Algebra & Number Theory

Volume 19
2025
No. 12

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2025 is US $565/year for the electronic version, and $820/year (+$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004, is published continuously online.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 19:12 (2025)
https://doi.org/10.2140/ant.2025.19.2307

Perfectoid towers and their tilts: with an application to the étale cohomology groups of local log-regular rings

Shinnosuke Ishiro, Kei Nakazato and Kazuma Shimomoto

To initiate a systematic study on the applications of perfectoid methods to Noetherian rings, we introduce the notions of perfectoid towers and their tilts. We mainly show that the tilting operation preserves several homological invariants and finiteness properties. Using this, we also provide a comparison result on étale cohomology groups under the tilting. As an application, we prove finiteness of the prime-to-p-torsion subgroup of the divisor class group of a local log-regular ring that appears in logarithmic geometry in the mixed characteristic case.

MSC2020: primary 14G45; secondary 13B02, 13B40, 13F35, 14A21.
Keywords: perfectoid tower, perfectoid ring, local log-regular ring, Frobenius map, tilting, small tilts, étale cohomology, log-regularity.

1. Introduction

In recent years, the perfectoid technique has become one of the most effective tools in commutative ring theory and singularity theory in mixed characteristic. The *tilting operation* $S \rightsquigarrow S^\flat$ for a perfectoid ring S is a central notion in this method, which makes a bridge between objects in mixed characteristic and objects in positive characteristic. However, perfectoid rings themselves are too big to fit into Noetherian ring theory. Hence, for applications, one often requires distinguished Noetherian ring extensions that approximate perfectoids. Indeed, in many earlier works (such as [7], [8] and [17]), one constructs a highly ramified tower of regular local rings or local log-regular rings:

$$R_0 \subseteq R_1 \subseteq R_2 \subseteq \cdots$$

that converges to a (pre)perfectoid ring. Our purposes in this paper are to axiomatize the above towers and establish a kind of Noetherization of perfectoid theory. As an application, we show a finiteness result on the divisor class groups of local log-regular rings.

Fix a prime p. The highly ramified towers in the positive characteristic case are of the form

$$R \subseteq R^{1/p} \subseteq R^{1/p^2} \subseteq \cdots .$$

This type of tower also appears when one considers the perfect closure of a reduced $\mathbb{F}_p$-algebra. Thus we formulate this class as a tower-theoretic analogue of perfect $\mathbb{F}_p$-algebras, and call them *perfect towers* (Definition 3.2). Next, we introduce *perfectoid towers* as a generalization of perfect towers, which includes the towers applied so far (cf. Proposition 3.58 and Example 3.62). A perfectoid tower is given by a direct system of rings $R_0 \xrightarrow{t_0} R_1 \xrightarrow{t_1} \cdots$ satisfying seven axioms in Definition 3.4 and Definition 3.21. If we assume that each R_i is Noetherian, then these axioms are essential to cope with two main difficulties which we explain below.

The first difficulty is that the residue ring R_i/pR_i on each layer is not necessarily semiperfect. We overcome it by axioms (b), (c), and (d); these ensure the existence of a surjective ring map $F_i : R_{i+1}/pR_{i+1} \to R_i/pR_i$ which gives a decomposition of the Frobenius endomorphism. We call F_i the *i-th Frobenius projection*, and define a ring $R_j^{s.\flat}$ ($j \geq 0$) as the inverse limit of Frobenius projections starting at R_j/pR_j. Then the resulting tower

$$R_0^{s.\flat} \xrightarrow{t_0^{s.\flat}} R_1^{s.\flat} \xrightarrow{t_1^{s.\flat}} \cdots$$

is perfect, and thus we obtain the tilting operation $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0}) \rightsquigarrow (\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$. We remark that this strategy is an axiomatization of the principal arguments in [37].

The second one is that each $R_i^{s.\flat}$ could be imperfect. Because of this, the Witt ring $W(R_i^{s.\flat})$ is often uncontrollable. On the other hand, the definition of Bhatt–Morrow–Scholze's perfectoid rings [5] contains an axiom involving Fontaine's theta map $\theta_S : W(S^\flat) \to S$ (see Definition 3.49(3)), where perfectness of $S^\flat$ is quite effective. Our axioms (f) and (g) are the substitutes for it; these require the Frobenius projections to behave well, especially on the p-torsion parts. This idea is closely related to Gabber and Ramero's characterization of perfectoid rings ([17, Corollary 16.3.75]; see also Theorem 3.50). Indeed, we apply it to deduce that the completed direct limit of a perfectoid tower is a perfectoid ring (Corollary 3.52).

We then verify fundamental properties of the tilting operation for towers. For example, the tilt $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$ is a perfectoid tower with respect to an ideal $I_0^{s.\flat} \subseteq R_0^{s.\flat}$ which is the kernel of the 0-th projection $R_0^{s.\flat} \to R_0/pR_0$ (Proposition 3.41). It induces an isomorphism between two perfectoid objects of different characteristics modulo the defining ideals (Lemma 3.39). Moreover, this operation preserves several finiteness properties such as Noetherianness on each layer (Proposition 3.42). A key to deducing these statements is the following result (see Remark 3.40 for homological interpretation).

Main Theorem 1 (see Theorem 3.35). *$I_0^{s.\flat}$ is a principal ideal. Moreover, we have isomorphisms of (possibly) nonunital rings $(R_i^{s.\flat})_{I_0^{s.\flat}\text{-tor}} \cong (R_i)_{p\text{-tor}}$ $(i \geq 0)$ that are compatible with $\{t_i^{s.\flat}\}_{i\geq 0}$ and $\{t_i\}_{i\geq 0}$.*

Under certain normality assumptions, we obtain a comparison result on the finiteness of étale cohomology groups under tilting for towers (Proposition 4.7). This proposition is considered to rework the crucial part of the proof of [8, Theorem 3.1.3] in a systematic way. Actually, our proposition applies beyond the regular case.

As a typical example, we investigate certain towers of *local log-regular rings*; this class of rings is defined by Kazuya Kato, and is central to logarithmic geometry (readers interested in logarithmic geometry can refer to [17], [26] and [34]). By Kato's structure theorem, a complete local log-regular ring $(R, \mathcal{Q}, \alpha)$ of mixed characteristic is of the form $C(k)[\![\mathcal{Q}\oplus\mathbb{N}^r]\!]/(p-f)$ where $C(k)$ is a Cohen ring of the residue field k of R (see Theorem 2.22). Gabber and Ramero gave a systematic way to build a perfectoid tower (in our sense) over it, which consists of local log-regular rings (Construction 3.56). In this paper, we reveal that its tilt also consists of local log-regular rings, and arises from $C(k)[\![\mathcal{Q}\oplus\mathbb{N}^r]\!]/(p)$ (Theorem 3.61). It says that these two rings on the starting layers fit into a Noetherian variant of the tilting correspondence in perfectoid theory (e.g. $\mathbb{Z}_p$ corresponds to $\mathbb{F}_p[\![x]\!]$).

We regard Theorem 3.61 to be of fundamental importance in the search on the singularities of Noetherian rings via perfectoid methods. For instance, we can investigate the *divisor class groups* of local log-regular rings.[1] The divisor class group of a Noetherian normal domain is an important invariant, but it is often hard to compute.[2] On the other hand, Polstra recently proved a remarkable result stating that the torsion subgroup of the divisor class group of an F-finite strongly F-regular domain is finite [35]. Based on this result, we obtain the following finiteness theorem.

Main Theorem 2 (Theorem 4.13). *Let $(R, \mathcal{Q}, \alpha)$ be a local log-regular ring of mixed characteristic with perfect residue field k of characteristic $p > 0$, and denote by $\mathrm{Cl}(R)$ the divisor class group with its torsion subgroup $\mathrm{Cl}(R)_{\mathrm{tor}}$. Assume that $\widehat{R^{\mathrm{sh}}}[\frac{1}{p}]$ is locally factorial, where $\widehat{R^{\mathrm{sh}}}$ is the completion of the strict Henselization R^{sh}. Then $\mathrm{Cl}(R)_{\mathrm{tor}} \otimes \mathbb{Z}[\frac{1}{p}]$ is a finite group. In other words, the ℓ-primary subgroup of $\mathrm{Cl}(R)_{\mathrm{tor}}$ is finite for all primes $\ell \neq p$ and vanishes for almost all primes $\ell \neq p$.*

[1] K. Kato proved that a local log-regular ring is a normal domain [26].

[2] Every abelian group is realized as a divisor class group of some Dedekind domain (due to Claborn's result [9]).

Our approach to the above theorem is a combination of Theorem 3.61 and Proposition 4.7.

Although we formulated the above theorem only in mixed characteristic, it has an analogue in characteristic $p > 0$, which is relatively easy as follows from the fact that F-finite log-regular rings are strongly F-regular (Lemma 2.25) combined with Polstra's theorem.

For a local log-regular ring $(R, \mathcal{Q}, \alpha)$, Gabber and Ramero constructed the isomorphism $\mathrm{Cl}(\mathcal{Q}) \cong \mathrm{Cl}(R)$ where $\mathrm{Cl}(\mathcal{Q})$ is the divisor class group of the associated monoid [17, Corollary 12.6.43]. It induces the finite generation of $\mathrm{Cl}(R)$.[3]

Recently, H. Cai, S. Lee, L. Ma, K. Schwede, and K. Tucker proved that the torsion part of the divisor class group of a BCM-regular ring is finite (see [6, Theorem 7.0.10.]). Since they also proved that local log-regular rings are BCM-regular, their result recovers Main Theorem 2. Although their approach relies on the evaluation of a certain inequality with the perfectoid signature which is defined in [6] as an analogue of F-signature, it does not use a reduction to positive characteristic and is therefore essentially different from our approach.

Outline. In Section 2, we discuss several properties of monoids and local log-regular rings needed in later sections. We also record a shorter proof of the result that *local log-regular rings are splinter* based on the direct summand theorem in Section 2C.

In Section 3, we introduce the notions of perfect towers, perfectoid towers, and their tilts. The most part of this section is devoted to studying fundamental properties of them; in particular, Section 3D deals with Main Theorem 1. In the last subsection Section 3F, we provide explicit examples of perfectoid towers consisting of local log-regular rings, and compute their tilts.

In Section 4, we give a proof for Main Theorem 2 using the tilting operation, which is an application of Sections 2 and 3.

In the Appendix, we review the notion of *maximal sequences* associated to certain differential modules due to Gabber and Ramero [17]. This plays an important role in the construction of perfectoid towers of local log-regular rings (Construction 3.56).

Conventions. • We consistently fix a prime $p > 0$. If we need to refer to another prime, we denote it by ℓ.

• All rings are assumed to be commutative and unital (unless otherwise stated; cf. Theorem 3.35(2)). We mean by a *ring map* a unital ring homomorphism.

• A local ring is a (not necessarily Noetherian) ring with a unique maximal ideal. When a ring R is local, then we use $\mathfrak{m}_R$ (or simply $\mathfrak{m}$ if no confusion is likely) to denote its unique maximal ideal. We say that a ring map $f : R \to S$ is *local* if R and S are local rings and $f^{-1}(\mathfrak{m}_S) = \mathfrak{m}_R$.

• Unless otherwise stated, a pair (A, I) consisting of a ring A and an ideal $I \subseteq A$ will be simply called a *pair*.

• The Frobenius endomorphism on an $\mathbb{F}_p$-algebra R is denoted by F_R. If there is no confusion, we denote it by Frob.

[3] The first-named author recently provided an elementary proof of [17, Corollary 12.6.43]. See [25].

2. Log-regularity

In this section, we discuss several properties of monoids and local log-regular rings. In Section 2A, we review basic terms on monoids, and examine the behavior of p-times maps which are effectively used in Gabber and Ramero's treatment of perfectoid towers (see Construction 3.56). In Section 2B, we review the definition of local log-regular rings and crucial results by K. Kato, and study the relationship with strong F-regularity. In Section 2C, we recall Gabber and Ramero's result which claims that *any local log-regular ring is a splinter* (Theorem 2.29), and give an alternative proof for it using the direct summand Theorem by Y. André [2] (its derived variant is proved by B. Bhatt [4]).

2A. *Preliminaries on monoids.*

2A1. *Basic terms.* Here we review the definition of several notions on monoids.

Definition 2.1. A *monoid* is a semigroup with a unit. A *homomorphism of monoids* is a semigroup homomorphism between monoids that sends a unit to a unit.

Throughout this paper, all monoids are assumed to be commutative. We denote by **Mnd** the category whose objects are (commutative) monoids and whose morphisms are homomorphisms of monoids.

We denote a unit by 0. Let $\mathcal{Q}$ be a monoid and $\mathcal{Q}^*$ denote the set of all $p \in \mathcal{Q}$ such that there exists $q \in \mathcal{Q}$ such that $p+q=0$. Let $\mathcal{Q}^{gp}$ denote the set of elements $a-b$ for all $a, b \in \mathcal{Q}$, where $a-b=a'-b'$ if and only if there exists $c \in \mathcal{Q}$ such that $a+b'+c=a'+b+c$. By definition, $\mathcal{Q}^{gp}$ is an abelian group. The following conditions yield good classes of monoids.

Definition 2.2. Let $\mathcal{Q}$ be a monoid.

(1) $\mathcal{Q}$ is called *integral* if for x, x' and $y \in \mathcal{Q}$, $x+y=x'+y$ implies $x=x'$.

(2) $\mathcal{Q}$ is called *fine* if it is finitely generated and integral.

(3) $\mathcal{Q}$ is called *sharp* if $\mathcal{Q}^*=0$.

(4) $\mathcal{Q}$ is called *saturated* if the following conditions hold.

  (a) $\mathcal{Q}$ is integral.

  (b) For any $x \in \mathcal{Q}^{gp}$, if $nx \in \mathcal{Q}$ for some $n \geq 1$, then $x \in \mathcal{Q}$.

For an integral monoid $\mathcal{Q}$, the map $\iota_{\mathcal{Q}} : \mathcal{Q} \to \mathcal{Q}^{gp}$; $q \mapsto q-0$ is injective (see [34, Chapter I, Proposition 1.3.3]). In Definition 2.2(4), we identify $\mathcal{Q}$ with its image in $\mathcal{Q}^{gp}$.

Next we recall the definition of a module over a monoid.[4]

Definition 2.3 ($\mathcal{Q}$-module). Let $\mathcal{Q}$ be a monoid.

(1) A *$\mathcal{Q}$-module* is a set M equipped with a binary operation

$$\mathcal{Q} \times M \to M \ ; \ (q, x) \mapsto q+x$$

having the following properties:

[4]This is called a *$\mathcal{Q}$-set in* [34]. We call it as above to follow the convention of the terminology in commutative ring theory.

(a) $0+x=x$ for any $x \in M$;

(b) $(p+q)+x=p+(q+x)$ for any $p,q \in \mathcal{Q}$ and $x \in M$.

(2) A *homomorphism of $\mathcal{Q}$-modules* is a (set-theoretic) map $f : M \to N$ between $\mathcal{Q}$-modules such that $f(q+x)=q+f(x)$ for any $q \in \mathcal{Q}$ and $x \in M$. We denote by $\mathcal{Q}$-**Mod** the category of $\mathcal{Q}$-modules and homomorphisms of $\mathcal{Q}$-modules.

For a monoid $\mathcal{Q}$ and a family of $\mathcal{Q}$-modules $\{M_i\}_{i\in I}$, we denote by $\coprod_{i\in I} M_i$ the disjoint union with the binary operation induced by that of each M_i. Then it is the coproduct in $\mathcal{Q}$-**Mod**.

Definition 2.4 (Monoid algebras). Let R be a ring and let $\mathcal{Q}$ be a monoid. Then the *monoid algebra* $R[\mathcal{Q}]$ is the R-algebra which is the free R-module $R^{\oplus\mathcal{Q}}$, endowed with the unique ring structure induced by the homomorphism of monoids

$$\mathcal{Q} \to R[\mathcal{Q}]\ ;\ q \mapsto e^q.$$

For a monoid $\mathcal{Q}$, one obtains the functor

$$\mathcal{Q}\text{-}\mathbf{Mod} \to R[\mathcal{Q}]\text{-}\mathbf{Mod}\ ;\ M \mapsto R[M], \tag{2-1}$$

which is a left adjoint of the forgetful functor $R[\mathcal{Q}]$-**Mod** $\to$ $\mathcal{Q}$-**Mod**. Notice that (2-1) preserves coproducts (we use this property to prove Proposition 2.8).

Like ideals (resp. prime ideals, the Krull dimension) of a ring, an ideal (resp. prime ideals, the dimension) of a monoid is defined as follows.

Definition 2.5. Let $\mathcal{Q}$ be a monoid.

(1) A $\mathcal{Q}$-submodule of $\mathcal{Q}$ is called an *ideal of* $\mathcal{Q}$.

(2) An ideal I is called *prime* if $I \neq \mathcal{Q}$ and $p+q \in I$ implies $p \in I$ or $q \in I$. Remark that the empty set $\varnothing$ is a prime ideal of $\mathcal{Q}$.

(3) The *dimension* of a monoid $\mathcal{Q}$ is the maximal length d of the ascending chain[5] of prime ideals

$$\varnothing = \mathfrak{q}_0 \subset \mathfrak{q}_1 \subset \cdots \subset \mathfrak{q}_d = \mathcal{Q}^+,$$

where $\mathcal{Q}^+$ is the set of non-unit elements of $\mathcal{Q}$ (i.e. $\mathcal{Q}^+ = \mathcal{Q} \setminus \mathcal{Q}^*$). We also denote it by $\dim \mathcal{Q}$.

Next we review a good class of homomorphisms of monoids, called *exact homomorphisms*.

Definition 2.6 (Exact homomorphisms). Let $\mathcal{P}$ and $\mathcal{Q}$ be monoids.

(1) A homomorphism of monoids $\varphi : \mathcal{P} \to \mathcal{Q}$ is said to be *exact* if the diagram of monoids

$$\begin{tikzcd} \mathcal{P} \arrow[r, "\varphi"] \arrow[d] & \mathcal{Q} \arrow[d] \\ \mathcal{P}^{gp} \arrow[r, "\varphi^{gp}"] & \mathcal{Q}^{gp} \end{tikzcd}$$

is cartesian.

[5] In this paper, the symbol $\subset$ is used to indicate *proper* inclusion for making an analogy to the inequality symbols as in [34].

(2) An *exact submonoid* of $\mathcal{Q}$ is a submonoid $\mathcal{Q}'$ of $\mathcal{Q}$ such that the inclusion map $\mathcal{Q}' \hookrightarrow \mathcal{Q}$ is exact (in other words, $(\mathcal{Q}')^{gp} \cap \mathcal{Q} = \mathcal{Q}'$).

There is a quite useful characterization of exact submonoids (Proposition 2.8). To see this, we recall a graded decomposition of a $\mathcal{Q}$-module attached to a submonoid. For a monoid $\mathcal{Q}$ and a submonoid $\mathcal{Q}' \subseteq \mathcal{Q}$, we denote by $\mathcal{Q} \to \mathcal{Q}/\mathcal{Q}'$ the cokernel of the inclusion map $\mathcal{Q}' \hookrightarrow \mathcal{Q}$.

Definition 2.7. Let $\mathcal{Q}$ be an integral monoid, and let $\mathcal{Q}' \subseteq \mathcal{Q}$ be a submonoid. Then for any $g \in \mathcal{Q}/\mathcal{Q}'$, we denote by $\mathcal{Q}_g$ a $\mathcal{Q}'$-module defined as follows.

- As a set, $\mathcal{Q}_g$ is the inverse image of $g \in \mathcal{Q}/\mathcal{Q}'$ under the cokernel $\mathcal{Q} \to \mathcal{Q}/\mathcal{Q}'$ of $\mathcal{Q}' \hookrightarrow \mathcal{Q}$.
- The operation $\mathcal{Q}' \times \mathcal{Q}_g \to \mathcal{Q}_g$ is defined by the rule $(q, x) \mapsto q + x$ (where $q + x$ denotes the sum of q and x in $\mathcal{Q}$).

By definition, $\mathcal{Q} = \coprod_{g \in \mathcal{Q}/\mathcal{Q}'} \mathcal{Q}_g$ in $\mathcal{Q}'$-**Mod**. Using this, one can refine a characterization of exact embeddings described in [34, Chapter I, Proposition 4.2.7].

