^k = (S, \ell)^k.$$

From this, Equation (2-4) follows immediately. □

Corollary 2.3. *If a cyclic formal power series S has integer coefficients, that is, if $(S, w) \in \mathbb{Z}$ for all $w \in A^*$, then so does ζ_S.*

3. Algebraic noncommutative series

This section is essentially a compilation of well-known results on algebraic noncommutative series.

Recall that a *system of proper algebraic noncommutative equations* is a finite set of equations

$$\xi_i = p_i, \quad i = 1, \ldots, n,$$

where $\xi_1, \ldots, \xi_n$ are noncommutative variables and $p_1, \ldots, p_n$ are elements of $K\langle \xi_1, \ldots, \xi_n, A\rangle$, where A is some alphabet. We assume that each p_i has no constant term and contains no monomial ξ_j. One can show that such a system has a unique solution $(S_1, \ldots, S_n)$, i.e., there exists a unique n-tuple $(S_1, \ldots, S_n) \in K\langle\langle A\rangle\rangle^n$ such that $S_i = p_i(S_1, \ldots, S_n, A)$ for all $i = 1, \ldots, n$ and each S_i has no constant term (see [Schützenberger 1962], [Salomaa and Soittola 1978, Theorem IV.1.1], or [Stanley 1999, Proposition 6.6.3]).

If a formal power series $S \in K\langle\langle A\rangle\rangle$ differs by a constant from such a formal power series S_i, we say that S is *algebraic*.

Example 3.1. Consider the proper algebraic noncommutative equation

$$\xi = a\xi^2 + b.$$

(Here $A = \{a, b\}$.) Its solution is of the form

$$S = b + abb + aabbb + ababb + \cdots .$$

One can show (see [Berstel 1979]) that S is the characteristic series of Łukasiewicz's language, namely of the set of words $w \in \{a, b\}^*$ such that $|w|_b = |w|_a + 1$ and $|u|_a \geq |u|_b$ for all proper prefixes u of w.

Recall also that $S \in K\langle\langle A\rangle\rangle$ is *rational* if it belongs to the smallest subalgebra of $K\langle\langle A\rangle\rangle$ containing $K\langle A\rangle$ and closed under inversion. By a theorem of Schützenberger (see [Berstel and Reutenauer 2011, Theorem I.7.1]), a formal power series $S \in K\langle\langle A\rangle\rangle$ is rational if and only if it is *recognizable*, i.e., there exist an integer $n \geq 1$, a representation μ of the free monoid A^* by matrices with entries in K, a row-matrix α and a column-matrix β such that, for all $w \in A^*$,

$$(S, w) = \alpha\mu(w)\beta.$$

We now record two well-known theorems.

Theorem 3.2. (1) *If $S \in K\langle\langle A\rangle\rangle$ is algebraic, then its generating series $g_S \in K[[t]]$ is algebraic in the usual sense.*

(2) *The set of algebraic power series is a subring of $K\langle\langle A\rangle\rangle$.*

(3) *A rational power series is algebraic.*

(4) *The Hadamard product of a rational power series and an algebraic power series is algebraic.*

(5) *Let* $A = \{a_1, \dots, a_n, a_1^{-1}, \dots, a_n^{-1}\}$ *and* L *be the language consisting of all words on the alphabet* A *whose image in the free group on* $a_1, \dots, a_n$ *is the neutral element. Then the characteristic series of* L *is algebraic.*

Items (1)–(4) of the previous theorem are due to Schützenberger [1962] and Item (5) to Chomsky and Schützenberger [1963] (see [Stanley 1999, Example 6.6.8]).

The second theorem is a criterion due to Jacob [1975].

Theorem 3.3. *A formal power series* $S \in K\langle\langle A\rangle\rangle$ *is algebraic if and only if there exist a free group* F, *a representation* μ *of the free monoid* A^* *by matrices with entries in* KF, *indices* i *and* j, *and an element* $\gamma \in F$ *such that, for all* $w \in A^*$,

$$(S, w) = ((\mu w)_{i,j}, \gamma).$$

The following is an immediate consequence of Theorem 3.3:

Corollary 3.4. *If* $S \in K\langle\langle A\rangle\rangle$ *is an algebraic power series and* $\varphi : B^* \to A^*$ *is a homomorphism of finitely generated free monoids, then the power series*

$$\sum_{w \in B^*} (S, \varphi(w)) w \in K\langle\langle B\rangle\rangle$$

is algebraic.

As a consequence of Theorem 3.2(5) and of Corollary 3.4, we obtain:

Corollary 3.5. *Let* $f : A^* \to F$ *be a homomorphism from* A^* *to a free group* F. *Then the characteristic series of* $f^{-1}(1) \in K\langle\langle A\rangle\rangle$ *is algebraic.*

4. Proof of Theorem 1.1

Let M be a $d \times d$ matrix. As observed in the introduction, it is enough to establish Theorem 1.1 when all the entries of M belong to $\mathbb{Z}F$.

We first reformulate the formal power series g_M and P_M of (1-2) as the generating series and the zeta function of a noncommutative formal power series, respectively.

Let A be the alphabet whose elements are triples $[g, i, j]$, where i and j are integers such that $1 \le i, j \le d$ and $g \in F$ appears in the (i, j)-entry $M_{i,j}$ of M, i.e., $(M_{i,j}, g) \neq 0$. We define the noncommutative formal power series $S_M \in K\langle\langle A\rangle\rangle$ as follows: for $w = [g_1, i_1, j_1] \cdots [g_n, i_n, j_n] \in A^+$, the scalar (S_M, w) vanishes unless we have

(a) $j_n = i_1$ and $j_k = i_{k+1}$ for all $k = 1, \dots, n-1$ and

(b) $g_1 \cdots g_n = 1$ in the group F, in which case (S_M, w) is given by

$$(S_M, w) = (M_{i_1,j_1}, g_1) \cdots (M_{i_n,j_n}, g_n) \in K.$$

By convention, $(S_M, 1) = d$.

Proposition 4.1. *The generating series and the zeta function of S_M are related to the formal power series g_M and P_M of* (1-2) *by*

$$g_{S_M} - d = g_M \quad \textit{and} \quad \zeta_{S_M} = P_M.$$

Proof. For $n \geq 1$, we have

$$\begin{aligned}\mathrm{Tr}(M^n) &= \sum M_{i_1,j_1} \cdots M_{i_n,j_n} \\ &= \sum (M_{i_1,j_1}, g_1) \cdots (M_{i_n,j_n}, g_n) g_1 \cdots g_n,\end{aligned}$$

where the sum runs over all indices $i_1, j_1, \ldots, i_n, j_n$ satisfying Condition (a) above and over all $g_1, \ldots, g_n \in F$. Then

$$a_n(M) = (\mathrm{Tr}(M^n), 1) = \sum (M_{i_1,j_1}, g_1) \cdots (M_{i_n,j_n}, g_n),$$

where Conditions (a) and (b) are satisfied. Hence,

$$a_n(M) = \sum_{w \in A^*,\ |w|=n} (S, w),$$

which proves the proposition in view of (1-2), (2-1) and (2-2). □

We next establish that S_M is both cyclic in the sense of Section 2 and algebraic in the sense of Section 3.

Proposition 4.2. *The noncommutative formal power series S_M is cyclic.*

Proof. (i) Conditions (a) and (b) above are clearly preserved under cyclic permutations. Hence, we also have

$$(S_M, w) = (M_{i_2,j_2}, g_2) \cdots (M_{i_n,j_n}, g_n)(M_{i_1,j_1}, g_1)$$

when $w = [g_1, i_1, j_1] \cdots [g_n, i_n, j_n]$ such that Conditions (a) and (b) are satisfied. It follows that $(S, uv) = S(vu)$ for all $u, v \in A^*$.