Proposition 2.8 (cf. [34, Chapter I, Proposition 4.2.7]). *Let $\mathcal{Q}$ be an integral monoid, and let $\mathcal{Q}' \subseteq \mathcal{Q}$ be a submonoid. Let $\theta : \mathcal{Q}' \hookrightarrow \mathcal{Q}$ be the inclusion map, and let $\mathbb{Z}[\theta] : \mathbb{Z}[\mathcal{Q}'] \to \mathbb{Z}[\mathcal{Q}]$ be the induced ring map. Set $G := \mathcal{Q}/\mathcal{Q}'$.*

(1) *The $\mathbb{Z}[\mathcal{Q}']$-module $\mathbb{Z}[\mathcal{Q}]$ admits a G-graded decomposition $\mathbb{Z}[\mathcal{Q}] = \bigoplus_{g \in G} \mathbb{Z}[\mathcal{Q}_g]$.*

(2) *The following conditions are equivalent.*

  (a) *The inclusion map $\theta : \mathcal{Q}' \hookrightarrow \mathcal{Q}$ is exact. In other words, $(\mathcal{Q}')^{gp} \cap \mathcal{Q} = \mathcal{Q}'$.*
  (b) *$\mathcal{Q}_0 = \mathcal{Q}'$.*
  (c) *$\mathbb{Z}[\theta]$ splits as a $\mathbb{Z}[\mathcal{Q}']$-linear map.*
  (d) *$\mathbb{Z}[\theta]$ is equal to the canonical embedding $\mathbb{Z}[\mathcal{Q}_0] \hookrightarrow \bigoplus_{g \in G} \mathbb{Z}[\mathcal{Q}_g]$.*
  (e) *$\mathbb{Z}[\theta]$ is universally injective.*

Proof. (1) By applying the functor (2-1) (which admits a right adjoint) to the decomposition $\mathcal{Q} = \coprod_{g \in G} \mathcal{Q}_g$, we find that the assertion follows.

(2) Since $\mathcal{Q}_0 = (\mathcal{Q}')^{gp} \cap \mathcal{Q}$ as sets by definition, the equivalence (a)$\Leftrightarrow$(b) follows. The assertion (a)$\Leftrightarrow$(c)$\Leftrightarrow$(e) is none other than [34, Chapter I, Proposition 4.2.7]. Moreover, (d) implies (c) obviously. Thus it suffices to show the implication (b)$\Rightarrow$(d). Assume that (b) is satisfied. Then one can decompose $\mathcal{Q}$ into the direct sum of $\mathcal{Q}'$-modules $\coprod_{g \in G} \mathcal{Q}_g$ with $\mathcal{Q}_0 = \mathcal{Q}'$. Hence the inclusion map $\mathcal{Q}' \hookrightarrow \mathcal{Q}$ is equal to the canonical embedding $\mathcal{Q}_0 \hookrightarrow \coprod_{g \in G} \mathcal{Q}_g$. Thus the induced homomorphism $\mathbb{Z}[\theta] : \mathbb{Z}[\mathcal{Q}_0] \hookrightarrow \mathbb{Z}\big[\coprod_{g \in G} \mathcal{Q}_g\big]$ satisfies (d), as desired. $\square$

Remark 2.9. In the situation of Proposition 2.8, assume that condition (d) is satisfied. Then the split surjection $\pi : \mathbb{Z}[\mathcal{Q}] \to \mathbb{Z}[\mathcal{Q}']$ has the property that $\pi(e^{\mathcal{Q}}) = e^{\mathcal{Q}'}$ by the construction of the G-graded decomposition $\mathbb{Z}[\mathcal{Q}] = \bigoplus_{g \in G} \mathbb{Z}[\mathcal{Q}_g]$. Moreover, $\pi(e^{\mathcal{Q}^+}) \subseteq e^{(\mathcal{Q}')^+}$ because $\mathcal{Q}^+ \cap \mathcal{Q}' \subseteq (\mathcal{Q}')^+$. We use this fact in our proof for Theorem 2.29.

Proposition 2.8 implies the following useful lemma.

Lemma 2.10. *Let $\mathcal{Q}$ be a fine, sharp, and saturated monoid. Let A be a ring. Then there is an embedding of monoids $\mathcal{Q} \hookrightarrow \mathbb{N}^d$ such that the induced map of monoid algebras*

$$A[\mathcal{Q}] \to A[\mathbb{N}^d] \tag{2-2}$$

splits as an $A[\mathcal{Q}]$-linear map.

Proof. Since $\mathcal{Q}$ is saturated, there exists an embedding $\mathcal{Q}$ into some $\mathbb{N}^d$ as an exact submonoid in view of [34, Chapter I, Corollary 2.2.7]. Then by Proposition 2.8, the associated map of monoid algebras

$$\mathbb{Z}[\mathcal{Q}] \to \mathbb{Z}[\mathbb{N}^d] \tag{2-3}$$

splits as a $\mathbb{Z}[\mathcal{Q}]$-linear map. By tensoring (2-3) with A, we get the desired split map. $\square$

2A2. *c-times maps on integral monoids.* For an integral monoid $\mathcal{Q}$, we denote by $\mathcal{Q}_{\mathbb{Q}}$ the submonoid of $\mathcal{Q}^{gp} \otimes_{\mathbb{Z}} \mathbb{Q}$ defined as

$$\mathcal{Q}_{\mathbb{Q}} := \{x \otimes r \in \mathcal{Q}^{gp} \otimes_{\mathbb{Z}} \mathbb{Q} \mid x \in \mathcal{Q},\ r \in \mathbb{Q}_{\geq 0}\}.$$

Using this, one can define the following monoid which plays a central role in Gabber and Ramero's construction of perfectoid towers consisting of local log-regular rings.

Definition 2.11. Let $\mathcal{Q}$ be an integral monoid. Let c and i be non-negative integers with $c > 0$.

(1) We denote by $\mathcal{Q}_c^{(i)}$ the submonoid of $\mathcal{Q}_{\mathbb{Q}}$ defined as

$$\mathcal{Q}_c^{(i)} := \{\gamma \in \mathcal{Q}_{\mathbb{Q}} \mid c^i\gamma \in \mathcal{Q}\}.$$

(2) We denote by $\iota_c^{(i)} : \mathcal{Q}_c^{(i)} \hookrightarrow \mathcal{Q}_c^{(i+1)}$ the inclusion map, and by $\mathbb{Z}[\iota_c^{(i)}] : \mathbb{Z}[\mathcal{Q}_c^{(i)}] \to \mathbb{Z}[\mathcal{Q}_c^{(i+1)}]$ the induced ring map.

In the rest of this subsection, we fix a positive integer $c > 0$. To prove several properties of $\mathcal{Q}_c^{(i)}$, the following one is important as a starting point.

Lemma 2.12. *Let $\mathcal{Q}$ be an integral monoid. Then for every $i \geq 0$, the following assertions hold.*

(1) *$\mathcal{Q}_c^{(i)}$ is integral.*

(2) $\mathcal{Q}_c^{(i+1)} = (\mathcal{Q}_c^{(i)})_c^{(1)}$.

(3) *The c-times map on $\mathcal{Q}_{\mathbb{Q}}$ restricts to an isomorphism of monoids*:

$$f_c : \mathcal{Q}_c^{(i+1)} \xrightarrow{\cong} \mathcal{Q}_c^{(i)}\ ;\ \gamma \mapsto c\gamma.$$

Proof. (1) Since $\mathcal{Q}^{gp} \otimes_{\mathbb{Z}} \mathbb{Q}$ is an integral monoid, so is $\mathcal{Q}_c^{(i)}$.

(2) Since any $g \in (\mathcal{Q}_c^{(i)})^{gp}$ satisfies $c^i g \in \mathcal{Q}^{gp}$, the inclusion map $\mathcal{Q}^{gp} \hookrightarrow (\mathcal{Q}_c^{(i)})^{gp}$ becomes an isomorphism $\varphi : \mathcal{Q}^{gp} \otimes_{\mathbb{Z}} \mathbb{Q} \xrightarrow{\cong} (\mathcal{Q}_c^{(i)})^{gp} \otimes_{\mathbb{Z}} \mathbb{Q}$ by extension of scalars along the flat ring map $\mathbb{Z} \to \mathbb{Q}$. The restriction $\widetilde{\varphi} : \mathcal{Q}_{\mathbb{Q}} \hookrightarrow (\mathcal{Q}_c^{(i)})_{\mathbb{Q}}$ of φ is also an isomorphism, and one can easily check that $\widetilde{\varphi}$ restricts to the desired canonical isomorphism $\mathcal{Q}_c^{(i+1)} \xrightarrow{\cong} (\mathcal{Q}_c^{(i)})_c^{(1)}$.

(3) It is easy to see that the c-times map on $\mathcal{Q}_{\mathbb{Q}}$ restricts to the homomorphism of monoids f_c. Since the abelian group $\mathcal{Q}_{\mathbb{Q}} = \mathcal{Q}^{gp} \otimes_{\mathbb{Z}} \mathbb{Q}$ is torsion-free, f_c is injective. Moreover, any element γ in $\mathcal{Q}_c^{(i)}$ is of the form $x \otimes r$ for some $x \in \mathcal{Q}^{gp}$ and $r \in \mathbb{Q}$, which satisfy $c\big(x \otimes \frac{r}{c}\big) = \gamma$ and $c^{i+1}\big(x \otimes \frac{r}{c}\big) \in \mathcal{Q}$. Hence f_c is also surjective, as desired. □

Let us inspect monoid-theoretic aspects of the inclusion $\iota_c^{(i)} : \mathcal{Q}_c^{(i)} \hookrightarrow \mathcal{Q}_c^{(i+1)}$.

Lemma 2.13. *Let $\mathcal{Q}$ be an integral monoid, and let $\boldsymbol{P} \in \{$fine, sharp, saturated$\}$. If $\mathcal{Q}$ satisfies $\boldsymbol{P}$, then $\mathcal{Q}_c^{(i)}$ also satisfies $\boldsymbol{P}$ for every $i \geq 0$.*

Proof. Assume that $\mathcal{Q}$ is sharp. Pick $x, y \in \mathcal{Q}_c^{(i)}$ such that $x + y = 0$. Then $c^i x = 0$ because $\mathcal{Q}$ is sharp. Since $\mathcal{Q}_c^{(i)}$ is a submonoid of the torsion-free group $\mathcal{Q}^{gp} \otimes_{\mathbb{Z}} \mathbb{Q}$, we have $x = 0$. Next, if $\mathcal{Q}$ is fine or saturated, then it suffices to show the case $i = 1$ by Lemma 2.12(2). If $\mathcal{Q}$ is fine, then there exists a finite system of generators $\{x_1, \dots, x_r\}$ of $\mathcal{Q}$. Hence $\mathcal{Q}_c^{(1)}$ also has a finite system of generators $\big\{x_j \otimes \frac{1}{c}\big\}_{j=1,\dots,r}$. Finally, assume that $\mathcal{Q}_c^{(1)}$ is saturated. Pick an element x of $(\mathcal{Q}_c^{(1)})^{gp}$ such that $nx \in \mathcal{Q}_c^{(1)}$. Then the element cx of $\mathcal{Q}^{gp}$ satisfies $n(cx) = c(nx) \in \mathcal{Q}$. Hence $cx \in \mathcal{Q}$ because $\mathcal{Q}$ is saturated. □

The assumption of fineness on $\mathcal{Q}$ induces several finiteness properties.

Lemma 2.14. *Let $\mathcal{Q}$ be a fine monoid. Then for every $i \geq 0$, the following assertions hold.*

(1) *The ring map $\mathbb{Z}[\iota_c^{(i)}] : \mathbb{Z}[\mathcal{Q}_c^{(i)}] \to \mathbb{Z}[\mathcal{Q}_c^{(i+1)}]$ is module-finite.*

(2) *$\mathcal{Q}_c^{(i+1)}/\mathcal{Q}_c^{(i)} \cong (\mathcal{Q}_c^{(i+1)})^{gp}/(\mathcal{Q}_c^{(i)})^{gp}$ as monoids. Moreover, $\mathcal{Q}_c^{(i+1)}/\mathcal{Q}_c^{(i)}$ forms a finite abelian group.*

(3) *For a prime $p > 0$, we have $\big|\mathcal{Q}_p^{(i+1)}/\mathcal{Q}_p^{(i)}\big| = p^s$ for some $s \geq 0$.*

Proof. In view of Lemma 2.12(2), it suffices to deal with the case when $i = 0$ only. Here notice that $\mathcal{Q}_c^{(0)} = \mathcal{Q}$.

(1) Let $\big\{\frac{1}{c}x_1, \dots, \frac{1}{c}x_r\big\}$ be the system of generators of $\mathcal{Q}_c^{(1)}$ obtained in the proof of Lemma 2.13 where $\frac{1}{c}x_j := x_j \otimes \frac{1}{c}$. Then the $\mathbb{Z}[\mathcal{Q}]$-algebra $\mathbb{Z}[\mathcal{Q}_c^{(1)}]$ is generated by $\{e^{\frac{1}{c}x_1}, \dots, e^{\frac{1}{c}x_r}\}$, and each $e^{\frac{1}{c}x_j} \in \mathbb{Z}[\mathcal{Q}_c^{(1)}]$ is integral over $\mathbb{Z}[\mathcal{Q}]$. Hence $\mathbb{Z}[\iota_c^{(0)}]$ is module-finite, as desired.

(2) By [34, Chapter I, Proposition 1.3.3], $\mathcal{Q}_c^{(1)}/\mathcal{Q}$ is identified with the image of the composition

$$\mathcal{Q}_c^{(1)} \hookrightarrow (\mathcal{Q}_c^{(1)})^{gp} \twoheadrightarrow (\mathcal{Q}_c^{(1)})^{gp}/\mathcal{Q}^{gp}. \tag{2-4}$$

Since $\mathcal{Q}_c^{(1)}$ is generated by $\frac{1}{c}x_1, \dots, \frac{1}{c}x_r$, we see $(\mathcal{Q}_c^{(1)})^{gp}$ is generated by $\frac{1}{c}x_1, \dots, \frac{1}{c}x_r, -\frac{1}{c}x_1, \dots, -\frac{1}{c}x_r$ as a monoid. On the other hand, $-\frac{1}{c}x_j \equiv (c-1)\frac{1}{c}x_j \bmod \mathcal{Q}^{gp}$ for $j = 1, \dots, r$. Hence $(\mathcal{Q}_c^{(1)})^{gp}/\mathcal{Q}^{gp}$ is generated by $\big\{\frac{1}{c}x_j \bmod \mathcal{Q}^{gp}\big\}_{j=1,\dots,r}$ as a monoid. Therefore, the composite map (2-4) is surjective, and $(\mathcal{Q}_c^{(1)})^{gp}/\mathcal{Q}^{gp}$ is a finitely generated torsion abelian group. Thus, $\mathcal{Q}_c^{(1)}/\mathcal{Q}$ coincides with $(\mathcal{Q}_c^{(1)})^{gp}/\mathcal{Q}^{gp}$, which is a finite abelian group, as desired.

(3) Since there exists a surjective group homomorphism

$$f : \underbrace{\mathbb{Z}/p\mathbb{Z} \times \cdots \times \mathbb{Z}/p\mathbb{Z}}_{r} \twoheadrightarrow (\mathcal{Q}_p^{(1)})^{gp}/\mathcal{Q}^{gp}\ ;\ (\overline{n}_1, \dots, \overline{n}_r) \mapsto n_1\Big(\frac{1}{p}x_1\Big) + \cdots + n_r\Big(\frac{1}{p}x_r\Big) \mod \mathcal{Q}^{gp},$$

we have $p^r = \left|(\mathcal{Q}_p^{(1)})^{gp}/\mathcal{Q}^{gp}\right| |\operatorname{Ker}(f)|$. Hence $\left|(\mathcal{Q}_p^{(1)})^{gp}/\mathcal{Q}^{gp}\right| = p^s$ for some $s \geq 0$. Thus the assertion follows from (2). □

By assuming saturatedness, one finds the exactness of $\iota_c^{(i)} : \mathcal{Q}_c^{(i)} \hookrightarrow \mathcal{Q}_c^{(i+1)}$.

Lemma 2.15. *Let $\mathcal{Q}$ be a saturated monoid. Then for every $i \geq 0$, $\iota_c^{(i)} : \mathcal{Q}_c^{(i)} \hookrightarrow \mathcal{Q}_c^{(i+1)}$ is exact (i.e., $\mathcal{Q}_c^{(i+1)} \cap (\mathcal{Q}_c^{(i)})^{gp} = \mathcal{Q}_c^{(i)}$).*

Proof. It suffices to show that $\mathcal{Q}_c^{(i+1)} \cap (\mathcal{Q}_c^{(i)})^{gp} \subseteq \mathcal{Q}_c^{(i)}$. Pick an element $a \in \mathcal{Q}_c^{(i+1)} \cap (\mathcal{Q}_c^{(i)})^{gp}$. Then $ca \in \mathcal{Q}_c^{(i)}$. Since $\mathcal{Q}_c^{(i)}$ is saturated by Lemma 2.13, it implies that $a \in \mathcal{Q}_c^{(i)}$, as desired. □

If further $\mathcal{Q}$ is fine, one can learn more about $\mathbb{Z}[\iota_c^{(i)}] : \mathbb{Z}[\mathcal{Q}_c^{(i)}] \to \mathbb{Z}[\mathcal{Q}_c^{(i+1)}]$ using the exactness of $\iota_c^{(i)}$ assured by Lemma 2.15.

Lemma 2.16. *Let $\mathcal{Q}$ be a fine and saturated monoid. For every $i \geq 0$, set $G_i := \mathcal{Q}_c^{(i+1)}/\mathcal{Q}_c^{(i)}$ (which is a finite abelian group by Lemma 2.14(2)) and $K_i := \operatorname{Frac}(\mathbb{Z}[\mathcal{Q}_c^{(i)}])$.*

(1) *For any $g \in G_i$, we have an isomorphism of $\mathbb{Z}[\mathcal{Q}_c^{(i)}]$-modules $\mathbb{Z}[(\mathcal{Q}_c^{(i+1)})_g] \otimes_{\mathbb{Z}[\mathcal{Q}_c^{(i)}]} K_i \cong K_i$.*

(2) *The base extension $K_i \to \mathbb{Z}[\mathcal{Q}_c^{(i+1)}] \otimes_{\mathbb{Z}[\mathcal{Q}_c^{(i)}]} K_i$ of $\mathbb{Z}[\iota_c^{(i)}]$ is isomorphic to the split injection*
$$K_i \hookrightarrow (K_i)^{\oplus |G_i|} \ ; \ a \mapsto (a, 0, \dots, 0)$$
as a K_i-linear map. In particular, $\dim_{K_i}\big(\mathbb{Z}[\mathcal{Q}_c^{(i+1)}] \otimes_{\mathbb{Z}[\mathcal{Q}_c^{(i)}]} K_i\big) = |\mathcal{Q}_c^{(i+1)}/\mathcal{Q}_c^{(i)}|$.

Proof. In view of Lemma 2.12(2) and Lemma 2.13, it suffices to show the assertions only for the case when $i = 0$.

(1) Let $y_g \in \mathcal{Q}_c^{(1)}$ be an element whose image in $\mathcal{Q}_c^{(1)}/\mathcal{Q}$ is equal to g. Then we obtain an injective homomorphism of $\mathcal{Q}$-modules
$$\iota_g : \mathcal{Q} \hookrightarrow (\mathcal{Q}_c^{(1)})_g \ ; \ x \mapsto x + y_g, \tag{2-5}$$
which induces an injective $\mathbb{Z}[\mathcal{Q}]$-linear map $\mathbb{Z}[\iota_g] : \mathbb{Z}[\mathcal{Q}] \hookrightarrow \mathbb{Z}[(\mathcal{Q}_c^{(1)})_g]$. Thus it suffices to show that $\operatorname{Coker}(\mathbb{Z}[\iota_g]) \otimes_{\mathbb{Z}[\mathcal{Q}]} K_0 = (0)$, i.e., $\operatorname{Coker}(\mathbb{Z}[\iota_g])$ is a torsion $\mathbb{Z}[\mathcal{Q}]$-module. On the other hand, we also have a homomorphism of $\mathcal{Q}$-modules
$$(\mathcal{Q}_c^{(1)})_g \to \mathcal{Q}^{gp} \ ; \ y \mapsto y - y_g,$$
which induces an embedding of $\mathbb{Z}[\mathcal{Q}]$-modules $\operatorname{Coker}(\mathbb{Z}[\iota_g]) \hookrightarrow \mathbb{Z}[\mathcal{Q}^{gp}]/\mathbb{Z}[\mathcal{Q}]$. Since $\mathbb{Z}[\mathcal{Q}^{gp}]/\mathbb{Z}[\mathcal{Q}]$ is $\mathbb{Z}[\mathcal{Q}]$-torsion, the assertion follows.

(2) This follows from the combination of part (1) with Lemma 2.15 and Proposition 2.8(2). □

2B. *Local log-regular rings.*

2B1. *Definition of local log-regular rings.* We review the definition and fundamental properties of local log-regular rings. Unless otherwise stated, we always assume that the monoid structure of a commutative ring is specified by the multiplicative structure.