(ii) If w satisfies Conditions (a) and (b), so does w^r for $r \geq 2$. Conversely, if w^r $(r \geq 2)$ satisfies Condition (a), then since

$$w^r = [g_1, i_1, j_1] \cdots [g_n, i_n, j_n][g_1, i_1, j_1] \cdots$$

we must have $j_n = i_1$ and $j_k = i_{k+1}$ for all $k = 1, \ldots, n-1$, and so w satisfies Condition (a).

If w^r $(r \geq 2)$ satisfies Condition (b), i.e., $(g_1 \cdots g_n)^r = 1$, then $g_1 \cdots g_n = 1$ since F is torsion-free. Hence, w satisfies Condition (b). It follows that $(S, w^r) = ((M_{i_1,j_1}, g_1) \cdots (M_{i_n,j_n}, g_n))^r = (S, w)^r$. □

Proposition 4.3. *The noncommutative formal power series S_M is algebraic.*

Proof. We write S_M as the Hadamard product of three noncommutative formal power series S_1, S_2 and S_3.

The series $S_1 \in K\langle\langle A\rangle\rangle$ is defined for $w = [g_1, i_1, j_1] \cdots [g_n, i_n, j_n] \in A^+$ by

$$(S_1, w) = (M_{i_1,j_1}, g_1) \cdots (M_{i_n,j_n}, g_n)$$

and by $(S_1, 1) = 1$. This is a recognizable, hence rational, series with one-dimensional representation $A^* \to K$ given by $[g, i, j] \mapsto (M_{i,j}, g)$.

Next consider the representation μ of the free monoid A^* defined by

$$\mu([g, i, j]) = E_{i,j},$$

where $E_{i,j}$ denotes as usual the $d \times d$ matrix with all entries vanishing except the (i, j)-entry, which is equal to 1. Set

$$S_2 = \sum_{w \in A^*} \mathrm{Tr}((\mu w))w \in K\langle\langle A\rangle\rangle.$$

The power series S_2 is recognizable and hence rational. Let us describe S_2 more explicitly. For $w = 1$, $\mu(w)$ is the identity $d \times d$ matrix; hence, $(S_2, 1) = d$. For $w = [g_1, i_1, j_1] \cdots [g_n, i_n, j_n] \in A^+$, we have

$$\mathrm{Tr}((\mu w)) = \mathrm{Tr}(E_{i_1,j_1} \cdots E_{i_n,j_n}).$$

It follows that $\mathrm{Tr}((\mu w)) \neq 0$ if and only if $\mathrm{Tr}(E_{i_1,j_1} \cdots E_{i_n,j_n}) \neq 0$, which is equivalent to $j_n = i_1$ and $j_k = i_{k+1}$ for all $k = 1, \dots, n-1$, in which case $\mathrm{Tr}((\mu w)) = 1$. Thus,

$$S_2 = d + \sum_{n\geq 1} \sum [g_1, i_1, i_2][g_2, i_2, i_3] \cdots [g_n, i_n, i_1],$$

where the second sum runs over all elements $g_1, \dots, g_n \in F$ and all indices $i_1, \dots, i_n$.

Finally, consider the homomorphism $f : A^* \to F$ sending $[g, i, j]$ to g. Then by Corollary 3.5 the characteristic series $S_3 \in K\langle\langle A\rangle\rangle$ of $f^{-1}(1)$ is algebraic.

It is now clear that S_M is the Hadamard product of S_1, S_2 and S_3:

$$S_M = S_1 \odot S_2 \odot S_3.$$

Since, by [Berstel and Reutenauer 2011, Theorem I.5.5] the Hadamard product of two rational series is rational, $S_1 \odot S_2$ is rational as well. It then follows from Theorem 3.2(4) and the algebraicity of S_3 that $S_M = S_1 \odot S_2 \odot S_3$ is algebraic. □

Since M has entries in $\mathbb{Z}F$, the power series $g_{S_M} = g_M + d$ belongs to $\mathbb{Z}[[t]]$. It follows by Corollary 2.3 and Proposition 4.2 that the power series $P_M = \zeta_{S_M}$ has integer coefficients as well. Moreover, by Theorem 3.2(1) and Proposition 4.3,

$$t\frac{d\log(P_M)}{dt} = g_M$$

is algebraic.

To complete the proof of Theorem 1.1, it suffices to apply the following algebraicity theorem:

Theorem 4.4. *If $f \in \mathbb{Z}[\![t]\!]$ is a formal power series with integer coefficients such that $t\, d\log f/dt$ is algebraic, then f is algebraic.*

Note that the integrality condition for f is essential: for the transcendental formal power series $f = \exp(t)$, we have $t\, d\log f/dt = t$, which is even rational.

Proof. This result follows from cases of the Grothendieck–Katz conjecture proved in [André 2004] and in [Bost 2001]. The conjecture states that, if $Y' = AY$ is a linear system of differential equations with $A \in M_d(\mathbb{Q}(t))$, then far from the poles of A it has a basis of solutions that are algebraic over $\mathbb{Q}(t)$ if and only if for almost all prime numbers p the reduction mod p of the system has a basis of solutions that are algebraic over $\mathbb{F}_p(t)$.

Let us now sketch a proof of the theorem (see also Exercise 5 of [André 1989, p. 160]). Set $g = tf'/f$, and consider the system $y' = (g/t)y$; it defines a differential form ω on an open set S of the smooth projective complete curve $\bar{S}$ associated to g. We now follow [André 2004, §6.3], which is inspired from [Chudnovsky and Chudnovsky 1985]. First, extend ω to a section (still denoted ω) of $\Omega^1_{\bar{S}}(-D)$, where D is the divisor of poles of ω. For any $n \geq 2$, we have a differential form $\sum_{i=1}^n p_i^*(\omega)$ on S^n, where $p_i : S^n \to S$ is the i-th canonical projection; this form goes down to the symmetric power $S^{(n)}$. Now let J be the generalized Jacobian of S parametrizing invertible fiber bundles over $\bar{S}$ that are rigidified over D. There is a morphism $\varphi : S \to J$ and a unique invariant differential form ω_J on J such that $\omega = \varphi^*(\omega_J)$. For any $n \geq 2$, φ induces a morphism $\varphi^{(n)} : S^{(n)} \to J$ such that $(\varphi^{(n)})^*(\omega_J) = \sum_{i=1}^n p_i^*(\omega)$. For n large enough, $\varphi^{(n)}$ is dominant, and if ω_J is exact, then so is ω. To prove that ω_J is exact, we note that J, being a scheme of commutative groups, is uniformized by $\mathbb{C}^n$. We can now apply Theorem 5.4.3 of [André 2004], whose hypotheses are satisfied because the solution f of the system has integer coefficients.

Alternatively, one can use a special case of a generalized Grothendieck–Katz conjecture proved by Bost, namely Corollary 2.8 in [Bost 2001, §2.4]: the vanishing of the p-curvatures in Condition (i) follows by a theorem of Cartier from the fact that the system has a solution in $\mathbb{F}_p(t)$, namely the reduction mod p of f for all prime numbers p for which such a reduction of the system exists (see Exercise 3 of [André 1989, p. 84] or Theorem 5.1 of [Katz 1970]); Condition (ii) is satisfied since $\mathbb{C}^n$ satisfies the Liouville property. □

A nice overview of such algebraicity results is given in the Bourbaki report of Chambert-Loir [2002]; see especially Theorem 2.6 and the following lines.

5. Examples

Kontsevich [2011] computed P_ω when $\omega = X_1 + X_1^{-1} + \cdots + X_n + X_n^{-1}$ considered as a 1×1 matrix, obtaining

$$P_\omega = \frac{2^n}{(2n-1)^{n-1}} \cdot \frac{(n-1+n(1-4(2n-1)t^2)^{1/2})^{n-1}}{(1+(1-4(2n-1)t^2)^{1/2})^n}, \tag{5-1}$$

which shows that P_ω belongs to a quadratic extension of $\mathbb{Q}(t)$.