Definition 2.17 [34, Chapter III, Definition 1.2.3]. Let R be a ring and let $\mathcal{Q}$ be a monoid with a homomorphism $\alpha : \mathcal{Q} \to R$ of monoids. Then we say that the triple $(R, \mathcal{Q}, \alpha)$ is a *log ring*. Moreover, we say that $(R, \mathcal{Q}, \alpha)$ is a *local log ring* if $(R, \mathcal{Q}, \alpha)$ is a log ring, where R is a local ring and $\alpha^{-1}(R^{\times}) = \mathcal{Q}^*$.

In order to preserve the locality of a log structure, we need the locality of a ring map.

Lemma 2.18. *Let $(R, \mathcal{Q}, \alpha)$ be a local log ring and let $(S, \mathfrak{m}_S)$ be a local ring with a local ring map $\phi : R \to S$. Then $(S, \mathcal{Q}, \phi \circ \alpha)$ is also a local log ring.*

Proof. By the locality of ϕ, we obtain the equality $(\phi \circ \alpha)^{-1}(S^{\times}) = \mathcal{Q}^*$, as desired. □

Now we define *log-regular rings* according to [34].

Definition 2.19. Let $(R, \mathcal{Q}, \alpha)$ be a local log ring, where R is Noetherian and $\overline{\mathcal{Q}} := \mathcal{Q}/\mathcal{Q}^*$ is fine and saturated. Let I_α be the ideal of R generated by the set $\alpha(\mathcal{Q}^+)$. Then $(R, \mathcal{Q}, \alpha)$ is called a *log-regular ring* if the following conditions hold.

(1) R/I_α is a regular local ring.

(2) $\dim R = \dim R/I_\alpha + \dim \mathcal{Q}$.

Remark 2.20. For a monoid $\mathcal{Q}$ such that $\overline{\mathcal{Q}}$ is fine and saturated, the natural projection $\pi : \mathcal{Q} \twoheadrightarrow \overline{\mathcal{Q}}$ splits (see [17, Lemma 6.2.10]). Thus, in the situation of Definition 2.19, α extends to the homomorphism of monoids $\overline{\alpha} : \overline{\mathcal{Q}} \to R$ along π. Namely, we obtain another local log-regular ring $(R, \overline{\mathcal{Q}}, \overline{\alpha})$ with the same underlying ring, where $\overline{\mathcal{Q}}$ is fine, sharp, and saturated.

In his monumental paper [26], Kato considered log structures of schemes on the étale sites, and he then considered them on the Zariski sites [27]. However, we do not need any deep part of logarithmic geometry and the present paper focuses on the local study of schemes with log structures. We should remark that if k is any fixed field and $\mathcal{Q} \subseteq \mathbb{N}^d$ is a fine and saturated monoid, then the monoid algebra $k[\mathcal{Q}]$ is known as an *affine normal semigroup ring* which is actively studied in combinatorial commutative algebra (see the book [30]). The following theorem is a natural extension of the Cohen–Macaulay property for the classical toric singularities over a field proved by Hochster [22].

Theorem 2.21 [27, Theorem 4.1]. *Every local log-regular ring is Cohen–Macaulay and normal.*

Let R be a ring and let $\mathcal{Q}$ be a fine sharp monoid. We denote by $R[\mathcal{Q}^+]$ the ideal of $R[\mathcal{Q}]$ generated by elements $\sum_{q \in \mathcal{Q}^+} a_q e^q$, where a_q is an element of R. Then we denote by $R[\![\mathcal{Q}]\!]$ the adic completion of $R[\mathcal{Q}]$ with respect to the ideal $R[\mathcal{Q}^+]$.

As to the structure of complete local log-regular rings, we have the following result analogous to the classical Cohen's structure theorem, originally proved in [27]. We borrow the presentation from [34, Chapter III, Theorem 1.11.2].

Theorem 2.22 (Kato). *Let $(R, \mathcal{Q}, \alpha)$ be a local log ring such that R is Noetherian and $\mathcal{Q}$ is fine, sharp, and saturated. Let k be the residue field of R and $\mathfrak{m}_R$ its maximal ideal. Let r be the dimension of R/I_α.*

(1) *Suppose that R contains a field. Then $(R, \mathcal{Q}, \alpha)$ is log-regular if and only if there exists a commutative diagram*

$$\begin{tikzcd} \mathcal{Q} \arrow[r] \arrow[d, "\alpha"] & k[[\mathcal{Q} \oplus \mathbb{N}^r]] \arrow[d, "\psi"] \\ R \arrow[r] & \hat{R} \end{tikzcd}$$

where $\hat{R}$ is the completion along the maximal ideal and ψ is an isomorphism of rings.

(2) *Assume that R is of mixed characteristic $p > 0$. Let $C(k)$ be a Cohen ring of k. Then $(R, \mathcal{Q}, \alpha)$ is log-regular if and only if there exists a commutative diagram*

$$\begin{tikzcd} \mathcal{Q} \arrow[r] \arrow[d, "\alpha"] & C(k)[[\mathcal{Q} \oplus \mathbb{N}^r]] \arrow[d, "\psi"] \\ R \arrow[r] & \hat{R} \end{tikzcd}$$

where $\hat{R}$ is the completion along the maximal ideal and ψ is a surjective ring map with $\operatorname{Ker}(\psi) = (\theta)$ for some element $\theta \in \mathfrak{m}_{\hat{R}}$ whose constant term is p. Moreover, $\operatorname{Ker}(\psi) = (\theta')$ for any element $\theta' \in \operatorname{Ker}(\psi)$ whose constant term is p.

Proof. Assertion (1) and the first part of (2) are [34, Chapter III, Theorem 1.11.2]. Pick an element $\theta' \in \operatorname{Ker}(\psi)$ whose constant term is p. Note that θ' is a regular element that is not invertible. By [34, Chapter III, Proposition 1.10.13], $C(k)[[\mathcal{Q} \oplus \mathbb{N}^r]]/(\theta')$ is a domain of $\dim \mathcal{Q} + r = \dim R = \dim \hat{R}$. Thus $\operatorname{Ker}(\psi) = (\theta')$ holds.[6] □

The completion of a normal affine semigroup ring with respect to the ideal generated by elements of the semigroup is a typical example of local log-regular rings:

Lemma 2.23. *Let $\mathcal{Q}$ be a fine, sharp and saturated monoid and let k be a field. Then $(k[[\mathcal{Q}]], \mathcal{Q}, \iota)$ is a local log-regular ring, where $\iota : \mathcal{Q} \hookrightarrow k[[\mathcal{Q}]]$ is the natural injection.*

Proof. By [34, Chapter I, Proposition 3.6.1], $(k[[\mathcal{Q}]], \mathcal{Q}, \iota)$ is a local log ring. Now applying Theorem 2.22, it is a local log-regular ring. □

2B2. *Log-regularity and strong F-regularity.* Strongly F-regular rings are one of the important classes appearing in the study of F-singularities. Let us recall the definition.

Definition 2.24 (strong F-regularity). Let R be a Noetherian reduced $\mathbb{F}_p$-algebra that is F-finite. Let $F^e_* R$ be the same as R as its underlying abelian groups with its R-module structure via restriction of scalars via the e-th iterated Frobenius endomorphism F^e_R on R. Then we say that R is *strongly F-regular*, if for any element $c \in R$ that is not in any minimal prime of R, there exist an $e > 0$ and a map $\phi \in \operatorname{Hom}_R(F^e_* R, R)$ such that $\phi(F^e_* c) = 1$.

[6] This argument is due to Ogus. See the proof of [34, Chapter III, Theorem 1.11.2(2)].

It is known that strongly F-regular rings are Cohen–Macaulay and normal (for example, see [28, Proposition 4.4 and Theorem 4.6]). Let us show that log-regularity implies strong F-regularity (in positive characteristic cases).

Lemma 2.25. *Let $(R, \mathcal{Q}, \alpha)$ be a local log-regular ring of characteristic $p > 0$ such that R is F-finite. Then R is strongly F-regular.*

Proof. The completion of R with respect to its maximal ideal is isomorphic to the completion of $k[\mathcal{Q} \oplus \mathbb{N}^r]$, and $\mathcal{Q}$ is a fine, sharp and saturated monoid by Theorem 2.22 and [34, Chapter I, Proposition 3.4.1]. Then it follows from Lemma 2.10 that $\mathcal{Q} \oplus \mathbb{N}^r$ can be embedded into $\mathbb{N}^d$ for $d > 0$, and $k[\mathcal{Q} \oplus \mathbb{N}^r] \to k[\mathbb{N}^d] \cong k[x_1, \dots, x_d]$ splits as a $k[\mathcal{Q} \oplus \mathbb{N}^r]$-linear map. Applying [23, Theorem 3.1], we see that $k[\mathcal{Q} \oplus \mathbb{N}^r]$ is strongly F-regular. After completion, the complete local ring $k[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]$ is strongly F-regular in view of [1, Theorem 3.6]. Then by faithful flatness of $R \to k[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]$, [23, Theorem 3.1] applies to yield strong F-regularity of R. □

Under the hypothesis in the following proposition, one can easily establish the finiteness of the torsion part of the divisor class group, which is the first assertion of Theorem 4.13.

Proposition 2.26. *Assume that $R \cong C(k)[\![\mathcal{Q}]\!]$, where $C(k)$ is a Cohen ring with F-finite residue field k and $\mathcal{Q}$ is a fine, sharp, and saturated monoid. Let $\mathrm{Cl}(R)_{\mathrm{tor}}$ be the torsion subgroup of $\mathrm{Cl}(R)$. Then $\mathrm{Cl}(R)_{\mathrm{tor}} \otimes \mathbb{Z}_{(\ell)}$ is finite for all $\ell \neq p$, and vanishes for almost all $\ell \neq p$.*

Proof. Since $R \cong C(k)[\![\mathcal{Q}]\!]$, we have

$$R/pR \cong k[\![\mathcal{Q}]\!],$$

which is a local F-finite log-regular ring. There is an induced map $\mathrm{Cl}(R) \to \mathrm{Cl}(R/pR)$. By restriction, we have $\mathrm{Cl}(R)_{\mathrm{tor}} \to \mathrm{Cl}(R/pR)_{\mathrm{tor}}$. Then Lemma 2.25 together with Polstra's result [35] says that $\mathrm{Cl}(R/pR)_{\mathrm{tor}}$ is finite. Let C_ℓ be the maximal ℓ-subgroup of $\mathrm{Cl}(R)_{\mathrm{tor}}$. Since $\ell \neq p$, we find that the map $\mathrm{Cl}(R)_{\mathrm{tor}} \to \mathrm{Cl}(R/pR)_{\mathrm{tor}}$ restricted to C_ℓ is injective in view of [18, Theorem 1.2]. In conclusion, C_ℓ is finite for all $\ell \neq p$, and C_ℓ vanishes for almost all $\ell \neq p$, as desired. □

2C. ***Log-regularity and splinters.*** Local log-regular rings have another notable property; they are *splinters*. Let us recall the definition of splinters.

Definition 2.27. A Noetherian ring A is a *splinter* if every finite ring map $f : A \to B$ such that $\mathrm{Spec}(B) \to \mathrm{Spec}(A)$ is surjective admits an A-linear map $h : B \to A$ such that $h \circ f = \mathrm{id}_A$.

In general, it is not easy to see which algebraic operations preserve splinters. In fact, it remains unsolved whether polynomial rings over a splinter are splinters (see [10, Question 1']). Regarding these issues, Datta and Tucker proved remarkable results such as [10, Theorem B], [10, Theorem C], and [10, Example 3.2.1]. See also Murayama's work [32] for the study of purity of ring extensions.

In order to prove the splinter property, we need a lemma on splitting a map under completion.

Lemma 2.28. *Let R be a ring and let $f : M \to N$ be an R-linear map. Consider a decreasing filtration by R-submodules $\{M_\lambda\}_{\lambda\in\Lambda}$ of M and a decreasing filtration by R-submodules $\{N_\lambda\}_{\lambda\in\Lambda}$ of N such that $f(M_\lambda) \subseteq N_\lambda$ for each $\lambda \in \Lambda$. Set*

$$\widehat{M} := \varprojlim_{\lambda\in\Lambda} M/M_\lambda \quad \textit{and} \quad \hat{N} := \varprojlim_{\lambda\in\Lambda} N/N_\lambda.$$

Finally, assume that f is a split injection that admits an R-linear map $g : N \to M$ such that $g \circ f = \mathrm{id}_M$, $g(N_\lambda) \subseteq M_\lambda$ for each $\lambda \in \Lambda$. Then f extends to a split injection $\widehat{M} \to \hat{N}$.

Proof. By assumption, there is an induced map

$$M/M_\lambda \xrightarrow{\bar{f}} N/N_\lambda \xrightarrow{\bar{g}} M/M_\lambda$$

which is an identity on M/M_λ. Taking inverse limits, we get an identity map $\widehat{M} \to \hat{N} \to \widehat{M}$, which proves the lemma. □

The next result is originally due to Gabber and Ramero [17, Theorem 17.3.12].[7] We give an alternative and short proof, using the direct summand theorem by André [2].

Theorem 2.29. *A local log-regular ring $(R, \mathcal{Q}, \alpha)$ is a splinter.*

Proof. First, we prove the theorem when R is complete. By Remark 2.20, we may assume that $\mathcal{Q}$ is fine, sharp, and saturated. By Theorem 2.22, we have

$$R \cong k[\![\mathcal{Q} \oplus \mathbb{N}^r]\!], \text{ or } R \cong C(k)[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]/(p - f),$$

depending on whether R contains a field or not. Let us consider the mixed characteristic case. By Lemma 2.10, there is a split injection $C(k)[\mathcal{Q} \oplus \mathbb{N}^r] \to C(k)[\mathbb{N}^d]$ for some $d > 0$, which comes from an injection $\delta : \mathcal{Q} \oplus \mathbb{N}^r \to \mathbb{N}^d$ that realizes $\delta(\mathcal{Q} \oplus \mathbb{N}^r)$ as an exact submonoid of $\mathbb{N}^d$. After dividing out by the ideal $(p - f)$, we find that the map

$$C(k)[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]/(p - f) \to C(k)[\![\mathbb{N}^d]\!]/(p - f)$$

splits as a $C(k)[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]/(p - f)$-linear map by Remark 2.9 and Lemma 2.28. Hence, R becomes a direct summand of the complete regular local ring $A := C(k)[\![x_1, \ldots, x_d]\!]/(p - f)$. By invoking [10, Proposition 2.2.8] and the Direct Summand Theorem [2], we see that R is a splinter. The case where $R = k[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]$ can be treated similarly.

Next let us consider the general case. Then the completion map $R \to \hat{R}$ is faithfully flat and $\hat{R}$ is a complete local log-regular ring (see Theorem 2.22). Hence applying the complete case as above and [10, Proposition 2.2.8] shows that R is a splinter, as desired. □

[7] One notices that the treatment of logarithmic geometry in [17] is topos-theoretic, while [27] considers mostly the Zariski sites.

3. Perfectoid towers and small tilts

In this section, we establish a tower-theoretic framework to deal with perfectoid objects using the notion of *perfectoid towers*. We first introduce the class of *perfect towers* (Definition 3.2) in Section 3A, and then define *inverse perfection of towers* (Definition 3.8) in Section 3B. These notions are tower-theoretic variants of perfect $\mathbb{F}_p$-algebras and inverse perfection of rings, respectively. In Section 3C, we give a set of axioms for perfectoid towers. In Section 3D, we adopt the process of inverse perfection for perfectoid towers as a new tilting operation. Indeed, we verify the invariance of several good properties under the tilting; Main Theorem 1 is discussed here. In Section 3E, we describe the relationship between perfectoid towers and perfectoid rings. This subsection also includes an alternative characterization of perfectoid rings without $\mathbb{A}_{\mathrm{inf}}$. In Section 3F, we calculate the tilts of perfectoid towers consisting of local log-regular rings.

3A. *Perfect towers.* First of all, we consider the category of *towers of rings*.

Definition 3.1 (towers of rings).

(1) A *tower of rings* is a direct system of rings of the form

$$R_0 \xrightarrow{t_0} R_1 \xrightarrow{t_1} R_2 \xrightarrow{t_2} \cdots \xrightarrow{t_{i-1}} R_i \xrightarrow{t_i} \cdots,$$

and we denote it by $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ or $\{R_0 \xrightarrow{t_0} R_1 \xrightarrow{t_1} \cdots\}$.

(2) A morphism of towers of rings $f : (\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0}) \to (\{R'_i\}_{i\geq 0}, \{t'_i\}_{i\geq 0})$ is defined as a collection of ring maps $\{f_i : R_i \to R'_i\}_{i\geq 0}$ that is compatible with the transition maps; in other words, f represents the commutative diagram

$$\begin{tikzcd} R_0 \arrow[r] \arrow[d, "f_0"'] & R_1 \arrow[r] \arrow[d, "f_1"'] & R_2 \arrow[r] \arrow[d, "f_2"'] & \cdots \arrow[r] & R_i \arrow[r] \arrow[d, "f_i"'] & \cdots \\ R'_0 \arrow[r] & R'_1 \arrow[r] & R'_2 \arrow[r] & \cdots \arrow[r] & R'_i \arrow[r] & \cdots. \end{tikzcd}$$

For a tower of rings $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$, we often denote by R_∞ an inductive limit $\varinjlim_{i\geq 0} R_i$. Clearly, an isomorphism of towers of rings $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0}) \to (\{R'_i\}_{i\geq 0}, \{t'_i\}_{i\geq 0})$ induces the isomorphism of rings $R_\infty \xrightarrow{\cong} R'_\infty$. For every $i \geq 0$, we regard R_{i+1} as an R_i-algebra via the transition map t_i.

Recall that the direct perfection of an $\mathbb{F}_p$-algebra R, which we denote by R^{perf}, is the direct limit of the tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ where $R_i = R$ and $t_i = F_R$ for every $i \geq 0$. We denote by $\phi_R : R \to R^{\mathrm{perf}}$ the natural map $R_0 \to \varinjlim_{i\geq 0} R_i$. If R is reduced, this tower can be regarded as ring extensions obtained by adjoining p^i-th roots (cf. Example 3.3). We formulate such towers as follows, and call them *perfect towers*.

Definition 3.2 (perfect towers). A *perfect $\mathbb{F}_p$-tower* (or, *perfect tower* as an abbreviated form) is a tower of rings that is isomorphic to a tower of the following form, where R is a reduced $\mathbb{F}_p$-algebra:

$$R \xrightarrow{F_R} R \xrightarrow{F_R} R \xrightarrow{F_R} \cdots. \tag{3-1}$$

Example 3.3. Let R be a reduced $\mathbb{F}_p$-algebra. Let R^{1/p^i} be the ring of p^i-th roots of elements of R for every $i \geq 0$.[8] Then the tower $R \xrightarrow{t_0} R^{1/p} \xrightarrow{t_1} R^{1/p^2} \xrightarrow{t_2} \cdots$ is a perfect tower. Indeed, we have an isomorphism $F_i : R^{1/p^{i+1}} \to R^{1/p^i}$; $x \mapsto x^p$. By putting $F_{0,i+1} := F_0 \circ \cdots \circ F_i$, we obtain the desired commutative ladder:

$$\begin{tikzcd}
R^{1/p^0} \arrow[r, "t_0"] \arrow[d, "F_{0,0}"] & R^{1/p} \arrow[r, "t_1"] \arrow[d, "F_{0,1}"] & \cdots \arrow[r, "t_{i-1}"] & R^{1/p^i} \arrow[r, "t_i"] \arrow[d, "F_{0,i}"] & \cdots \\
R \arrow[r, "F_R"'] & R \arrow[r, "F_R"'] & \cdots \arrow[r, "F_R"'] & R \arrow[r, "F_R"'] & \cdots .
\end{tikzcd}$$

3B. *Purely inseparable towers and inverse perfection.* In this subsection, we define *inverse perfection for towers*, which assigns a perfect tower to a tower by arranging a certain type of inverse limits of rings. For this, we introduce the following class of towers that admit distinguished inverse systems of rings.

Definition 3.4 (purely inseparable towers). Let R be a ring, and let $I \subseteq R$ be an ideal.

(1) A tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is called a *p-purely inseparable tower arising from* (R, I) if it satisfies the following axioms.

  (a) $R_0 = R$ and $p \in I$.

  (b) For any $i \geq 0$, the ring map $\bar{t}_i : R_i/IR_i \to R_{i+1}/IR_{i+1}$ induced by t_i is injective.

  (c) For any $i \geq 0$, the image of the Frobenius endomorphism on R_{i+1}/IR_{i+1} is contained in the image of $\bar{t}_i : R_i/IR_i \to R_{i+1}/IR_{i+1}$.

(2) Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a p-purely inseparable tower arising from (R, I). For any $i \geq 0$, we denote by $F_i : R_{i+1}/IR_{i+1} \to R_i/IR_i$ the ring map (which uniquely exists by axioms (b) and (c)) such that the following diagram commutes:

$$\begin{tikzcd}
R_{i+1}/IR_{i+1} \arrow[r, "F_{R_{i+1}/IR_{i+1}}"] \arrow[dr, "F_i"'] & R_{i+1}/IR_{i+1} \\
& R_i/IR_i . \arrow[u, "\bar{t}_i"']
\end{tikzcd} \tag{3-2}$$

We call F_i *the i-th Frobenius projection* (of $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ associated to (R, I)).