We now present similar results for the zeta functions of two matrices: the first one of order 2 and the second one of order $d \geq 3$.

Computing P_M for a 2×2 matrix. Consider the following matrix with entries in the ring $\mathbb{Z}\langle a, a^{-1}, b, b^{-1}, d, d^{-1}\rangle$, where a, b and d are noncommuting variables:

$$M = \begin{pmatrix} a + a^{-1} & b \\ b^{-1} & d + d^{-1} \end{pmatrix}. \tag{5-2}$$

Proposition 5.1. *We have*

$$g_M = 3\frac{(1-8t^2)^{1/2} - 1 + 6t^2}{1 - 9t^2}, \tag{5-3}$$

$$P_M = \frac{(1-8t^2)^{3/2} - 1 + 12t^2 - 24t^4}{32t^6}. \tag{5-4}$$

Expanding P_M as a formal power series, we obtain

$$P_M = 1 + \sum_{n \geq 1} \frac{3 \cdot 2^n}{(n+2)(n+3)} \binom{2n+2}{n+1} t^{2n}.$$

Proof. View the matrix M under the form of the graph of Figure 1 with two vertices 1 and 2 and six labeled oriented edges. We identify paths in this graph and words on the alphabet $A = \{a, a^{-1}, b, b^{-1}, d, d^{-1}\}$. Let B denote the set of nonempty words on A that become trivial in the corresponding free group on a, b and d and whose corresponding path is a closed path. Then the integer $a_n(M)$ is the number of words in B of length n. We have $\varepsilon(B) = g_M$, where $\varepsilon : K\langle\langle A\rangle\rangle \to K[[t]]$ is the algebra map defined in Section 2.

We define B_i ($i = 1, 2$) as the set of paths in B starting from and ending at the vertex i; we have $B = B_1 + B_2$. Each set B_i is a free subsemigroup of A^*, freely generated by the set C_i of closed paths not passing through i (except at their ends).

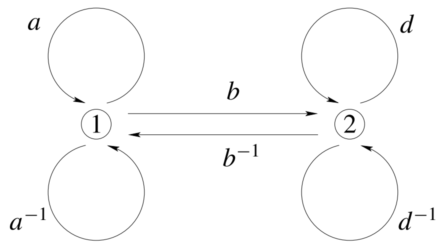


Figure 1. A graph representing M.

The sets C_i do not contain the empty word. We have

$$B_i = C_i^+ = \sum_{n\geq 1} C_i^n, \quad i = 1, 2.$$

Given a letter x, we denote by $C_i(x)$ the set of closed paths in C_i starting with x. Any word of $C_i(x)$ is of the form xwx^{-1}, where $w \in B_j$ when $i \xrightarrow{x} j$; such w does not start with x^{-1}. Identifying a language with its characteristic series and using the standard notation $L^* = 1 + \sum_{n\geq 1} L^n$ for any language L, we obtain the equations

$$C_1(a) = a(C_1(a) + C_1(b))^* a^{-1}, \tag{5-5}$$

$$C_1(b) = b(C_2(d) + C_2(d^{-1}))^* b^{-1}. \tag{5-6}$$

Applying the algebra map ε and taking into account the symmetries of the graph, we see that the four noncommutative formal power series $C_1(a)$, $C_1(a^{-1})$, $C_2(d)$ and $C_2(d^{-1})$ are sent to the same formal power series $u \in \mathbb{Z}[[t]]$ while $C_1(b)$ and $C_2(b^{-1})$ are sent to the same formal power series v. It follows from (5-5) and (5-6) that u and v satisfy the equations

$$u = t^2(u+v)^* = \frac{t^2}{1-u-v} \quad \text{and} \quad v = t^2(2u)^* = \frac{t^2}{1-2u}, \tag{5-7}$$

from which we deduce

$$t^2 = u(1-u-v) = v(1-2u).$$

The second equality is equivalent to $(u-v)(u-1) = 0$. Since $C_1(a)$ does not contain the empty word, the constant term of u vanishes; hence, $u - 1 \neq 0$. Therefore, $u = v$.

Since $C_1 = C_1(a) + C_1(a^{-1}) + C_1(b)$ and $C_2 = C_2(d) + C_2(d^{-1}) + C_2(b^{-1})$, we have $\varepsilon(C_1) = \varepsilon(C_2) = 2u + v = 3u$. Therefore, $\varepsilon(B_1) = \varepsilon(B_2) = 3u/(1-3u)$ and

$$g_M = \varepsilon(B) = \frac{6u}{1-3u}. \tag{5-8}$$

Let us now compute u using (5-7) and the equality $u = v$. The formal power series u satisfies the quadratic equation $2u^2 - u + t^2 = 0$. Since u has zero constant term, we obtain

$$u = \frac{1-(1-8t^2)^{1/2}}{4}.$$

From this and (5-8), we obtain the desired form for g_M.

Let $P(t)$ be the right-hand side in Equation (5-4). To prove $P_M = P(t)$, we checked that $tP'(t)/P(t) = g_M$ and the constant term of $P(t)$ is 1. □

Remark 5.2. We found Equation (5-4) for $P(t)$ as follows. We first computed the lowest coefficients of g_M up to degree 10:

$$g_M = 6(t^2+5t^4+29t^6+181t^8+1181t^{10}) + O(t^{12}).$$

From this, it was not difficult to find that

$$P_M = 1+3t^2+12t^4+56t^6+288t^8+1584t^{10}+O(t^{12}). \tag{5-9}$$

Up to a shift, the sequence (5-9) of nonzero coefficients of P_M is the same as the sequence of numbers of "new" intervals in a Tamari lattice computed in [Chapoton 2006, §9]. (We learnt this from [OEIS 2010], where this sequence is listed as A000257.) Chapoton gave an explicit formula for the generating function ν of these "new" intervals (see Equation (73) in [Chapoton 2006]). Rescaling ν, we found that $P(t) = (\nu(t^2)-t^4)/t^6$ has up to degree 10 the same expansion as (5-9). It then sufficed to check that $tP'(t)/P(t) = g_M$.

By [OEIS 2010], the integers in the sequence A000257 also count the number of planar rooted bicubic maps with $2n$ vertices (see [Tutte 1963, p. 269]). Planar maps also come up in the combinatorial interpretation of (5-1) given in [Reutenauer and Robado 2012, §5] for $n=2$.

Note that the sequence of nonzero coefficients of $g_M/6$ is listed as A194723 in [OEIS 2010].

A similar d × d matrix. Fix an integer $d \geq 3$, and let M be the $d\times d$ matrix with entries $M_{i,j}$ defined by

$$M_{i,i} = a_i + a_i^{-1} \quad \text{and} \quad M_{i,j} = \begin{cases} b_{ij} & \text{if } i<j, \\ b_{ji}^{-1} & \text{if } j<i, \end{cases}$$

where $a_1,\dots,a_d$, b_{ij} $(1\leq i<j\leq d)$ are noncommuting variables. This matrix is a straightforward generalization of (5-2).

Proceeding as above, we obtain two formal power series u and v satisfying the following equations similar to (5-7):

$$u = t^2(u+(d-1)v)^* = \frac{t^2}{1-u-(d-1)v},$$

$$v = t^2(2u+(d-2)v)^* = \frac{t^2}{1-2u-(d-2)v}.$$

We deduce the equality $u = v$ and the quadratic equation $u(1-du) = t^2$. We finally have

$$g_M = \frac{d(d+1)u}{1-(d+1)u},$$

which leads to

$$g_M = \frac{d(d+1)}{2}\,\frac{(1-4dt^2)^{1/2} - 1 + 2(d+1)t^2}{1-(d+1)^2t^2}.$$

Its expansion as a formal power series is the following:

$$g_M = d(d+1)\big\{t^2 + (2d+1)t^4 + (5d^2+4d+1)t^6 + (14d^3+14d^2+6d+1)t^8 \\ + (42d^4+48d^3+27d^2+8d+1)t^{10}\big\} + O(t^{12}).$$

When $d = 2, 3, 4$, the sequence of nonzero coefficients of $g_M/d(d+1)$ is listed respectively as A194723, A194724 and A194725 in [OEIS 2010] (it is also the d-th column in Sequence A183134). These sequences count the d-ary words, either empty or beginning with the first letter of the alphabet, that can be built by inserting n doublets into the initially empty word.