Hereafter, we leave out "p-" from "p-purely inseparable towers" if no confusion occurs. Similarly, we omit the parenthetic phrase "of … associated to (R, I)" subsequent to "the i-th Frobenius projection" (but we should be careful in some situations; cf. Remark 3.38).

Throughout this paper, when a purely inseparable tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is given and its starting layer (R, I) is clear from the context, we denote R_i/IR_i by $\bar{R}_i$ for every $i \geq 0$.

Example 3.5. Any perfect tower is a purely inseparable tower. More precisely, $(\{R\}_{i\geq 0}, \{F_R\}_{i\geq 0})$ appearing in Definition 3.2 is a purely inseparable tower arising from $(R, (0))$. Indeed, axioms (a) and (c)

[8] For more details of the ring of p-th roots of elements of a reduced ring, we refer to [28].

are obvious, and axiom (b) follows from reducedness of R. The i-th Frobenius projection is given by the identity map on R.

To develop the theory of perfectoid towers, we often use a combination of diagram (3-2) in Definition 3.4 and diagram (3-3) in the following lemma.

Lemma 3.6. *Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a purely inseparable tower arising from some pair (R, I). Then for every $i \geq 0$, the following assertions hold.*

(1) $\mathrm{Ker}(F_i) = \mathrm{Ker}(F_{\overline{R}_{i+1}})$. *In particular, F_i is injective if and only if $\overline{R}_{i+1}$ is reduced.*

(2) *Any element of $\overline{R}_{i+1}$ is a root of a polynomial of the form $X^p - \overline{t}_i(a)$ with $a \in \overline{R}_i$. In particular, the ring map $\overline{t}_i : \overline{R}_i \hookrightarrow \overline{R}_{i+1}$ is integral.*

(3) *The following diagram commutes:*

$$\begin{tikzcd} \overline{R}_{i+1} \arrow[dr, "F_i"] & \\ \overline{R}_i \arrow[u, "\overline{t}_i"] \arrow[r, "F_{\overline{R}_i}"'] & \overline{R}_i . \end{tikzcd} \tag{3-3}$$

Proof. Since $\overline{t}_i$ is injective, the commutative diagram (3-2) yields assertion (1). Moreover, (3-2) also yields the equality $x^p - \overline{t}_i(F_i(x)) = 0$ for every $x \in \overline{R}_{i+1}$. Hence assertion (2) follows. To prove (3), let us recall the equalities

$$\overline{t}_i \circ F_{\overline{R}_i} = F_{\overline{R}_{i+1}} \circ \overline{t}_i = \overline{t}_i \circ F_i \circ \overline{t}_i,$$

where the second one follows from the commutative diagram (3-2). Since $\overline{t}_i$ is injective, we obtain the equality $F_{\overline{R}_i} = F_i \circ \overline{t}_i$, as desired. □

Lemma 3.6(3) is essential for defining inverse perfection of towers (cf. Definition 3.8(2)). Moreover, it provides a useful tool for studying direct perfection on each layer. Recall that for an $\mathbb{F}_p$-algebra homomorphism $f : R \to S$, there exists a unique ring map $f^{\mathrm{perf}} : R^{\mathrm{perf}} \to S^{\mathrm{perf}}$ such that the following diagram commutes (the notations are explained just before Definition 3.2):

$$\begin{tikzcd} R \arrow[r, "f"] \arrow[d, "\phi_R"'] & S \arrow[d, "\phi_S"] \\ R^{\mathrm{perf}} \arrow[r, "f^{\mathrm{perf}}"] & S^{\mathrm{perf}}. \end{tikzcd}$$

Corollary 3.7. *Keep the notation as in Lemma 3.6. Then $(\overline{t}_i)^{\mathrm{perf}} : (\overline{R}_i)^{\mathrm{perf}} \to (\overline{R}_{i+1})^{\mathrm{perf}}$ is an isomorphism of rings whose inverse map is $(F_i)^{\mathrm{perf}} : (\overline{R}_{i+1})^{\mathrm{perf}} \to (\overline{R}_i)^{\mathrm{perf}}$ up to the Frobenius automorphisms.*

Proof. By Lemma 3.6(3), $F_{(\overline{R}_{i+1})^{\mathrm{perf}}}$ is described as $(F_{\overline{R}_{i+1}})^{\mathrm{perf}} = (\overline{t}_i)^{\mathrm{perf}} \circ F_i^{\mathrm{perf}}$, and it is an automorphism. Similarly, it follows from the commutative diagram (3-2) that $F_i^{\mathrm{perf}} \circ (\overline{t}_i)^{\mathrm{perf}}$ is the Frobenius automorphism of $(\overline{R}_i)^{\mathrm{perf}}$. Hence the assertion follows. □

Definition 3.8 (inverse perfection of towers). Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a (p-)purely inseparable tower arising from some pair (R, I).

(1) For any $j \geq 0$, we define the *j-th inverse quasi-perfection of* $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ *associated to* (R, I) as the limit

$$(R_j)_I^{q.\,\mathrm{frep}} := \varprojlim\{\cdots \to \overline{R}_{j+i+1} \xrightarrow{F_{j+i}} \overline{R}_{j+i} \to \cdots \xrightarrow{F_j} \overline{R}_j\}.$$

(2) For any $j \geq 0$, we define an injective ring map $(t_j)_I^{q.\,\mathrm{frep}} : (R_j)_I^{q.\,\mathrm{frep}} \hookrightarrow (R_{j+1})_I^{q.\,\mathrm{frep}}$ by the rule

$$(t_j)_I^{q.\,\mathrm{frep}}((a_i)_{i\geq 0}) := (\overline{t}_{j+i}(a_i))_{i\geq 0}.$$

We call the resulting tower $\big(\{(R_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0}, \{(t_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0}\big)$ *the inverse perfection of* $\big(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0}\big)$ *associated to* (R, I).

(3) For any $j \geq 0$, we define a ring map $(F_j)_I^{q.\,\mathrm{frep}} : (R_{j+1})_I^{q.\,\mathrm{frep}} \to (R_j)_I^{q.\,\mathrm{frep}}$ by the rule

$$(F_j)_I^{q.\,\mathrm{frep}}((a_i)_{i\geq 0}) := (F_{j+i}(a_i))_{i\geq 0}. \tag{3-4}$$

(4) For any $j \geq 0$ and for any $m \geq 0$, we denote by $\Phi_m^{(j)}$ the m-th projection map:

$$(R_j)_I^{q.\,\mathrm{frep}} \to \overline{R}_{j+m}\ ;\ (a_i)_{i\geq 0} \mapsto a_m.$$

If no confusion occurs, we also abbreviate $(R_j)_I^{q.\,\mathrm{frep}}$, $(t_j)_I^{q.\,\mathrm{frep}}$, $(F_j)_I^{q.\,\mathrm{frep}}$ to $R_j^{q.\,\mathrm{frep}}$, $t_j^{q.\,\mathrm{frep}}$, $F_j^{q.\,\mathrm{frep}}$.

Example 3.9. Let R be an $\mathbb{F}_p$-algebra. Set $R_i := R$ and $t_i := \mathrm{id}_R$ for every $i \geq 0$. Then the tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a purely inseparable tower arising from $(R, (0))$. Moreover, for every $j \geq 0$, the attached j-th inverse quasi-perfection is a limit

$$R_j^{q.\,\mathrm{frep}} = \varprojlim\{\cdots \xrightarrow{F_R} R \xrightarrow{F_R} R \xrightarrow{F_R} R\},$$

which is none other than the inverse perfection of R.

In the situation of Definition 3.8, we have the commutative diagram:

$$\begin{tikzcd}
(R_{j+1})_I^{q.\,\mathrm{frep}} \arrow[r, "F_{(R_{j+1})_I^{q.\,\mathrm{frep}}}"] \arrow[dr, "(F_j)_I^{q.\,\mathrm{frep}}"'] & (R_{j+1})_I^{q.\,\mathrm{frep}} \\
& (R_j)_I^{q.\,\mathrm{frep}}. \arrow[u, "(t_j)_I^{q.\,\mathrm{frep}}"']
\end{tikzcd} \tag{3-5}$$

Therefore the tower $\big(\{(R_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0}, \{(t_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0}\big)$ is also a purely inseparable tower associated to $((R_0)_I^{q.\,\mathrm{frep}}, (0))$.

In the rest of this subsection, we fix a purely inseparable tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ arising from some pair (R, I). Keep in mind that the inverse perfection $(\{(R_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0}, \{(t_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0})$ is given in Definition 3.8(2), and its Frobenius projections $\{(F_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0}$ are described in Definition 3.8(3). Some basic properties of inverse quasi-perfection are contained in the following proposition.

Proposition 3.10. *The following assertions hold.*

(1) *For any $j \geq 0$, the following assertions hold.*

(a) *Let $J \subseteq (R_j)_I^{q.\,\mathrm{frep}}$ be a finitely generated ideal such that $J^k \subseteq \mathrm{Ker}(\Phi_0^{(j)})$ for some $k > 0$ (see Definition 3.8(4) for $\Phi_0^{(j)}$). Then $(R_j)_I^{q.\,\mathrm{frep}}$ is J-adically complete and separated.*

(b) *Let $x = (x_i)_{i \geq 0}$ be an element of $(R_j)_I^{q.\,\mathrm{frep}}$. Then x is a unit if and only if $x_0 \in R_j/IR_j$ is a unit.*

(c) *The ring map $(F_j)_I^{q.\,\mathrm{frep}}$ is an isomorphism.*

(2) *$(\{(R_i)_I^{q.\,\mathrm{frep}}\}_{i \geq 0}, \{(t_i)_I^{q.\,\mathrm{frep}}\}_{i \geq 0})$ is a perfect tower. In particular, each $(R_i)_I^{q.\,\mathrm{frep}}$ is reduced.*

Proof. (1) Since $(\{(R_{j+i})^{q.\,\mathrm{frep}}\}_{i \geq 0}, \{(t_{j+i})_I^{q.\,\mathrm{frep}}\}_{i \geq 0})$ is the inverse perfection of $(\{R_{j+i}\}_{i \geq 0}, \{t_{j+i}\}_{i \geq 0})$, we are reduced to showing the assertions in the case when $j = 0$.

(a): By definition, $(R_0)_I^{q.\,\mathrm{frep}}$ is complete and separated with respect to the linear topology induced by the descending filtration

$$\mathrm{Ker}(\Phi_0^{(0)}) \supseteq \mathrm{Ker}(\Phi_1^{(0)}) \supseteq \mathrm{Ker}(\Phi_2^{(0)}) \supseteq \cdots .$$

Moreover, since $J^k \subseteq \mathrm{Ker}(\Phi_0^{(0)})$, we have $(J^k)^{[p^i]} \subseteq \mathrm{Ker}(\Phi_i^{(0)})$ for every $i \geq 0$ by the commutative diagram (3-2).[9] On the other hand, since J^k is finitely generated, $(J^k)^{p^i r} \subseteq (J^k)^{[p^i]}$ for some $r > 0$. Thus the assertion follows from [15, Lemma 2.1.1].

(b): It is obvious that $x_0 \in \bar{R}_0$ is a unit if $x \in (R_0)_I^{q.\,\mathrm{frep}}$ is a unit. Conversely, assume that $x_0 \in \bar{R}_0$ is a unit. Then for every $i \geq 0$, $x_i^{p^i}$ is a unit because it is the image of x_0 in $\bar{R}_i$. Hence x_i is also a unit. Therefore, we have isomorphisms $R_i/IR_i \xrightarrow{\times x_i} R_i/IR_i$ ($i \geq 0$) that are compatible with the Frobenius projections. Thus we obtain the isomorphism between inverse limits $(R_0)_I^{q.\,\mathrm{frep}} \xrightarrow{\times x} (R_0)_I^{q.\,\mathrm{frep}}$, which yields the assertion.

(c): Consider the shifting map $s_0 : (R_0)_I^{q.\,\mathrm{frep}} \to (R_1)_I^{q.\,\mathrm{frep}}$ defined by the rule $s_0((a_i)_{i \geq 0}) := (a_{i+1})_{i \geq 0}$. Then one can easily check that s_0 is the inverse map of $(F_0)_I^{q.\,\mathrm{frep}}$.

(2) Define $F_{0,i}^{q.\,\mathrm{frep}} : (R_i)_I^{q.\,\mathrm{frep}} \to (R_0)_I^{q.\,\mathrm{frep}}$ as the composite map $(F_0)_I^{q.\,\mathrm{frep}} \circ \cdots \circ (F_{i-1})_I^{q.\,\mathrm{frep}}$ (if $i \geq 1$) or the identity map (if $i = 0$). Then the collection $\{F_{0,i}^{q.\,\mathrm{frep}}\}_{i \geq 0}$ gives a morphism of towers from $(\{(R_i)_I^{q.\,\mathrm{frep}}\}_{i \geq 0}, \{(t_i)_I^{q.\,\mathrm{frep}}\}_{i \geq 0})$ to $\left\{(R_0)_I^{q.\,\mathrm{frep}} \xrightarrow{F_{(R_0)_I^{q.\,\mathrm{frep}}}} (R_0)_I^{q.\,\mathrm{frep}} \xrightarrow{F_{(R_0)_I^{q.\,\mathrm{frep}}}} \cdots\right\}$. Using assertion (1-c) and Lemma 3.6(1), we complete the proof. □

The operation of inverse quasi-perfection preserves the locality of rings and ring maps.

Lemma 3.11. *Assume that R_i is a local ring for any $i \geq 0$, and $I \neq R$. Then for any $j \geq 0$, the following assertions hold.*

(1) *The ring maps $t_j, \bar{t}_j$, and F_j are local.*

(2) *$(R_j)_I^{q.\,\mathrm{frep}}$ is a local ring.*

(3) *The ring map $(t_j)_I^{q.\,\mathrm{frep}} : (R_j)_I^{q.\,\mathrm{frep}} \to (R_{j+1})_I^{q.\,\mathrm{frep}}$ is local.*

[9] The symbol $I^{[p^n]}$ for an ideal I in an $\mathbb{F}_p$-algebra A is the ideal generated by the elements x^{p^n} for $x \in I$.

Proof. As in Proposition 3.10(1), it suffices to show the assertions in the case when $j = 0$.

(1) Since the diagrams (3-2) and (3-3) are commutative, $F_0 \circ \overline{t}_0$ and $\overline{t}_0 \circ F_0$ are local. Hence $\overline{t}_0$ and F_0 are local. In particular, the composition $R_0 \twoheadrightarrow \overline{R}_0 \xrightarrow{\overline{t}_0} \overline{R}_1$ is local. Since this map factors through t_0, t_0 is also local, as desired.

(2) Let $\mathfrak{m}_0$ be the maximal ideal of R_0. Consider the ideal

$$(\mathfrak{m}_0)_I^{q.\,\mathrm{frep}} = \{(x_i)_{i\geq 0} \in (R_0)_I^{q.\,\mathrm{frep}} \mid x_0 \in \mathfrak{m}_0/IR_0\},$$

where $\mathfrak{m}_0/IR_0$ is the maximal ideal of $\overline{R}_0$. Then by Proposition 3.10(1-b), $(\mathfrak{m}_0)_I^{q.\,\mathrm{frep}}$ is a unique maximal ideal of $(R_0)_I^{q.\,\mathrm{frep}}$. The assertion follows.

(3) By assertion (2), $(\{(R_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0}, \{(t_i)_I^{q.\,\mathrm{frep}}\}_{i\geq 0})$ is a purely inseparable tower of local rings. Hence by (1), $(t_0)_I^{q.\,\mathrm{frep}}$ is local. □

A purely inseparable tower also satisfies the following amusing property. This is well-known in positive characteristic, in which case $R_i \to R_{i+1}$ gives a universal homeomorphism (i.e. the induced morphism of schemes $\operatorname{Spec} R_{i+1} \to \operatorname{Spec} R_i$ is a universally homeomorphism). See also Proposition 3.45.

Lemma 3.12. *For every $i \geq 0$, assume that R_i is I-adically Henselian.*[10] *Then the ring map t_i induces an equivalence of categories*:

$$\mathbf{F.\acute{E}t}(R_i) \xrightarrow{\cong} \mathbf{F.\acute{E}t}(R_{i+1}),$$

where $\mathbf{F.\acute{E}t}(A)$ is the category of finite étale A-algebras for a ring A.

Proof. By Corollary 3.7, we obtain the commutative diagram of rings

$$\begin{tikzcd} R_i \arrow[r, "t_i"] \arrow[d, "\pi_i"'] & R_{i+1} \arrow[d, "\pi_{i+1}"] \\ \overline{R}_i \arrow[r, "\overline{t}_i"] \arrow[d, "\phi_{\overline{R}_i}"'] & \overline{R}_{i+1} \arrow[d, "\phi_{\overline{R}_{i+1}}"] \\ (\overline{R}_i)^{\mathrm{perf}} \arrow[r, "(\overline{t}_i)^{\mathrm{perf}}"] & (\overline{R}_{i+1})^{\mathrm{perf}} \end{tikzcd} \tag{3-6}$$

where π_j ($j \in \{i, i+1\}$) is the natural projection, and the bottom map is an isomorphism. Since the Frobenius endomorphism on any $\mathbb{F}_p$-algebra gives a universal homeomorphism [38, Tag 0CC6], so does $\phi_{\overline{R}_j}$ by [38, Tag 01YW] and [38, Tag 01YZ]. Hence $\phi_{\overline{R}_j}$ induces an equivalence of categories of finite étale algebras over respective rings in view of [38, Tag 0BQN]. The same assertion holds for π_j by the lifting property of a henselian pair [38, Tag 09ZL]. By going around the diagram (3-6), we finish the proof. □

[10] This condition is realized if R_0 is I-adically Henselian and each $t_i : R_i \to R_{i+1}$ is integral.

3C. *Axioms for perfectoid towers.*

3C1. *Remarks on torsion.* In the subsequent Section 3C2, we introduce the class of *perfectoid towers* as a generalization of perfect towers. For this purpose, we need to deal with a purely inseparable tower arising from (R, I) in the case when $I = (0)$ at least, and hence plenty of I-torsion elements. Thus we begin by giving several preliminary lemmas on torsion of modules over rings.

Definition 3.13. Let R be a ring, and let M be an R-module.

(1) Let $x \in R$ be an element. We say that an element $m \in M$ is x-*torsion* if $x^n m = 0$ for some $n > 0$. We denote by $M_{x\text{-tor}}$ the R-submodule of M consisting of all x-torsion elements in M.

(2) Let $I \subseteq R$ be an ideal. We say that an element $m \in M$ is I-*torsion* if m is x-torsion for every $x \in I$. We denote by $M_{I\text{-tor}}$ the R-submodule of M consisting of all I-torsion elements in M. Note that $M_{(x)\text{-tor}} = M_{x\text{-tor}} = M_{x^n\text{-tor}}$ for every $n > 0$.

(3) For an element $x \in R$ (resp. an ideal $I \subseteq R$), we say that M *has bounded* x-*torsion* (resp. *bounded* I-*torsion*) if there exists some $l > 0$ such that $x^l M_{x\text{-tor}} = (0)$ $(I^l M_{I\text{-tor}} = (0))$.

(4) For an ideal $I \subseteq R$, we denote by $\varphi_{I,M} : M_{I\text{-tor}} \to M/IM$ the composition of natural R-linear maps:

$$M_{I\text{-tor}} \hookrightarrow M \twoheadrightarrow M/IM. \tag{3-7}$$

First we record the following fundamental lemma.

Lemma 3.14. *Let R be a ring, and let M be an R-module. Let $x \in R$ be an element. Then for every $n > 0$, we have*

$$M_{x\text{-tor}} \cap x^n M = x^n M_{x\text{-tor}}.$$

Proof. Pick an element $m \in M_{x\text{-tor}} \cap x^n M$. Then $m = x^n m_0$ for some $m_0 \in M$, and $x^l m = 0$ for some $l > 0$. Hence $x^{l+n} m_0 = 0$, which implies that $m_0 \in M_{x\text{-tor}}$ and thus $m \in x^n M_{x\text{-tor}}$. The containment $x^n M_{x\text{-tor}} \subseteq M_{x\text{-tor}} \cap x^n M$ is clear. □

Corollary 3.15. *Keep the notation as in Lemma 3.14, and suppose further that $x M_{x\text{-tor}} = (0)$. Then the map $\varphi_{(x),M} : M_{x\text{-tor}} \to M/xM$* (see Definition 3.13(4)) *is injective.*

Proof. It is clear from Lemma 3.14. □

Lemma 3.14 is also applied to show a half part of the following useful result.