We were not able to find a closed formula for P_M analogous to (5-4). Using Maple, we found that, for instance up to degree 10, the expansion of P_M is

$$1 + \frac{d(d+1)}{2}t^2 + \frac{d(d+1)(d^2+5d+2)}{8}t^4 \\ + \frac{d(d+1)(d^4+14d^3+59d^2+38d+8)}{48}t^6 \\ + \frac{d(d+1)(d^6+27d^5+271d^4+1105d^3+904d^2+332d+48)}{384}t^8.$$

Acknowledgements

We are most grateful to Yves André, Jean-Benoît Bost and Carlo Gasbarri for their help in the proof of Theorem 4.4. We are also indebted to François Bergeron and Pierre Guillot for assisting us with computer computations in the process detailed in Remark 5.2, to Frédéric Chapoton for his comments on the last part of Section 5 and to an anonymous referee for having spotted slight inaccuracies. Thanks also to Stavros Garoufalidis for pointing out references [Garoufalidis and Bellissard 2007; Sauer 2003], in which the algebraicity of g_M had been proved.

Kassel was partially funded by the Laboratoire International Franco-Québécois de Recherche en Combinatoire (LIRCO) and Université du Québec à Montréal (UQAM). Reutenauer is supported by NSERC (Canada).

References

[André 1989] Y. André, *G-functions and geometry*, Aspects of Mathematics **13**, Friedr. Vieweg & Sohn, Braunschweig, 1989. MR 90k:11087 Zbl 0688.10032

[André 2004] Y. André, "Sur la conjecture des p-courbures de Grothendieck–Katz et un problème de Dwork", pp. 55–112 in *Geometric aspects of Dwork theory*, vol. I, edited by A. Adolphson et al., Walter de Gruyter GmbH & Co. KG, Berlin, 2004. MR 2006d:12005 Zbl 1102.12004

[Berstel 1979] J. Berstel, *Transductions and context-free languages*, Leitfäden der Angewandten Mathematik und Mechanik **38**, B. G. Teubner, Stuttgart, 1979. MR 80j:68056 Zbl 0424.68040

[Berstel and Reutenauer 1990] J. Berstel and C. Reutenauer, "Zeta functions of formal languages", *Trans. Amer. Math. Soc.* **321**:2 (1990), 533–546. MR 91f:68110 Zbl 0797.68092

[Berstel and Reutenauer 2011] J. Berstel and C. Reutenauer, *Noncommutative rational series with applications*, Encyclopedia of Mathematics and its Applications **137**, Cambridge University Press, 2011. MR 2012b:68152 Zbl 1250.68007

[Bost 2001] J.-B. Bost, "Algebraic leaves of algebraic foliations over number fields", *Publ. Math. Inst. Hautes Études Sci.* 93 (2001), 161–221. MR 2002h:14037 Zbl 1034.14010

[Bousquet-Mélou 2005] M. Bousquet-Mélou, "Algebraic generating functions in enumerative combinatorics and context-free languages", pp. 18–35 in *STACS 2005*, edited by V. Diekert and B. Durand, Lecture Notes in Comput. Sci. **3404**, Springer, Berlin, 2005. MR 2006c:05012 Zbl 1118.05300

[Chambert-Loir 2002] A. Chambert-Loir, "Théorèmes d'algébricité en géométrie diophantienne (d'après J.-B. Bost, Y. André, D. & G. Chudnovsky)", pp. 175–209 in *Séminaire Bourbaki 2000/2001*, Astérisque **282**, Société Mathématique de France, Paris, 2002. MR 2004f:11062 Zbl 1044.11055