Lemma 3.16. *Keep the notation as in Lemma 3.14, and suppose further that M has bounded x-torsion. Let $\widehat{M}$ be the x-adic completion of M, and let $\psi : M \to \widehat{M}$ be the natural map. Then the restriction $\psi_{\mathrm{tor}} : M_{x\text{-tor}} \to (\widehat{M})_{x\text{-tor}}$ of ψ is an isomorphism of R-modules.*

Proof. By assumption, there exists some $l > 0$ such that $x^l M_{x\text{-tor}} = (0)$. On the other hand, $\mathrm{Ker}(\psi_{\mathrm{tor}}) = M_{x\text{-tor}} \cap \bigcap_{n=0}^{\infty} x^n M$ is contained in $M_{x\text{-tor}} \cap x^l M$, which is equal to $x^l M_{x\text{-tor}}$ by Lemma 3.14. Hence ψ_{tor} is injective.

Let us prove the surjectivity. Let $\hat{N}$ denote the x-adic completion of N for every R-module N. Then we obtain the commutative diagram of R-modules:

$$\begin{tikzcd} 0 \arrow[r] & M_{x\text{-tor}} \arrow[r, "\iota"] \arrow[d, "\psi_{M_{x\text{-tor}}}"'] & M \arrow[r, "\pi"] \arrow[d, "\psi"] & M/M_{x\text{-tor}} \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & \widehat{M_{x\text{-tor}}} \arrow[r, "\hat{\iota}"] & \widehat{M} \arrow[r, "\widehat{\pi}"] & \widehat{M/M_{x\text{-tor}}} \arrow[r] & 0 \end{tikzcd} \tag{3-8}$$

where ι is the inclusion map and π is the natural projection. Since $\psi \circ \iota$ factors through ψ_{tor}, it suffices to show that $(\widehat{M})_{x\text{-tor}} \subseteq \operatorname{Im}(\hat{\iota} \circ \psi_{M_{x\text{-tor}}})$. First, $\psi_{M_{x\text{-tor}}}$ is bijective because it is isomorphic to the canonical isomorphism $M_{x\text{-tor}}/(x^l) \xrightarrow{\cong} \widehat{M_{x\text{-tor}}}/(x^l)$. To show that $(\widehat{M})_{x\text{-tor}} \subseteq \operatorname{Im}(\hat{\iota})$, note that the top row of (3-8) forms an exact sequence, and it consists of R-modules that have bounded x-torsion. Then by [38, Tag 0923] and right exactness of derived completion functors, $\operatorname{Ker}(\hat{\pi}) = \operatorname{Im}(\hat{\iota})$ (in fact, the bottom sequence is also exact because ψ_{tor} is injective). Since $\widehat{M/M_{x\text{-tor}}}$ is x-torsion free by [14, Chapter II, Lemma 1.1.5], $(\widehat{M})_{x\text{-tor}} \subseteq \operatorname{Ker}(\widehat{\pi})$. The assertion follows. ☐

The following lemma is used for proving Main Theorem 1 (cf. Lemma 3.48).

Lemma 3.17. *Let R be a ring, and let M be an R-module. Let $x \in R$ be an element. Then for every $n > 0$, we have*

$$\operatorname{Ann}_{M/x^nM}(x) \subseteq \operatorname{Im}(\varphi_{(x^n),M}) + x^{n-1}(M/x^nM). \tag{3-9}$$

Proof. Pick an element $m \in M$ such that $xm \in x^nM$. Then $x(m - x^{n-1}m') = 0$ for some $m' \in M$. In particular, $m - x^{n-1}m' \in M_{x^n\text{-tor}}$. Hence $m \bmod x^nM$ lies in the right-hand side of (3-9), as desired. ☐

In the case when $M = R$, we can regard $M_{I\text{-tor}}$ as a (possibly) nonunital subring of R. This point of view provides valuable insights. For example, "reducedness" for $R_{I\text{-tor}}$ induces a good property on boundedness of torsions.

Lemma 3.18. *Let (R, I) be a pair such that $R_{I\text{-tor}}$ does not contain any non-zero nilpotent element of R. Then $IR_{I\text{-tor}} = (0)$.*

Proof. It suffices to show that $xR_{I\text{-tor}} = 0$ for every $x \in I$. Pick an element $a \in R_{I\text{-tor}}$. Then for a sufficiently large $n > 0$, $x^na = 0$. Hence $(xa)^n = x^na \cdot a^{n-1} = 0$. Thus we have $xa = 0$ by assumption, as desired. ☐

Corollary 3.19. *Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a purely inseparable tower arising from some pair (R, I). Then for every $i \geq 0$ and every ideal $J \subseteq (R_i)_I^{q.\,\text{frep}}$, we have $J((R_i)_I^{q.\,\text{frep}})_{J\text{-tor}} = (0)$.*

Proof. Since $(R_i)_I^{q.\,\text{frep}}$ is reduced by Proposition 3.10(2), the assertion follows from Lemma 3.18. ☐

Furthermore, we can treat $R_{I\text{-tor}}$ as a positive characteristic object (in the situation of our interest), even if R is not an $\mathbb{F}_p$-algebra.

Lemma-Definition 3.20. *If (R, I) is a pair such that $p \in I$ and $IR_{I\text{-tor}} = (0)$, the multiplicative map*

$$R_{I\text{-tor}} \to R_{I\text{-tor}};\ x \mapsto x^p \tag{3-10}$$

is also additive. We denote by $F_{R_{I\text{-tor}}}$ the map (3-10).

Proof. This immediately follows from the binomial theorem. □

3C2. *Perfectoid towers and pillars.*

Definition 3.21 (perfectoid towers). Let R be a ring, and let $I_0 \subseteq R$ be an ideal. A tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is called a *(p)-perfectoid tower arising from* (R, I_0) if it is a p-purely inseparable tower arising from (R, I_0) (cf. Definition 3.4(1)) and satisfies the following additional axioms.

(d) For every $i \geq 0$, the i-th Frobenius projection $F_i : R_{i+1}/I_0R_{i+1} \to R_i/I_0R_i$ (cf. Definition 3.4(2)) is surjective.

(e) For every $i \geq 0$, R_i is an I_0-adically Zariskian ring (in other words, I_0R_i is contained in the Jacobson radical of R_i).

(f) I_0 is a principal ideal, and R_1 contains a principal ideal I_1 that satisfies the following axioms.

 (f-1) $I_1^p = I_0R_1$.

 (f-2) For every $i \geq 0$, $\mathrm{Ker}(F_i) = I_1(R_{i+1}/I_0R_{i+1})$.

(g) For every $i \geq 0$, $I_0(R_i)_{I_0\text{-tor}} = (0)$. Moreover, there exists a (unique) bijective map $(F_i)_{\mathrm{tor}} : (R_{i+1})_{I_0\text{-tor}} \to (R_i)_{I_0\text{-tor}}$ such that the diagram

$$\begin{array}{ccc} (R_{i+1})_{I_0\text{-tor}} & \xrightarrow{\varphi_{I_0,R_{i+1}}} & R_{i+1}/I_0R_{i+1} \\ {\scriptstyle (F_i)_{\mathrm{tor}}}\downarrow & & \downarrow{\scriptstyle F_i} \\ (R_i)_{I_0\text{-tor}} & \xrightarrow[\varphi_{I_0,R_i}]{} & R_i/I_0R_i \end{array} \tag{3-11}$$

commutes (see Definition 3.13 for the notation; see also Corollary 3.15).

Remark 3.22. If I_0 is generated by an element whose image in R_i is a non-zerodivisor for every $i \geq 0$, then axiom (g) is satisfied automatically. If R_1 is reduced and $I_0 = (0)$, then axiom (g) follows from axioms (d) and (f). Consequently, if every t_i is injective and $\varinjlim_{i\geq 0} R_i$ is a domain, one can ignore axiom (g) when checking that $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower.

We have some examples of perfectoid towers.

Example 3.23. (1) (cf. [37, Definition 4.4]) Let $(R, \mathfrak{m}, k)$ be a d-dimensional complete unramified regular local ring of mixed characteristic $p > 0$ whose residue field is perfect. Then we have

$$R \cong W(k)[[x_2, \ldots, x_d]].$$

For every $i \geq 0$, set $R_i := R[p^{1/p^i}, x_2^{1/p^i}, \dots, x_d^{1/p^i}]$, and let $t_i : R_i \to R_{i+1}$ be the inclusion map. Then the tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower arising from $(R, (p))$. Indeed, the Frobenius projection $F_i : R_{i+1}/pR_{i+1} \to R_i/pR_i$ is given as the p-th power map.[11]

(2) For some generalization of (1), one can build a perfectoid tower arising from a complete local log-regular ring. For details, see Section 3F.

(3) We note that t_i (resp. F_i) of a perfectoid tower is not necessarily the inclusion map (resp. the p-th power map). For instance, let R be a reduced $\mathbb{F}_p$-algebra. Set $R_i := R$, $t_i := F_R$, and $F_i := \mathrm{id}_R$ for every $i \geq 0$. Then $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower arising from $(R, (0))$.

The class of perfectoid towers is a generalization of perfect towers.

Lemma 3.24. *Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a tower of $\mathbb{F}_p$-algebras. The following conditions are equivalent.*

(1) *$(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfect $\mathbb{F}_p$-tower* (cf. Definition 3.2).

(2) *$(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a p-perfectoid tower arising from $(R_0, (0))$.*

Proof. (1) ⇒ (2) We may assume that $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is of the form $R \xrightarrow{F_R} R \xrightarrow{F_R} R \xrightarrow{F_R} \cdots$ (see Definition 3.2). By Example 3.5, this is a purely inseparable tower arising from $(R, (0))$. Axiom (e) in Definition 3.21 is obvious. Axioms (d), (f), and (g) are also satisfied, since the Frobenius projection F_i (cf. Example 3.5) is an isomorphism for any $i \geq 0$. This yields the assertion.

(2) ⇒ (1) Conversely, assume that $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower arising from $(R_0, (0))$. Since F_i is identified with $(F_i)_{\mathrm{tor}}$ in axiom (g), the injectivity of $(F_i)_{\mathrm{tor}}$ implies that F_i is injective. In other words, R_i is reduced by Lemma 3.6(1). Furthermore, F_i is an isomorphism by axiom (d) or the surjectivity of $(F_i)_{\mathrm{tor}}$. Hence we obtained the desired isomorphism of towers:

$$\begin{tikzcd} R_0 \arrow[r, "t_0"] \arrow[d, "\mathrm{id}_{R_0}"] & R_1 \arrow[r, "t_1"] \arrow[d, "F_0"] & R_2 \arrow[r, "t_2"] \arrow[d, "F_0\circ F_1"] & R_3 \arrow[r, "t_3"] \arrow[d, "F_0\circ F_1\circ F_2"] & \cdots \\ R_0 \arrow[r, "F_{R_0}"'] & R_0 \arrow[r, "F_{R_0}"'] & R_0 \arrow[r, "F_{R_0}"'] & R_0 \arrow[r, "F_{R_0}"'] & \cdots . \end{tikzcd} \tag{3-12}$$

The assertion follows. □

For a perfectoid tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ arising from (R, I_0), an ideal $I_1 \subseteq R_1$ appearing in axiom (f) in Definition 3.21 is unique. Indeed, it contains I_0R_1, and its image via the projection $R_1 \to \overline{R}_1$ is a fixed ideal $\mathrm{Ker}(F_0)$.

Definition 3.25. We call I_1 the *first perfectoid pillar of $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ arising from (R, I_0).*

The relationship between I_0 and I_1 can be observed also in higher layers (see Proposition 3.26 below). In the rest of this section, we fix a perfectoid tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ arising from some pair (R, I_0), and let I_1 denote the first perfectoid pillar.

[11] Axiom (f-2) follows from the normality of R_i. The other axioms are clearly satisfied.

Proposition 3.26. (1) *For a sequence of principal ideals* $\{I_i \subseteq R_i\}_{i\geq 2}$, *the following conditions are equivalent.*

(a) $F_i^{-1}(I_i\overline{R}_i) = I_{i+1}\overline{R}_{i+1}$ *for every* $i \geq 0$.

(b) $F_i(I_{i+1}\overline{R}_{i+1}) = I_i\overline{R}_i$ *for every* $i \geq 0$.

(2) *Each one of the equivalent conditions in* (1) *implies that* $I_{i+1}^p = I_iR_{i+1}$ *for every* $i \geq 0$.

(3) *There exists a unique sequence of principal ideals* $\{I_i \subseteq R_i\}_{i\geq 0}$ *that satisfies one of the equivalent conditions in* (1). *Moreover, there exists a sequence of elements* $\{\overline{f}_i \in \overline{R}_i\}_{i\geq 0}$ *such that* $I_i\overline{R}_i = (\overline{f}_i)$ *and* $F_i(\overline{f}_{i+1}) = \overline{f}_i$ *for every* $i \geq 0$.

Proof. (1) Since the implication (a) $\Rightarrow$ (b) follows from axiom (d) in Definition 3.21, it suffices to show the converse. Assume that condition (b) is satisfied. Then for every $i \geq 0$, the compatibility $\overline{t}_i \circ F_i = F_{\overline{R}_{i+1}}$ implies

$$I_{i+1}^p\overline{R}_{i+1} = I_i\overline{R}_{i+1} \tag{3-13}$$

because I_{i+1} is principal. In particular, $\mathrm{Ker}(F_i) = I_1\overline{R}_{i+1} \subseteq I_{i+1}\overline{R}_{i+1}$ (cf. axiom (f-2)). On the other hand, by the surjectivity of F_i and the assumption again, we have $F_i(F_i^{-1}(I_i\overline{R}_i)) = I_i\overline{R}_i = F_i(I_{i+1}\overline{R}_{i+1})$. Hence

$$F_i^{-1}(I_i\overline{R}_i) \subseteq I_{i+1}\overline{R}_{i+1} + \mathrm{Ker}(F_i) \subseteq I_{i+1}\overline{R}_{i+1} \subseteq F_i^{-1}(I_i\overline{R}_i),$$

which yields the assertion.

(2) Let us deduce the assertion from (3-13) by induction. By definition, $I_1^p = I_0R_1$. We then fix some $i \geq 1$. Suppose that for every $1 \leq k \leq i$, $I_k^p = I_{k-1}R_k$. Then $I_0R_i = I_i^{p^i}$. In particular, R_i is I_i-adically Zariskian by axiom (e). Moreover, by (3-13), we have the equalities of R_i-modules:

$$I_iR_{i+1} = I_{i+1}^p + I_0R_{i+1} = I_{i+1}^p + I_i^{p^i-1}(I_iR_{i+1}).$$

Hence by Nakayama's lemma, we obtain $I_{i+1}^p = I_iR_{i+1}$ as desired.

(3) By the axiom of (dependent) choice, the existence follows from axiom (d) in Definition 3.21. Let us show the uniqueness of $\{I_i \subseteq R_i\}_{i\geq 0}$ that satisfies condition (a) in (1). For every $i \geq 0$, $I_iR_{i+1} \subseteq I_{i+1}$ by (2), and hence I_{i+1} is the inverse image of $F_i^{-1}(I_i\overline{R}_i)$ via the projection $R_{i+1} \to \overline{R}_{i+1}$. Since I_0 is fixed, the assertion follows. $\square$

Definition 3.27. In the situation described in Proposition 3.26(3), we call I_i the *i-th perfectoid pillar of* $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ *arising from* (R_0, I_0).

The following property of perfectoid pillars is applied to prove our main result.

Lemma 3.28. *Let* $\{I_i\}_{i\geq 0}$ *denote the system of perfectoid pillars of* $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$, *and let* $\pi_i : R_i/I_0R_i \to R_i/I_iR_i$ $(i \geq 0)$ *be the natural projections. Then for every* $i \geq 0$, *there exists a unique isomorphism of rings*

$$F_i' : R_{i+1}/I_{i+1}R_{i+1} \xrightarrow{\cong} R_i/I_iR_i$$

such that $\pi_i \circ F_i = F'_i \circ \pi_{i+1}$.

Proof. Since F_i and π_i are surjective, the assertion follows from $\mathrm{Ker}(\pi_i \circ F_i) = F_i^{-1}(I_i(R_i/I_0R_i)) = I_{i+1}(R_{i+1}/I_0R_{i+1})$. □

3D. *Tilts of perfectoid towers.*

3D1. *Invariance of some properties.* Here we establish tilting operation for perfectoid towers. For this, we first introduce the notion of *small tilt*, which originates in [37].

Definition 3.29 (small tilts). Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a perfectoid tower arising from some pair (R, I_0).

(1) For any $j \geq 0$, we define the *j-th small tilt* of $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ associated to (R, I_0) as the j-th inverse quasi-perfection of $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ associated to (R, I_0) and denote it by $(R_j)^{s.\flat}_{I_0}$.

(2) Let the notation be as in Lemma 3.28. Then we define $I_i^{s.\flat} := \mathrm{Ker}(\pi_i \circ \Phi_0^{(i)})$ for every $i \geq 0$.

Note that the ideal $I_i^{s.\flat} \subseteq R_i^{s.\flat}$ has the following property.

Lemma 3.30. *Keep the notation as in Definition 3.29. Then for every $i \geq 0$ and $j \geq 0$, we have* $\Phi_i^{(j)}(I_j^{s.\flat}) = I_{j+i}\bar{R}_{j+i}$.

Proof. Since $\Phi_0^{(j)}$ is surjective, we have $\Phi_0^{(j)}(I_j^{s.\flat}) = I_j\bar{R}_j$. On the other hand, since $\Phi_0^{(j)} = F_j \circ \Phi_1^{(j)}$, we have

$$F_j^{-1}(\Phi_0^{(j)}(I_j^{s.\flat})) = \Phi_1^{(j)}(I_j^{s.\flat}) + \mathrm{Ker}(F_j) = \Phi_1^{(j)}(I_j^{s.\flat}).$$

Hence by condition (a) in Proposition 3.26(1), $\Phi_1^{(j)}(I_j^{s.\flat}) = I_{j+1}\bar{R}_{j+1}$. By repeating this procedure recursively, we obtain the assertion. □

The next lemma provides some completeness of the small tilts attached to a perfectoid tower of characteristic $p > 0$ (see also Remark 3.33).

Lemma 3.31. *Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a perfectoid tower arising from $(R, (0))$. Then, for any element $f \in R$ and any $j \geq 0$, the inverse limit $\varprojlim\{\cdots \xrightarrow{\bar{F}}_{j+1} R_{j+1}/fR_{j+1} \xrightarrow{\bar{F}_j} R_j/fR_j\}$ is isomorphic to the f-adic completion of R_j.*

Proof. It suffices to show the assertion when $j = 0$. Let $(\{R'_i\}_{i\geq 0}, \{t'_i\}_{i\geq 0})$ denote the standard perfect tower (3-1) arising from R. By Lemma 3.24, (3-12) gives a canonical isomorphism $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0}) \xrightarrow{\cong} (\{R'_i\}_{i\geq 0}, \{t'_i\}_{i\geq 0})$. If we put $J_0 = fR'_0$, then $R'_i/J_0R'_i = R/f^{p^i}R$ for every $i \geq 0$. Hence we have the desired canonical isomorphisms:

$$\varprojlim\{\cdots \xrightarrow{\bar{F}_1} R_1/fR_1 \xrightarrow{\bar{F}_0} R_0/fR_0\} \xrightarrow{\cong} \varprojlim_{n\geq 0} R/f^{p^n}R \xrightarrow{\cong} \varprojlim_{n\geq 0} R/f^nR.$$ □

Example 3.32. Let S be a perfect $\mathbb{F}_p$-algebra. Pick an arbitrary $f \in S$, and let $\hat{S}$ denote the f-adic completion. We obtain a canonical isomorphism $\varprojlim_{\mathrm{Frob}} S/fS \xrightarrow{\cong} \hat{S}$ by applying the above proof to the tower

$$S \xrightarrow{\mathrm{id}_S} S \xrightarrow{\mathrm{id}_S} S \xrightarrow{\mathrm{id}_S} \cdots.$$

Remark 3.33. In the situation of Lemma 3.31, assume further that $\bar{t}_i : R_i/fR_i \to R_{i+1}/fR_{i+1}$ is injective for every $i \geq 0$. Then $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a purely inseparable tower arising from $(R, (f))$ with Frobenius projections $\{\overline{F}_i : R_{i+1}/fR_{i+1} \to R_i/fR_i\}_{i\geq 0}$. Furthermore, it satisfies axioms (d), (f), and (g) in Definition 3.21. To check this, we may assume that $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is the standard perfect tower (3-1). Then $\overline{F}_i$ is the natural projection $R/f^{p^{i+1}}R \twoheadrightarrow R/f^{p^i}R$. It is clearly surjective, and its kernel is $f^{p^i}(R/f^{p^{i+1}}R)$. Let I_i be the ideal of R_i generated by $f \in R_i$ $(= R)$. Then $I_0R_i = f^{p^i}R$ and $I_1R_{i+1} = f^{p^i}R$. Hence $I_1^p = I_0R_1$ and $\mathrm{Ker}(\overline{F}_i) = I_1\overline{R}_{i+1}$. Finally, note that $(R_i)_{I_0\text{-tor}} = R_{f\text{-tor}}$. Then $I_0(R_i)_{I_0\text{-tor}} = f^{p^i}R_{f\text{-tor}} = (0)$ by Lemma 3.18, and we can take $\mathrm{id}_{R_{f\text{-tor}}}$ as the bijection $(\overline{F}_i)_{\mathrm{tor}}$ fitting into the diagram (3-11).

Definition 3.34 (tilts of perfectoid towers). Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a perfectoid tower arising from some pair (R, I). Then we define *the tilt of* $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ associated to (R, I) as the inverse perfection of $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ associated to (R, I), and denote it by $(\{(R_i)_I^{s.\flat}\}_{i\geq 0}, \{(t_i)_I^{s.\flat}\}_{i\geq 0})$. If no confusion occurs, we can abbreviate $(R_i)_I^{s.\flat}$ and $(t_i)_I^{s.\flat}$) to $R_i^{s.\flat}$ and $t_i^{s.\flat}$.

After discussing several basic properties of this tilting operation, we illustrate how to compute the tilts of perfectoid towers in some specific cases; when they consist of *log-regular rings* (see Theorem 3.61 and Example 3.62).

We should remark that all results on the perfection of purely inseparable towers (established in Section 3B) can be applied to the tilts of perfectoid towers.

Let us state Main Theorem 1 in a more refined form. This is an important tool when one wants to see that a certain correspondence holds between Noetherian rings of mixed characteristic and those of positive characteristic.

Theorem 3.35. *Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a perfectoid tower arising from some pair (R, I_0), and let $\{I_i\}_{i\geq 0}$ be the system of perfectoid pillars. Let $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$ denote the tilt associated to (R, I_0).*

(1) *For every $j \geq 0$ and every element $f_j^{s.\flat} \in R_j^{s.\flat}$, the following conditions are equivalent.*
   (a) *$f_j^{s.\flat}$ is a generator of $I_j^{s.\flat}$.*
   (b) *For every $i \geq 0$, $\Phi_i^{(j)}(f_j^{s.\flat})$ is a generator of $I_{j+i}\overline{R}_{j+i}$.*

   In particular, $I_j^{s.\flat}$ is a principal ideal, and $(I_{j+1}^{s.\flat})^p = I_j^{s.\flat}R_{j+1}^{s.\flat}$.

(2) *We have isomorphisms of (possibly) nonunital rings $(R_j^{s.\flat})_{I_0^{s.\flat}\text{-tor}} \cong (R_j)_{I_0\text{-tor}}$ that are compatible with $\{t_j\}_{j\geq 0}$ and $\{t_j^{s.\flat}\}_{j\geq 0}$.*

We give its proof in the subsequent Section 3D2. Before that, let us observe that this theorem induces many good properties of tilting. In the rest of this subsection, we keep the notation as in Theorem 3.35.

Lemma 3.36. *For every $i \geq 0$, $R_i^{s.\flat}$ is $I_0^{s.\flat}$-adically complete and separated.*

Proof. By Theorem 3.35, the ideal $I_0^{s.\flat}R_i^{s.\flat} \subseteq R_i^{s.\flat}$ is principal. Hence one can apply Proposition 3.10(1-a) to deduce the assertion. □

To discuss perfectoidness for the tilt $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$, we introduce the following maps.

Definition 3.37. For every $i \geq 0$, we define a ring map $(F_i)^{s.\flat}_{I_0} : (R_{i+1})^{s.\flat}_{I_0}/I_0^{s.\flat}(R_{i+1})^{s.\flat}_{I_0} \to (R_i)^{s.\flat}_{I_0}/I_0^{s.\flat}(R_i)^{s.\flat}_{I_0}$ by the rule

$$(F_i)^{s.\flat}_{I_0}(\alpha_{i+1} \bmod I_0^{s.\flat}(R_{i+1})^{s.\flat}_{I_0}) = (F_i)^{q.\,\mathrm{frep}}_{I_0}(\alpha_{i+1}) \bmod I_0^{s.\flat}(R_i)^{s.\flat}_{I_0},$$

where $\alpha_{i+1} \in (R_{i+1})^{s.\flat}_{I_0}$. If no confusion occurs, we can abbreviate $(F_i)^{s.\flat}_{I_0}$ to $F_i^{s.\flat}$.

Remark 3.38. Although the symbols $(\;\cdot\;)^{s.\flat}$ and $(\;\cdot\;)^{q.\,\mathrm{frep}}$ had been used interchangeably before Definition 3.37, $(F_i)^{s.\flat}_{I_0}$ differs from $(F_i)^{q.\,\mathrm{frep}}_{I_0}$ in general.

The following lemma is an immediate consequence of Theorem 3.35(1), but quite useful.

Lemma 3.39. *For every $j \geq 0$, $\Phi_0^{(j)}$ induces an isomorphism*

$$\overline{\Phi_0^{(j)}} : R_j^{s.\flat}/I_0^{s.\flat}R_j^{s.\flat} \xrightarrow{\cong} R_j/I_0R_j;\; a \bmod I_0^{s.\flat}R_j^{s.\flat} \mapsto \Phi_0^{(j)}(a). \tag{3-14}$$

Moreover, $\{\overline{\Phi_0^{(i)}}\}_{i\geq 0}$ is compatible with $\{t_i\}_{i\geq 0}$ (resp. $\{F_{R_i^{s.\flat}/I_0^{s.\flat}R_i^{s.\flat}}\}_{i\geq 0}$, resp. $\{F_i^{s.\flat}\}_{i\geq 0}$) and $\{t_i^{s.\flat}\}_{i\geq 0}$ (resp. $\{F_{R_i/I_0R_i}\}_{i\geq 0}$, resp. $\{F_i\}_{i\geq 0}$).

Proof. By axiom (d) in Definition 3.21, (3-14) is surjective. We check the injectivity. By Theorem 3.35(1), $I_0^{s.\flat}$ is generated by an element $f_0^{s.\flat} \in R_0^{s.\flat}$ such that $\Phi_i^{(0)}(f_0^{s.\flat})$ is a generator of $I_i\bar{R}_i$ $(i \geq 0)$. Note that $(\{R_{j+i}\}_{i\geq 0}, \{t_{j+i}\}_{i\geq 0})$ is a perfectoid tower arising from (R_j, I_0R_j). Moreover, $\{I_iR_{j+i}\}_{i\geq 0}$ is the system of perfectoid pillars associated to (R_j, I_0R_j) (cf. condition (b) in Proposition 3.26(1)). Put $J_0 := I_0R_j$. Then by Theorem 3.35(1) again, we find that $J_0^{s.\flat} = f_0^{s.\flat}R_j^{s.\flat} = I_0^{s.\flat}R_j^{s.\flat}$. Since $J_0^{s.\flat} = \operatorname{Ker}\Phi_0^{(j)}$, we obtain the first assertion.

One can deduce that $\{\overline{\Phi_0^{(i)}}\}_{i\geq 0}$ is compatible with the Frobenius projections from the commutativity of (3-2), because the other compatibility assertions immediately follow from the construction. ☐

Remark 3.40. Theorem 3.35(2) and Lemma 3.39 can be interpreted as a correspondence of homological invariants between R_i and $R_i^{s.\flat}$ by using Koszul homologies. Indeed, for any generator f_0 (resp. $f_0^{s.\flat}$) of I_0 (resp. $I_0^{s.\flat}$), the Koszul homology $H_q(f_0^{s.\flat}; R_i^{s.\flat})$ is isomorphic to $H_q(f_0; R_i)$ for any $q \geq 0$ as an abelian group.[12]

Now we can show the invariance of several properties of perfectoid towers under tilting. The first one is perfectoidness, which is most important in our framework.

Proposition 3.41. $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$ *is a perfectoid tower arising from* $(R_0^{s.\flat}, I_0^{s.\flat})$.

Proof. By Lemma 3.39 and Remark 3.33, $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$ is a purely inseparable tower arising from $(R_0^{s.\flat}, I_0^{s.\flat})$ that also satisfies axioms (d), (f), and (g). Moreover, axiom (e) holds by Lemma 3.36. Hence the assertion follows. ☐

Next we focus on finiteness properties. "Small" in the name of small tilts comes from the following fact.

[12]Note that $(R_i)_{I_0\text{-tor}} = \operatorname{Ann}_{R_i}(I_0)$ by axiom (g), and $(R_i^{s.\flat})_{I_0^{s.\flat}\text{-tor}} = \operatorname{Ann}_{R_i^{s.\flat}}(I_0^{s.\flat})$ by Corollary 3.19.

Proposition 3.42. *For every $j \geq 0$, the following assertions hold.*

(1) *If $t_j : R_j \to R_{j+1}$ is module-finite, then so is $t_j^{s.\flat} : R_j^{s.\flat} \to R_{j+1}^{s.\flat}$. Moreover, the converse holds true when R_j is I_0-adically complete and separated.*

(2) *If R_j is a Noetherian ring, then so is $R_j^{s.\flat}$. Moreover, the converse holds true when R_j is I_0-adically complete and separated.*

(3) *Assume that R_j is a Noetherian local ring, and a generator of I_0R_j is regular. Then the dimension of R_j is equal to that of $R_j^{s.\flat}$.*

Proof. (1) By Lemma 3.39, $\overline{t}_j : R_j/I_0R_j \to R_{j+1}/I_0R_{j+1}$ is module-finite if and only if $\overline{t_j^{s.\flat}} : R_j^{s.\flat}/I_0^{s.\flat}R_j^{s.\flat} \to R_{j+1}^{s.\flat}/I_0^{s.\flat}R_{j+1}^{s.\flat}$ is so. Thus by Lemma 3.36 and [29, Theorem 8.4], the assertion follows.

(2) One can prove this assertion by applying Lemma 3.36, Lemma 3.39, and [38, Tag 05GH].

(3) By Theorem 3.35, $I_0^{s.\flat}R_j^{s.\flat}$ is also generated by a regular element. Thus we obtain the equalities $\dim R_j = \dim R_j/I_0R_j + 1$ and $\dim R_j^{s.\flat} = \dim R_j^{s.\flat}/I_0^{s.\flat}R_j^{s.\flat} + 1$. By combining these equalities with Lemma 3.39, we deduce assertion. □

Proposition 3.42(2) says that Noetherianness for a perfectoid tower is preserved under tilting.

Definition 3.43. We say that $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is *Noetherian* if R_i is Noetherian for each $i \geq 0$.

Corollary 3.44. *If $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is Noetherian, then so is the tilt $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$. Moreover, the converse holds true when R_i is I_0-adically complete and separated for each $i \geq 0$.*

Proof. This immediately follows from Proposition 3.42(2). □

Finally, let us consider perfectoid towers of henselian rings. Then we obtain the equivalence of categories of finite étale algebras over each layer.

Proposition 3.45. *Assume that R_i is I_0-adically Henselian for any $i \geq 0$. Then we obtain the following equivalences of categories*:

$$\mathbf{F.\acute{E}t}(R_i^{s.\flat}) \xrightarrow{\cong} \mathbf{F.\acute{E}t}(R_i).$$

Proof. This follows from Lemma 3.36, Lemma 3.39 and [38, Tag 09ZL]. □

3D2. *Proof of Main Theorem 1.* We keep the notation as above. Furthermore, we set $\overline{I}_i := I_i\overline{R}_i$ for every $i \geq 0$. To prove Theorem 3.35, we investigate some relationship between $(R_i)_{I_0\text{-tor}}$ and $\mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)$. First recall that we can regard $(R_i)_{I_0\text{-tor}}$ as a nonunital subring of $\overline{R}_i$ by Corollary 3.15. Moreover, the map $\overline{t}_i$ naturally restricts to $(R_i)_{I_0\text{-tor}} \hookrightarrow (R_{i+1})_{I_0\text{-tor}}$, as follows.

Lemma 3.46. *For every $i \geq 0$, let $(t_i)_{\mathrm{tor}} : (R_i)_{I_0\text{-tor}} \to (R_{i+1})_{I_0\text{-tor}}$ be the restriction of t_i.*

(1) *$(t_i)_{\mathrm{tor}}$ is the unique map such that $\varphi_{I_0,R_{i+1}} \circ (t_i)_{\mathrm{tor}} = \overline{t}_i \circ \varphi_{I_0,R_i}$.*

(2) $(t_i)_{\mathrm{tor}} \circ (F_i)_{\mathrm{tor}} = (F_{i+1})_{\mathrm{tor}} \circ (t_{i+1})_{\mathrm{tor}} = F_{(R_{i+1})_{I_0\text{-tor}}}$.

Proof. Since φ_{I_0,R_i} is injective by Corollary 3.15, assertion (1) is clear from the construction. Hence we can regard $(t_i)_{\mathrm{tor}}$ and $(F_i)_{\mathrm{tor}}$ as the restrictions of $\overline{t}_i$ and F_i. Thus assertion (2) follows from the compatibility $\overline{t}_i \circ F_i = F_{i+1} \circ \overline{t}_{i+1} = F_{\overline{R}_{i+1}}$ induced by Lemma 3.6(3). $\square$

The map $\varphi_{I_0,R_i} : (R_i)_{I_0\text{-tor}} \hookrightarrow R_i/I_0R_i$ restricts to $\mathrm{Ann}_{R_i}(I_i) \hookrightarrow \mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)$. On the other hand, $\mathrm{Ann}_{R_i}(I_i)$ turns out to be equal to $(R_i)_{I_0\text{-tor}}$ by the following lemma.

Lemma 3.47. *For every $i \geq 0$, $I_i(R_i)_{I_0\text{-tor}} = 0$. In particular, $\mathrm{Im}(\varphi_{I_0,R_i}) \subseteq \mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)$.*

Proof. By Lemma 3.46(2) and axiom (g) in Definition 3.21, we find that $F_{(R_i)_{I_0\text{-tor}}}$ is injective. In other words, $(R_i)_{I_0\text{-tor}}$ does not contain any non-zero nilpotent element. Moreover, $(R_i)_{I_0\text{-tor}} = (R_i)_{I_i\text{-tor}}$. Hence the assertion follows from Lemma 3.18. $\square$

The following lemma is essential for proving Theorem 3.35.

Lemma 3.48. *For every $i \geq 0$, F_i restricts to a $\mathbb{Z}$-linear map $\mathrm{Ann}_{\overline{R}_{i+1}}(\overline{I}_{i+1}) \to \mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)$. Moreover, the resulting inverse system $\{\mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)\}_{i\geq 0}$ has the following properties.*

(1) *For every $j \geq 0$, $\varprojlim^1_{i\geq 0} \mathrm{Ann}_{\overline{R}_{j+i}}(\overline{I}_{j+i}) = (0)$.*

(2) *There are isomorphisms of $\mathbb{Z}$-linear maps $\varprojlim_{i\geq 0} \mathrm{Ann}_{\overline{R}_{j+i}}(\overline{I}_{j+i}) \cong (R_j)_{I_0\text{-tor}}$ $(j \geq 0)$ that are multiplicative, and compatible with $\{t_j^{s,\flat}\}_{j\geq 0}$ and $\{t_j\}_{j\geq 0}$.*

Proof. Since $F_i(\overline{I}_{i+1}) = \overline{I}_i$, F_i restricts to a $\mathbb{Z}$-linear map $(F_i)_{\mathrm{ann}} : \mathrm{Ann}_{\overline{R}_{i+1}}(\overline{I}_{i+1}) \to \mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)$. Let $\varphi_i : (R_i)_{I_0\text{-tor}} \hookrightarrow \mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)$ be the restriction of φ_{I_0,R_i}. By Lemma 3.17 and Lemma 3.47, we can write $\mathrm{Ann}_{\overline{R}_i}(\overline{I}_i) = \mathrm{Im}(\varphi_i) + \overline{I}_i^{p^i-1}$. Moreover, $\mathrm{Im}(\varphi_i) \cap \overline{I}_i^{p^i-1} = (0)$ by Lemma 3.14 and Lemma 3.47. Hence we have the following ladder with exact rows:

$$\begin{tikzcd} 0 \arrow[r] & (R_{i+1})_{I_0\text{-tor}} \arrow[r, "\varphi_{i+1}"] \arrow[d, "(F_i)_{\mathrm{tor}}"] & \mathrm{Ann}_{\overline{R}_{i+1}}(\overline{I}_{i+1}) \arrow[r] \arrow[d] & \overline{I}_{i+1}^{p^{i+1}-1} \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & (R_i)_{I_0\text{-tor}} \arrow[r, "\varphi_i"] & \mathrm{Ann}_{\overline{R}_i}(\overline{I}_i) \arrow[r] & \overline{I}_i^{p^i-1} \arrow[r] & 0 \end{tikzcd} \tag{3-15}$$

where the second and third vertical maps are the restrictions of F_i. Since $F_i(\overline{I}_{i+1}^{p^{i+1}-1}) = 0$, both functors $\varprojlim_{i\geq 0}$ and $\varprojlim^1_{i\geq 0}$ assign (0) to the inverse system $\{\overline{I}_{j+i}^{p^{j+i}-1}\}_{i\geq 0}$. Moreover, since $(F_i)_{\mathrm{tor}}$ is bijective, $\varprojlim_{i\geq 0}(R_{j+i})_{I_0\text{-tor}} \cong (R_j)_{I_0\text{-tor}}$ and $\varprojlim^1_{i\geq 0}(R_{j+i})_{I_0\text{-tor}} = (0)$. Hence we find that $\varprojlim^1_{i\geq 0} \mathrm{Ann}_{\overline{R}_{j+i}}(\overline{I}_{j+i}) = (0)$, which is assertion (1). Furthermore, we obtain the isomorphisms of $\mathbb{Z}$-modules:

$$(R_j)_{I_0\text{-tor}} \xleftarrow{(\Phi_0^{(j)})_{\mathrm{tor}}} \varprojlim_{i\geq 0}(R_{j+i})_{I_0\text{-tor}} \xrightarrow{\varprojlim_{i\geq 0}\varphi_{j+i}} \varprojlim_{i\geq 0} \mathrm{Ann}_{\overline{R}_{j+i}}(\overline{I}_{j+i}) \tag{3-16}$$

(where $(\Phi_0^{(j)})_{\mathrm{tor}}$ denotes the 0-th projection map), which are also multiplicative. Let us deduce (2) from it. Since $t_j^{s,\flat} = \varprojlim_{i\geq 0} \overline{t}_{j+i}$ by definition, the maps $\varprojlim_{i\geq 0} \varphi_{j+i}$ $(j \geq 0)$ are compatible with $\{\varprojlim_{i\geq 0}(t_{j+i})_{\mathrm{tor}}\}_{j\geq 0}$ (induced by Lemma 3.46(2)) and $\{t_j^{s,\flat}\}_{j\geq 0}$ by Lemma 3.46(1). On the other hand, the projections $(\Phi_0^{(j)})_{\mathrm{tor}}$ $(j \geq 0)$ are compatible with $\{\varprojlim_{i\geq 0}(t_{j+i})_{\mathrm{tor}}\}_{j\geq 0}$ and $\{(t_j)_{\mathrm{tor}}\}_{j\geq 0}$. The assertion follows. $\square$

Let us complete the proof of Theorem 3.35.