[Chapoton 2006] F. Chapoton, "Sur le nombre d'intervalles dans les treillis de Tamari", *Sém. Lothar. Combin.* **55** (2006), Art. B55f. MR 2007g:05009 Zbl 1207.05011

[Chomsky and Schützenberger 1963] N. Chomsky and M. P. Schützenberger, "The algebraic theory of context-free languages", pp. 118–161 in *Computer programming and formal systems*, edited by P. Braffort and D. Hirschberg, North-Holland, Amsterdam, 1963. MR 27 #2371 Zbl 0148.00804

[Chudnovsky and Chudnovsky 1985] D. V. Chudnovsky and G. V. Chudnovsky, "Applications of Padé approximations to the Grothendieck conjecture on linear differential equations", pp. 52–100 in *Number theory* (New York, 1983–1984), edited by D. V. Chudnovsky et al., Lecture Notes in Math. **1135**, Springer, Berlin, 1985. MR 87d:11053 Zbl 0565.14010

[Garoufalidis and Bellissard 2007] S. Garoufalidis and J. Bellissard, "Algebraic G-functions associated to matrices over a group-ring", preprint, 2007. arXiv 0708.4234v4

[Jacob 1975] G. Jacob, "Sur un théorème de Shamir", *Information and Control* **27** (1975), 218–261. MR 51 #2361 Zbl 0318.68053

[Katz 1970] N. M. Katz, "Nilpotent connections and the monodromy theorem: applications of a result of Turrittin", *Inst. Hautes Études Sci. Publ. Math.* 39 (1970), 175–232. MR 45 #271 Zbl 0221.14007

[Kontsevich 2011] M. Kontsevich, "Noncommutative identities", notes of talk at *Mathematische Arbeitstagung*, Bonn, 2011. arXiv 1109.2469v1

[OEIS 2010] "The on-line encyclopedia of integer sequences", 2010, http://oeis.org.

[Reutenauer and Robado 2012] C. Reutenauer and M. Robado, "On an algebraicity theorem of Kontsevich", pp. 239–246 in *24th International Conference on Formal Power Series and Algebraic Combinatorics* (Nagoya, 2012), Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012. MR 2958001

[Salomaa and Soittola 1978] A. Salomaa and M. Soittola, *Automata-theoretic aspects of formal power series*, Springer, New York, 1978. MR 58 #3698 Zbl 0377.68039

[Sauer 2003] R. Sauer, "Power series over the group ring of a free group and applications to Novikov–Shubin invariants", pp. 449–468 in *High-dimensional manifold topology*, edited by F. T. Farrell and W. Lück, World Sci. Publ., 2003. MR 2005b:16078 Zbl 1051.16013

[Schützenberger 1962] M. P. Schützenberger, "On a theorem of R. Jungen", *Proc. Amer. Math. Soc.* **13** (1962), 885–890. MR 26 #350 Zbl 0107.03102

[Stanley 1999] R. P. Stanley, *Enumerative combinatorics*, vol. 2, Cambridge Studies in Advanced Mathematics **62**, Cambridge University Press, 1999. MR 2000k:05026 Zbl 0928.05001

[Tutte 1963] W. T. Tutte, "A census of planar maps", *Canad. J. Math.* **15** (1963), 249–271. MR 26 #4343 Zbl 0115.17305

Communicated by Victor Reiner
Received 2013-04-25 Revised 2013-07-15 Accepted 2013-07-24

kassel@math.unistra.fr *Institut de Recherche Mathématique Avancée, CNRS, Université de Strasbourg, 7 rue René Descartes, 67084 Strasbourg, France*

reutenauer.christophe@uqam.ca *Mathématiques, Université du Québec à Montréal, CP 8888 succursale Centre Ville, Montréal QC H3C 3P8, Canada*

mathematical sciences publishers msp

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 8 No. 2 2014

1937-0652(2014)8:2;1-9