Proof of Theorem 3.35. (1) The implication (a) $\Rightarrow$ (b) follows from Lemma 3.30. Let us show the converse (b)$\Rightarrow$(a). For every $i \geq 0$, put $\overline{f}_{j+i} := \Phi_i^{(j)}(f_j^{s.\flat})$, and let π_i and F_i' be as in Lemma 3.28. Then, by the assumption, we have the following commutative ladder with exact rows:

$$\begin{tikzcd}
0 \arrow[r] & (\overline{f}_{i+1}) \arrow[r, "\iota_{i+1}"] \arrow[d] & \overline{R}_{i+1} \arrow[r, "\pi_{i+1}"] \arrow[d, "F_i"] & R_{i+1}/I_{i+1} \arrow[r] \arrow[d, "F_i'"] & 0 \\
0 \arrow[r] & (\overline{f}_i) \arrow[r, "\iota_i"] & \overline{R}_i \arrow[r, "\pi_i"] & R_i/I_i \arrow[r] & 0
\end{tikzcd}$$

where ι_i is the inclusion map. Let us consider the exact sequence obtained by taking inverse limits for all columns of the above ladder. Then, since each F_i' is an isomorphism, the map $\varprojlim_{i\geq 0} \pi_{j+i} : R_j^{s.\flat} \to \varprojlim_{i\geq 0} R_{j+i}/I_{j+i}$ is isomorphic to $\pi_j \circ \Phi_0^{(j)}$. Thus we find that $I_j^{s.\flat} = \mathrm{Im}(\varprojlim_{i\geq 0} \iota_{j+i})$. Let us show that the ideal $\mathrm{Im}(\varprojlim_{i\geq 0} \iota_{j+i}) \subseteq R_j^{s.\flat}$ is generated by $f_j^{s.\flat}$. For $i \geq 0$, let $\mu_i : \overline{R}_i \to (\overline{f}_i)$ be the $\overline{R}_i$-linear map induced by multiplication by $\overline{f}_i$. Then we obtain the commutative ladder

$$\begin{tikzcd}
\overline{R}_{i+1} \arrow[r, "\mu_{i+1}"] \arrow[d, "F_i"] & (\overline{f}_{i+1}) \arrow[r, "\iota_{i+1}"] \arrow[d] & \overline{R}_{i+1} \arrow[d, "F_i"] \\
\overline{R}_i \arrow[r, "\mu_i"] & (\overline{f}_i) \arrow[r, "\iota_i"] & \overline{R}_i .
\end{tikzcd}$$

Then, since $\mathrm{Ker}\,\mu_i = \mathrm{Ann}_{\overline{R}_i}(\overline{I}_i)$ for every $i \geq 0$, $\varprojlim_{i\geq 0} \mu_{j+i}$ is surjective by Lemma 3.48(1). Hence we have $\mathrm{Im}(\varprojlim_{i\geq 0} \iota_{j+i}) = \mathrm{Im}(\varprojlim_{i\geq 0}(\iota_{j+i} \circ \mu_{j+i}))$, where the right hand side is the ideal of $R_j^{s.\flat}$ generated by $f_j^{s.\flat}$. Thus we obtain the desired implication. Finally, note that by Proposition 3.26(3), we can take a system of elements $\{f_j^{s.\flat} \in R_j^{s.\flat}\}_{j\geq 0}$ satisfying condition (b) and such that $(f_{j+1}^{s.\flat})^p = f_j^{s.\flat}$ $(j \geq 0)$.

(2) We have $I_j^{s.\flat}(R_j^{s.\flat})_{I_j^{s.\flat}\text{-tor}} = (0)$ by Corollary 3.19. Hence, by assertion (1),

$$(R_j^{s.\flat})_{I_0^{s.\flat}\text{-tor}} = (R_j^{s.\flat})_{I_j^{s.\flat}\text{-tor}} = \mathrm{Ann}_{R_j^{s.\flat}}(I_j^{s.\flat}) = \mathrm{Ker}(\varprojlim_{i\geq 0} \mu_{j+i}) = \varprojlim_{i\geq 0} \mathrm{Ann}_{\overline{R}_{j+i}}(\overline{I}_{j+i}).$$

Thus by Lemma 3.48(2), we obtain an isomorphism $(R_j^{s.\flat})_{I_0^{s.\flat}\text{-tor}} \cong (R_j)_{I_0\text{-tor}}$ with the desired property. $\square$

3E. *Relation with perfectoid rings.* In the rest of this paper, for a ring R, we use the following notation. Set the inverse limit

$$R^\flat := \varprojlim\{\cdots \to R/pR \to R/pR \to \cdots \to R/pR\},$$

where each transition map is the Frobenius endomorphism on R/pR. It is called the *tilt* (or *tilting*) of R. Moreover, we denote by $W(R)$ the ring of p-typical Witt vectors over R. If R is p-adically complete and separated, we denote by $\theta_R : W(R^\flat) \to R$ the ring map such that the diagram

$$\begin{tikzcd}
W(R^\flat) \arrow[r, "\theta_R"] \arrow[d] & R \arrow[d] \\
R^\flat \arrow[r] & R/pR
\end{tikzcd} \tag{3-17}$$

(where the vertical maps are induced by reduction modulo p and the bottom map is the first projection) commutes.

Recall the definition of perfectoid rings.

Definition 3.49 [5, Definition 3.5]. A ring S is *perfectoid* if the following conditions hold.

(1) S is ϖ-adically complete and separated for some element $\varpi \in S$ such that ϖ^p divides p.

(2) The Frobenius endomorphism on S/pS is surjective.

(3) The kernel of $\theta_S : W(S^\flat) \to S$ is principal.

We have a connection between perfectoid towers and perfectoid rings. To see this, we use the following characterization of perfectoid rings.

Theorem 3.50 (cf. [17, Corollary 16.3.75]). *Let S be a ring. Then S is a perfectoid ring if and only if S contains an element ϖ with the following properties.*

(1) *ϖ^p divides p, and S is ϖ-adically complete and separated.*

(2) *The ring map $S/\varpi S \to S/\varpi^p S$ induced by the Frobenius endomorphism on $S/\varpi^p S$ is an isomorphism.*

(3) *The multiplicative map*

$$S_{\varpi\text{-tor}} \to S_{\varpi\text{-tor}}\ ;\ s \mapsto s^p \tag{3-18}$$

is bijective.

Proof. The "if" part follows from [17, Corollary 16.3.75].

For the converse, let $\varpi \in S$ be as in Definition 3.49. Such a ϖ clearly has property (1) in the present theorem, and also has property (2) by [5, Lemma 3.10(i)]. To show the remaining part, we set $\widetilde{S} := S/S_{\varpi\text{-tor}}$. By [8, §2.1.3], the diagram of rings:

$$\begin{tikzcd} S \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & (S/\varpi S)_{\mathrm{red}} \arrow[d, "\pi_4"] \\ \widetilde{S} \arrow[r, "\pi_3"'] & (\widetilde{S}/\varpi\widetilde{S})_{\mathrm{red}} \end{tikzcd}$$

(where π_i is the canonical projection map for $i = 1, 2, 3, 4$) is cartesian. Hence $S_{\varpi\text{-tor}}$ ($= \mathrm{Ker}(\pi_1)$) is isomorphic to $\mathrm{Ker}(\pi_4)$ as a (possibly) nonunital ring. Since $(S/\varpi S)_{\mathrm{red}}$ is a perfect $\mathbb{F}_p$-algebra, it admits the Frobenius endomorphism and the inverse Frobenius. Moreover, $\mathrm{Ker}(\pi_4)$ is closed under these operations because $(\widetilde{S}/\varpi\widetilde{S})_{\mathrm{red}}$ is reduced. Consequently, there is a bijection (3-18). Hence ϖ has property (3), as desired. □

Remark 3.51. In view of the above proof, the "only if" part of Theorem 3.50 can be refined as follows. For a perfectoid ring S, an element $\varpi \in S$ such that $p \in \varpi^p S$ and S is ϖ-adically complete and separated satisfies the properties (2) and (3) in Theorem 3.50.

Corollary 3.52. *Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a perfectoid tower arising from some pair (R_0, I_0). Let $\widehat{R_\infty}$ denote the I_1-adic completion of R_∞. Then $\widehat{R_\infty}$ is a perfectoid ring.*

Proof. Since we have $\varinjlim_{i\geq 0} F_{R_i/I_0R_i} = (\varinjlim_{i\geq 0} \bar{t}_i) \circ (\varinjlim_{i\geq 0} F_i)$ and $\varinjlim_{i\geq 0} \bar{t}_i$ is a canonical isomorphism, the Frobenius endomorphism on $\widehat{R_\infty}/I_0\widehat{R_\infty}$ can be identified with $\varinjlim_{i\geq 0} F_i$. Hence one can immediately deduce from the axioms in Definition 3.21 that any generator of $I_1\widehat{R_\infty}$ has the all properties assumed on ϖ in Theorem 3.50. □

In view of Theorem 3.50, one can regard perfectoid rings as a special class of perfectoid towers.

Example 3.53. Let S be a perfectoid ring. Let $\varpi \in S$ be such that $p \in \varpi^p S$ and S is ϖ-adically complete and separated. Set $S_i = S$ and $t_i = \mathrm{id}_S$ for every $i \geq 0$, and $I_0 = \varpi^p S$. Then by Remark 3.51, the tower $(\{S_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower arising from (S, I_0). In particular, $I_0 S_{I_0\text{-tor}} = (0)$, and $F_{S_{I_0\text{-tor}}}$ is bijective.

Moreover, we can treat more general rings in a tower-theoretic way.

Example 3.54 (Zariskian preperfectoid rings). Let R be a ring that contains an element ϖ such that $p \in \varpi^p R$, R is ϖ-adically Zariskian, and R has bounded ϖ-torsion. Assume that the ϖ-adic completion $\hat{R}$ is a perfectoid ring. Set $R_i = R$ and $t_i = \mathrm{id}_R$ for every $i \geq 0$, and $I_0 = \varpi^p R$. Then the tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower arising from (R, I_0). Indeed, axioms (a) and (e) are clear from the assumption. Since $\hat{R}$ is perfectoid and $R/\varpi^p R \cong \hat{R}/\varpi^p \hat{R}$, axioms (b), (c), (d) and (f) hold by Example 3.53. Similarly, axiom (g) holds by Lemma 3.16 (the map $\psi_{\mathrm{tor}} : R_{I_0\text{-tor}} \to (\hat{R})_{I_0\text{-tor}}$ is also an isomorphism of nonunital rings).

Recall that we have two types of tilting operation at present; one is defined for perfectoid rings, and the other is for perfectoid towers. The following result asserts that they are compatible.

Lemma 3.55. *Let $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$ be the tilt of $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ associated to (R_0, I_0). Let $\widehat{R_\infty^{s.\flat}}$ be the $I_0^{s.\flat}$-adic completion of $R_\infty^{s.\flat} := \varinjlim_{i\geq 0} R_i^{s.\flat}$. Let $(I_0R_\infty)^\flat$ be the ideal of $R_\infty^\flat$ that is the inverse image of I_0R_∞ mod pR_∞ via the first projection. Then there exist canonical isomorphisms*

$$R_\infty^\flat \xleftarrow{\cong} \varprojlim_{\mathrm{Frob}} R_\infty^{s.\flat}/I_0^{s.\flat}R_\infty^{s.\flat} \xrightarrow{\cong} \widehat{R_\infty^{s.\flat}}$$

under which $(I_0R_\infty)^\flat \subseteq R_\infty^\flat$ corresponds to $I_0^{s.\flat}\widehat{R_\infty^{s.\flat}} \subseteq \widehat{R_\infty^{s.\flat}}$.

Proof. Note that $R_\infty^{s.\flat}$ is perfect. By Lemma 3.39 and Example 3.32, we obtain the commutative diagram of rings

$$\begin{tikzcd}
\varprojlim_{\mathrm{Frob}} R_\infty/I_0R_\infty \arrow[d] & \varprojlim_{\mathrm{Frob}} R_\infty^{s.\flat}/I_0^{s.\flat}R_\infty^{s.\flat} \arrow[l, "\cong"'] \arrow[r, "\cong"] \arrow[d] & \widehat{R_\infty^{s.\flat}} \arrow[d] \\
R_\infty/I_0R_\infty & R_\infty^{s.\flat}/I_0^{s.\flat}R_\infty^{s.\flat} \arrow[l, "\cong"'] \arrow[r, equal] & R_\infty^{s.\flat}/I_0^{s.\flat}R_\infty^{s.\flat}
\end{tikzcd} \tag{3-19}$$

where the vertical arrows denote the first projection maps. By [5, Lemma 3.2(i)], we can identify $R_\infty^\flat$ with $\varprojlim_{\mathrm{Frob}} R_\infty/I_0R_\infty$, and the ideal $(I_0R_\infty)^\flat \subseteq R_\infty^\flat$ corresponds to the kernel of the leftmost vertical map. Since the kernel of the rightmost vertical map is $I_0^{s.\flat}\widehat{R_\infty^{s.\flat}}$, the assertion follows. □

3F. *Examples: complete local log-regular rings.*

3F1. *Calculation of the tilts.* As an example of tilts of Noetherian perfectoid towers, we calculate them for certain towers of local log-regular rings. Firstly, we review a perfectoid tower constructed in [17].

Construction 3.56. Let $(R, \mathcal{Q}, \alpha)$ be a *complete* local log-regular ring with perfect residue field of characteristic $p > 0$. Assume that $\mathcal{Q}$ is fine, sharp, and saturated (see Remark 2.20). Let $I_\alpha \subseteq R$ be the ideal defined in Definition 2.19. Set $A := R/I_\alpha$. Let $(f_1, \dots, f_r)$ be a sequence of elements of R whose image in A is *maximal* (see Definition A.4). Since the residue field of R is perfect, r is the dimension of A (see the Appendix). For every $i \geq 0$, we consider the ring

$$A_i := A[T_1, \dots, T_r]/(T_1^{p^i} - \overline{f}_1, \dots, T_r^{p^i} - \overline{f}_r),$$

where each $\overline{f}_j$ denotes the image of f_j in A ($j = 1, \dots, r$). Notice that A_i is regular by Theorem A.3. Moreover, we set $\mathcal{Q}^{(i)} := \mathcal{Q}_p^{(i)}$ (see Definition 2.11). Furthermore, we define

$$R_i' := \mathbb{Z}[\mathcal{Q}^{(i)}] \otimes_{\mathbb{Z}[\mathcal{Q}]} R, \ \ R_i'' := R[T_1, \dots, T_r]/(T_1^{p^i} - f_1, \dots, T_r^{p^i} - f_r), \tag{3-20}$$

and

$$R_i := R_i' \otimes_R R_i''. \tag{3-21}$$

Let $t_i : R_i \to R_{i+1}$ be the ring map that is naturally induced by the inclusion map $\iota^{(i)} : \mathcal{Q}^{(i)} \hookrightarrow \mathcal{Q}^{(i+1)}$. Since R_{i+1}'' is a free R_i''-module, t_i is universally injective by Lemma 2.15 and condition (e) in Proposition 2.8(2).

Proposition 3.57. *Keep the notation as in Construction 3.56. Let $\alpha_i : \mathcal{Q}^{(i)} \to R_i$ be the natural map. Then $(R_i, \mathcal{Q}^{(i)}, \alpha_i)$ is a local log-regular ring.*

Proof. We refer the reader to [17, 17.2.5]. □

By construction, we obtain the tower of rings $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ (see Definition 3.1).

Proposition 3.58. *Keep the notation as in Construction 3.56. Then the tower $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ of local log-regular rings defined above is a perfectoid tower arising from $(R, (p))$.*

Proof. We verify (a)–(g) in Definition 3.4 and Definition 3.21. Axiom (a) is trivial. Since t_i is universally injective, axiom (b) follows. Axioms (c) and (d) follow from [17, (17.2.10) and Lemma 17.2.11]. Since R is of residual characteristic p, axiom (e) follows from the locality. Since t_i is injective and R_i is a domain for any $i \geq 0$, axiom (g) holds by Remark 3.22.

Finally, let us check that axiom (f) holds. In the case when $p = 0$, it follows from [17, Theorem 17.2.14(i)]. Otherwise, there exists an element $\varpi \in R_1$ that satisfies $\varpi^p = pu$ for some unit $u \in R_1$ by [17, Theorem 17.2.14(ii)]. Set $I_1 := (\varpi)$. Then axiom (f-1) holds. Axiom (f-2) follows from [17, Theorem 17.2.14(iii)]. Thus the assertion follows. □

For calculating the tilt of the perfectoid tower constructed above, the following lemma is quite useful.

Lemma 3.59. *Keep the notation as in Proposition 3.57. Let k be the residue field of R. Then there exists a family of ring maps $\{\phi_i : C(k)[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!] \to R_i\}_{i\geq 0}$ which is compatible with the log structures of $\{(R_i, \mathcal{Q}^{(i)}, \alpha_i)\}_{i\geq 0}$ such that the following diagram commutes for every $i \geq 0$:*

$$\begin{tikzcd} C(k)[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!] \arrow[r, hook] \arrow[d, two heads, "\phi_i"] & C(k)[\![\mathcal{Q}^{(i+1)} \oplus (\mathbb{N}^r)^{(i+1)}]\!] \arrow[d, two heads, "\phi_{i+1}"] \\ R_i \arrow[r, hook, "t_i"] & R_{i+1} \end{tikzcd} \tag{3-22}$$

(*where the top arrow is the natural inclusion*). *Moreover, there exists an element $\theta \in C(k)[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]$ whose constant term is p such that the kernel of ϕ_i is generated by θ for every $i \geq 0$.*

Proof. First we remark the following. Let k_i be the residue field of R_i. Then by Lemma 3.11(1) and Lemma 3.6(2), the transition maps induce a purely inseparable extension $k \hookrightarrow k_i$. Moreover, this extension is trivial because k is perfect. Therefore, we can identify k_i with k, and the Cohen ring of R_i with $C(k)$.

Next, let us show the existence of a family of ring maps $\{\phi_i\}_{i\geq 0}$ with the desired compatibility. Since $(R_i, \mathcal{Q}^{(i)}, \alpha_i)$ is a complete local log-regular ring, we can take a surjective ring map $\psi_i : C(k)[\![\mathcal{Q}^{(i)} \oplus \mathbb{N}^r]\!] \to R_i$ as in Theorem 2.22; its kernel is generated by an element θ_i whose constant term is p, and the diagram

$$\begin{tikzcd} \mathcal{Q}^{(i)} \arrow[r] \arrow[dr, "\alpha_i"'] & C(k)[\![\mathcal{Q}^{(i)} \oplus \mathbb{N}^r]\!] \arrow[d, two heads, "\psi_i"] \\ & R_i \end{tikzcd}$$

commutes. For $j = 1, \dots, r$, let us denote by f_j^{1/p^i} the image of $T_j \in R[T_1, \dots, T_r]$ in R_i (see (3-20) and (3-21)). Note that the sequence $f_1^{1/p^i}, \dots, f_r^{1/p^i}$ in R_i becomes a regular system of parameters of R_i/I_{α_i} by the reduction modulo I_{α_i} (see [17, 17.2.3] and [17, 17.2.5]). Thus, for the set of the canonical basis $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_r\}$ of $\mathbb{N}^r$, we may assume $\psi_i(e^{\boldsymbol{e}_j}) = f_j^{1/p^i}$ by the construction of ψ_i (see the proof of [34, Chapter III, Theorem 1.11.2]). Hence we can choose $\{\psi_i\}_{i\geq 0}$ so that the diagram:

$$\begin{tikzcd} C(k)[\![\mathcal{Q}^{(i)} \oplus \mathbb{N}^r]\!] \arrow[r, hook] \arrow[d, two heads, "\psi_i"'] & C(k)[\![\mathcal{Q}^{(i+1)} \oplus \mathbb{N}^r]\!] \arrow[d, two heads, "\psi_{i+1}"] \\ R_i \arrow[r, hook, "t_i"] & R_{i+1} \end{tikzcd} \tag{3-23}$$

commutes. Thus it suffices to define $\phi_i : C(k)[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!] \to R_i$ as the composite map of the isomorphism $C(k)[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!] \xrightarrow{\cong} C(k)[\![\mathcal{Q}^{(i)} \oplus \mathbb{N}^r]\!]$ obtained by Lemma 2.12(3) and ψ_i.

Finally, note that the image of $\theta_0 \in \mathrm{Ker}(\psi_0)$ in $C(k)[\![\mathcal{Q}^{(i)} \oplus \mathbb{N}^r]\!]$ is contained in $\mathrm{Ker}(\psi_i)$, and its constant term is still p. Thus, by the latter assertion of Theorem 2.22(2), $\mathrm{Ker}(\psi_i)$ is generated by θ_0. Hence by taking θ_0 as θ, we complete the proof. $\square$

Let us consider the monoids $\mathcal{Q}^{(i)}$ for an integral sharp monoid $\mathcal{Q}$. Since there is the natural inclusion $\iota^{(i)} : \mathcal{Q}^{(i)} \hookrightarrow \mathcal{Q}^{(i+1)}$ for any $i \geq 0$, we obtain a direct system of monoids $(\{\mathcal{Q}^{(i)}\}_{i\geq 0}, \{\iota^{(i)}\}_{i\geq 0})$. Moreover, the p-times map on $\mathcal{Q}^{(i+1)}$ gives a factorization:

$$\begin{tikzcd} \mathcal{Q}^{(i+1)} \arrow[r, "\times p"] \arrow[dr, two heads, "\times p"'] & \mathcal{Q}^{(i+1)} \\ & \mathcal{Q}^{(i)}. \arrow[u, hook, "\iota^{(i)}"'] \end{tikzcd}$$

From this discussion, we define the small tilt of $\{\mathcal{Q}^{(i)}\}_{i\geq 0}$.

Definition 3.60. Let $\mathcal{Q}$ be an integral monoid, and let $(\{\mathcal{Q}^{(i)}\}_{i\geq 0}, \{\iota^{(i)}\}_{i\geq 0})$ be as above. Then for an integer $j \geq 0$, we define the *j-th small tilt of* $(\{\mathcal{Q}^{(i)}\}_{i\geq 0}, \{\iota^{(i)}\}_{i\geq 0})$ as the inverse limit

$$\mathcal{Q}_j^{s.\flat} := \varprojlim\{\cdots \to \mathcal{Q}^{(j+1)} \to \mathcal{Q}^{(j)}\}, \tag{3-24}$$

where the transition map $\mathcal{Q}^{(i+1)} \to \mathcal{Q}^{(i)}$ is the p-times map of monoids.

Now we can derive important properties of the tilt of the perfectoid tower given in Construction 3.56.

Theorem 3.61. *Keep the notation as in Lemma 3.59.*

(1) *The tower $(\{(R_i)_{(p)}^{s.\flat}\}_{i\geq 0}, \{(t_i)_{(p)}^{s.\flat}\}_{i\geq 0})$ is isomorphic to $(\{k[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!]\}_{i\geq 0}, \{u_i\}_{i\geq 0})$, where u_i is the ring map induced by the natural inclusion $\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)} \hookrightarrow \mathcal{Q}^{(i+1)} \oplus (\mathbb{N}^r)^{(i+1)}$.*

(2) *For every $j \geq 0$, there exists a homomorphism of monoids*

$$\alpha_j^{s.\flat} : \mathcal{Q}_j^{s.\flat} \to (R_j)_{(p)}^{s.\flat}$$

such that $((R_j)_{(p)}^{s.\flat}, \mathcal{Q}_j^{s.\flat}, \alpha_j^{s.\flat})$ is a local log-regular ring.

(3) *For every $j \geq 0$, $(t_j)_{(p)}^{s.\flat} : (R_j)_{(p)}^{s.\flat} \to (R_{j+1})_{(p)}^{s.\flat}$ is module-finite and $(R_j)_{(p)}^{s.\flat}$ is F-finite.*

Proof. (1) By Lemma 3.59, each R_i is isomorphic to $C(k)[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!]/(p-f)C(k)[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!]$ where f is an element of $C(k)[\![\mathcal{Q} \oplus \mathbb{N}^r]\!]$ which has no constant term. Set $S_i := k[\![\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)}]\!]$ for any $i \geq 0$ and let $u_i : S_i \hookrightarrow S_{i+1}$ be the inclusion map induced by the natural inclusion $\mathcal{Q}^{(i)} \oplus (\mathbb{N}^r)^{(i)} \hookrightarrow \mathcal{Q}^{(i+1)} \oplus (\mathbb{N}^r)^{(i+1)}$. Then the tower $(\{S_i\}_{i\geq 0}, \{u_i\}_{i\geq 0})$ is a perfect tower. Indeed, each S_i is reduced by Theorem 2.21; moreover, by the perfectness of k and Lemma 2.12(3), the Frobenius endomorphism on S_{i+1} factors through a surjection $G_i : S_{i+1} \to S_i$. In particular, $(\{S_i\}_{i\geq 0}, \{u_i\}_{i\geq 0})$ is a perfectoid tower arising from $(S_0, (0))$ and G_i is the i-th Frobenius projection (cf. Lemma 3.24).

Put $\overline{f} := f \bmod pC(k)[\![\mathcal{Q} \oplus \mathbb{N}^r]\!] \in S_0$. Then each S_i is $\overline{f}$-adically complete and separated by [15, Lemma 2.1.1]. Moreover, the commutative diagram (3-22) yields the commutative squares ($i \geq 0$):

$$\begin{tikzcd} S_{i+1}/\overline{f}S_{i+1} \arrow[r, "\cong"] \arrow[d, "\overline{G}_i"] & R_{i+1}/pR_{i+1} \arrow[d, "F_i"] \\ S_i/\overline{f}S_i \arrow[r, "\cong"] & R_i/pR_i \end{tikzcd}$$

that are compatible with $\{\overline{u}_i : S_i/\overline{f}S_i \to S_{i+1}/\overline{f}S_{i+1}\}_{i\geq 0}$ and $\{\overline{t}_i\}_{i\geq 0}$. Hence by Lemma 3.31, we obtain the isomorphisms

$$(R_j)^{s.\flat}_{(p)} \xleftarrow{\cong} \varprojlim\{\cdots \xrightarrow{\overline{G}_{j+1}} S_{j+1}/\overline{f}S_{j+1} \xrightarrow{\overline{G}_j} S_j/\overline{f}S_j\} \xrightarrow{\cong} S_j \qquad (j \geq 0) \tag{3-25}$$

that are compatible with the transition maps of the towers. Thus the assertion follows.

(2) Considering the inverse limit of the composite maps $\mathcal{Q}^{(j+i)} \xrightarrow{\alpha_{j+i}} R_{j+i} \twoheadrightarrow R_{j+i}/pR_{j+i}$ $(i \geq 0)$, we obtain a homomorphism of monoids $\alpha_j^{s.\flat} : \mathcal{Q}_j^{s.\flat} \to (R_j)^{s.\flat}_{(p)}$. On the other hand, let $\overline{\alpha}_j : \mathcal{Q}^{(j)} \to S_j$ be the natural inclusion. Then, since S_j is canonically isomorphic to $k[\![\mathcal{Q}^{(j)} \oplus \mathbb{N}^r]\!]$, $(S_j, \mathcal{Q}^{(j)}, \overline{\alpha}_j)$ is a local log-regular ring by Theorem 2.22(1). Thus it suffices to show that $((R_j)^{s.\flat}_{(p)}, \mathcal{Q}_j^{s.\flat}, \alpha_j^{s.\flat})$ is isomorphic to $(S_j, \mathcal{Q}^{(j)}, \overline{\alpha}_j)$ as a log ring. Since the transition maps in (3-24) are isomorphisms by Lemma 2.12(3), we obtain the isomorphisms of monoids

$$\mathcal{Q}_j^{s.\flat} \xleftarrow{\mathrm{id}_{\mathcal{Q}_j^{s.\flat}}} \mathcal{Q}_j^{s.\flat} \xrightarrow{\cong} \mathcal{Q}^{(j)} \quad (j \geq 0). \tag{3-26}$$

Then one can connect (3-26) to (3-25) to construct a commutative diagram using $\alpha_j^{s.\flat}$ and $\overline{\alpha}_j$. Hence the assertion follows.

(3) By Lemma 2.14(1), $t_j : R_j \to R_{j+1}$ is module-finite. Hence by Proposition 3.42(1), $(t_j)^{s.\flat}_{(p)} : (R_j)^{s.\flat}_{(p)} \to (R_{j+1})^{s.\flat}_{(p)}$ is also module-finite. Finally let us show that $(R_j)^{s.\flat}_{(p)}$ is F-finite. By assertion (2), $(R_j)^{s.\flat}_{(p)}$ is a complete Noetherian local ring, and the residue field is F-finite because it is perfect. Thus the assertion follows from [29, Theorem 8.4]. □

Example 3.62. (1) A tower of regular local rings which is treated in [7] and [8] is a perfectoid tower in our sense. Let $(R, \mathfrak{m}, k)$ be a d-dimensional regular local ring whose residue field k is perfect and let $x_1, \ldots, x_d$ be a regular sequence of parameters. Let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_d$ be the canonical basis of $\mathbb{N}^d$. Then $(R, \mathbb{N}^d, \alpha)$ is a local log-regular ring where $\alpha : \mathbb{N}^d \to R$ is a homomorphism of monoids which maps $\boldsymbol{e}_i$ to x_i. Furthermore, assume that R is $\mathfrak{m}$-adically complete. Then, by Cohen's structure theorem, R is isomorphic to

$$W(k)[\![x_1, \ldots, x_d]\!]/(p - f)$$

where $f = x_1$ or $f \in (p, x_1, \ldots, x_d)^2$ (the former case is called *unramified*, and the latter *ramified*). Let us construct a perfectoid tower arising from $(R, (p))$ along Construction 3.56. Since k is perfect, $\boldsymbol{\Omega}_k$ is zero by the short exact sequences (A-4) and the definition of itself. This implies that the image of the empty subset of R in k forms a maximal sequence. Hence R_i'' in Construction 3.56 is equal to R. Moreover, $(\mathbb{N}^d)^{(i)}$ is generated by $\frac{1}{p^i}\boldsymbol{e}_1, \ldots, \frac{1}{p^i}\boldsymbol{e}_d$. Applying Construction 3.56, we obtain

$$\begin{aligned} R_i = R_i' = \mathbb{Z}[(\mathbb{N}^d)^{(i)}] \otimes_{\mathbb{Z}[\mathbb{N}^d]} R &\cong R[T_1, \ldots, T_d]/(T_1^{p^i} - x_1, \ldots, T_d^{p^i} - x_d) \\ &\cong W(k)[\![x_1^{1/p^i}, \ldots, x_d^{1/p^i}]\!]/(p - f). \end{aligned}$$

Set the natural injection $t_i : R_i \to R_{i+1}$ for any $i \geq 0$. Then, by Proposition 3.58, $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower arising from $(R, (p))$. By Theorem 3.61, its tilt $(\{(R_i)^{s.\flat}_{(p)}\}_{i\geq 0}, \{(t_i)^{s.\flat}_{(p)}\}_{i\geq 0})$ is isomorphic to the tower $k[\![\mathbb{N}^d]\!] \hookrightarrow k[\![(\mathbb{N}^d)^{(1)}]\!] \hookrightarrow k[\![(\mathbb{N}^d)^{(2)}]\!] \hookrightarrow \cdots$, which can be written as

$$k[\![x_1, \dots, x_d]\!] \hookrightarrow k[\![x_1^{1/p}, \dots, x_d^{1/p}]\!] \hookrightarrow k[\![x_1^{1/p^2}, \dots, x_d^{1/p^2}]\!] \hookrightarrow \cdots .$$

(2) Consider the surjection

$$\begin{aligned} S := W(k)[\![x, y, z, w]\!]/(xy - zw) \twoheadrightarrow R &:= W(k)[\![x, y, z, w]\!]/(xy - zw, p - w) \\ &= W(k)[\![x, y, z]\!]/(xy - pz). \end{aligned}$$

where k is a perfect field. Let $\mathcal{Q} \subseteq \mathbb{N}^4$ be a saturated submonoid generated by $(1,1,0,0)$, $(0,0,1,1)$, $(1,0,0,1)$ and $(0,1,1,0)$. Then S admits a homomorphism of monoids $\alpha_S : \mathcal{Q} \to S$ by letting $(1,1,0,0) \mapsto x$, $(0,0,1,1) \mapsto y$, $(1,0,0,1) \mapsto z$ and $(0,1,1,0) \mapsto w$. With this, $(S, \mathcal{Q}, \alpha_S)$ is a local log-regular ring. The composite map $\alpha_R : \mathcal{Q} \to S \to R$ makes R into a local log ring. Indeed, we can write $R \cong W(k)[\![\mathcal{Q}]\!]/(p - e^{(0,1,1,0)})$; hence $(R, \mathcal{Q}, \alpha_R)$ is log-regular by Theorem 2.22.

Next, note that $R/I_{\alpha_R} \cong k$. Then, for the same reason in (1), R_i'' is equal to R. Moreover, $\mathcal{Q}^{(i)}$ is generated by

$$\Big(\frac{1}{p^i}, \frac{1}{p^i}, 0, 0\Big),\ \Big(0, 0, \frac{1}{p^i}, \frac{1}{p^i}\Big),\ \Big(\frac{1}{p^i}, 0, 0\frac{1}{p^i}\Big),\ \Big(0, \frac{1}{p^i}, \frac{1}{p^i}, 0\Big).$$

Thus, applying Construction 3.56, we obtain

$$\begin{aligned} R_i &= R[\![\mathcal{Q}^{(i)}]\!] \\ &\cong W(k)[\![\mathcal{Q}^{(i)}]\!]/(p - e^{(0,1,1,0)}) \\ &\cong W(k)[\![x^{1/p^i}, y^{1/p^i}, z^{1/p^i}, w^{1/p^i}]\!]/(x^{1/p^i} y^{1/p^i} - z^{1/p^i} w^{1/p^i}, p - w). \end{aligned}$$

Set a natural injection $t_i : R_i \to R_{i+1}$. Then, by Proposition 3.58, $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ is a perfectoid tower arising from $(R, (p))$. Hence

$$R_\infty = \varinjlim_{i\geq 0} R_i \cong \bigcup_{i\geq 0} W(k)[\![x^{1/p^i}, y^{1/p^i}, z^{1/p^i}, w^{1/p^i}]\!]/(x^{1/p^i} y^{1/p^i} - z^{1/p^i} w^{1/p^i}, p - w),$$

and its p-adic completion is perfectoid. One can calculate the tilt $\big(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0}\big)$ to be $k[\![\mathcal{Q}]\!] \hookrightarrow k[\![\mathcal{Q}^{(1)}]\!] \hookrightarrow k[\![\mathcal{Q}^{(2)}]\!] \hookrightarrow \cdots$ by Theorem 3.61, or, more explicitly,

$$k[\![x, y, z, w]\!]/(xy - zw) \hookrightarrow k[\![x^{1/p}, y^{1/p}, z^{1/p}, w^{1/p}]\!]/(x^{1/p} y^{1/p} - z^{1/p} w^{1/p}) \hookrightarrow \cdots .$$

3F2. *Towers of split maps and sousperfectoid rings.* Recall that Hansen and Kedlaya introduced a new class of topological rings that guarantees sheafiness on the associated adic spectra (see [21, Definition 7.1]).

Definition 3.63. Let A be a complete and separated Tate ring such that a prime $p \in A$ is topologically nilpotent. We say that A is *sousperfectoid* if there exists a perfectoid ring B in the sense of Fontaine (see [21, Definition 2.13]) with a continuous A-linear map $f : A \to B$ that splits in the category of topological A-modules. That is, there is a continuous A-linear map $\sigma : B \to A$ such that $\sigma \circ f = \mathrm{id}_A$.

Let us show that a perfectoid tower consisting of split maps induces sousperfectoid rings. In view of Theorem 2.29, one can apply this result to the towers discussed above. See [33] for detailed studies on algebraic aspects of Tate rings.

Proposition 3.64. *Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a perfectoid tower arising from some pair $(R, (f_0))$. Assume that f_0 is regular, R is f_0-adically complete and separated, and t_i splits as an R_i-linear map for every $i \geq 0$. We equip $R[\frac{1}{f_0}]$ with the linear topology in such a way that $\{f_0^n R\}_{n\geq 1}$ defines a fundamental system of open neighborhoods at $0 \in R[\frac{1}{f_0}]$. Then $R[\frac{1}{f_0}]$ is a sousperfectoid Tate ring, and hence stably uniform.*

In order to prove this, we need the following lemma.

Lemma 3.65. *Keep the notations and assumptions as in Proposition 3.64. Then the natural map $R_0 \to \varinjlim_{i\geq 0} R_i$ splits as an R_0-linear map.*

Proof. We use the fact that each $t_i : R_i \to R_{i+1}$ splits as an R_i-linear map by assumption. This implies that the short exact sequence of R-modules

$$0 \to R_0 \to R_i \to R_i/R \to 0$$

splits for any $i \geq 0$. It induces a commutative diagram of R-modules

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{Hom}_{R_0}(R_{i+1}/R_0, R_0) \arrow[r] \arrow[d, "\alpha_i"] & \mathrm{Hom}_{R_0}(R_{i+1}, R_0) \arrow[r] \arrow[d, "\beta_i"] & \mathrm{Hom}_{R_0}(R_0, R_0) \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & \mathrm{Hom}_{R_0}(R_i/R_0, R_0) \arrow[r] & \mathrm{Hom}_{R_0}(R_i, R_0) \arrow[r] & \mathrm{Hom}_{R_0}(R_0, R_0) \arrow[r] & 0 \end{tikzcd}$$

where each horizontal sequence is split exact and each vertical map forms an inverse system induced by $t_i : R_i \to R_{i+1}$. Thus β_i is surjective and it follows from the snake lemma that α_i is surjective as well. By taking inverse limits, we obtain the short exact sequence

$$0 \to \varprojlim_{i\geq 0} \mathrm{Hom}_{R_0}(R_i/R_0, R_0) \to \varprojlim_{i\geq 0} \mathrm{Hom}_{R_0}(R_i, R_0) \xrightarrow{h} \mathrm{Hom}_{R_0}(R_0, R_0) \to 0.$$

It follows from [36, Lemma 4.1] that h is the canonical surjection $\mathrm{Hom}_{R_0}(R_\infty, R_0) \twoheadrightarrow \mathrm{Hom}_{R_0}(R_0, R_0)$. Then choosing an inverse image of $\mathrm{id}_{R_0} \in \mathrm{Hom}_{R_0}(R_0, R_0)$ gives a splitting of $R_0 \to R_\infty$. □

Proof of Proposition 3.64. We have constructed an infinite extension $R \to R_\infty$ such that if $\widehat{R}_\infty$ is the f_0-adic completion, then the associated Tate ring $\widehat{R}_\infty[\frac{1}{f_0}]$ is a perfectoid ring in the sense of Fontaine by Corollary 3.52 and [5, Lemma 3.21].

By Lemma 2.28 and Lemma 3.65, it follows that the map $R[\frac{1}{f_0}] \to \widehat{R}_\infty[\frac{1}{f_0}]$ splits in the category of topological $R[\frac{1}{f_0}]$-modules (notice that R is f_0-adically complete and separated). Thus, $R[\frac{1}{f_0}]$ is a sousperfectoid Tate ring. The combination of [21, Corollary 8.10], [21, Proposition 11.3] and [21, Lemma 11.9] allows us to conclude that $R[\frac{1}{f_0}]$ is stably uniform. □

As a corollary, one can obtain the stable uniformity for complete local log-regular rings (see also Construction 3.56 and Theorem 2.29).

Corollary 3.66. *Let $(R, \mathcal{Q}, \alpha)$ is a complete local log-regular ring of mixed characteristic with perfect residue field. We equip $R[\frac{1}{p}]$ with the structure of a complete and separated Tate ring in such a way that $\{p^n R\}_{n\geq 1}$ defines a fundamental system of open neighborhoods at $0 \in R[\frac{1}{p}]$. Then $R[\frac{1}{p}]$ is stably uniform.*

4. Applications to étale cohomology of Noetherian rings

In this section, we establish several results on étale cohomology of Noetherian rings, as applications of the theory of perfectoid towers developed in Section 3. In Section 4A, for a ring that admits a certain type of perfectoid tower, we prove that finiteness of étale cohomology groups on the positive characteristic side carries over to the mixed characteristic side (Proposition 4.7). In Section 4B, we apply this result to a problem on divisor class groups of log-regular rings.

We prepare some notation. Let X be a scheme and let $X_{\text{ét}}$ denote the category of schemes that are étale over X, and for any étale X-scheme Y, we specify the covering $\{Y_i \to Y\}_{i\in I}$ so that Y_i is étale over Y and the family $\{Y_i\}_{i\in I}$ covers surjectively Y. For an abelian sheaf $\mathcal{F}$ on $X_{\text{ét}}$, we denote by $H^i(X_{\text{ét}}, \mathcal{F})$ the value of the i-th derived functor of $U \in X_{\text{ét}} \mapsto \Gamma(U, \mathcal{F})$. For the most part of applications, we consider *torsion* sheaves, such as $\mathbb{Z}/n\mathbb{Z}$ and μ_n for $n \in \mathbb{N}$. However, for the multiplicative group scheme $\mathbb{G}_m$, we often use the following isomorphism:

$$H^1(X_{\text{ét}}, \mathbb{G}_m) \cong \operatorname{Pic}(X).$$

For the basics on étale cohomology, we often use [12] or [31] as references.

4A. ***Tilting étale cohomology groups.*** Let A be a ring with an ideal J, let $\hat{A}$ be the J-adic completion of A, and let $U \subseteq \operatorname{Spec}(A)$ be an open subset. We define the J-adic completion of U to be the open subset $\hat{U} \subseteq \operatorname{Spec}(\hat{A})$, which is the inverse image of U via $\operatorname{Spec}(\hat{A}) \to \operatorname{Spec}(A)$. We will use the following result for deriving results on the behavior of étale cohomology under the tilting operation as well as some interesting results on the divisor class groups of Noetherian normal domains (see Proposition 4.10 and Proposition 4.11).

Theorem 4.1 (Fujiwara and Gabber). *Let (A, J) be a Henselian pair with $X := \operatorname{Spec}(A)$ and let $\hat{A}$ be the J-adic completion of A.*

(1) *For any abelian torsion sheaf $\mathscr{F}$ on $X_{\text{ét}}$, we have $\boldsymbol{R}\Gamma(\operatorname{Spec}(A)_{\text{ét}}, \mathscr{F}) \simeq \boldsymbol{R}\Gamma(\operatorname{Spec}(A/J)_{\text{ét}}, \mathscr{F}|_{\operatorname{Spec}(A/J)})$.*

(2) *Assume that J is finitely generated. Then for any abelian torsion sheaf $\mathscr{F}$ on $X_{\text{ét}}$ and any open subset $U \subseteq X$ such that $X \setminus V(J) \subseteq U$, we have $\boldsymbol{R}\Gamma(U_{\text{ét}}, \mathscr{F}) \simeq \boldsymbol{R}\Gamma(\hat{U}_{\text{ét}}, \mathscr{F})$.*

Proof. The first statement is known as *Affine analog of proper base change* in [16], while the second one is known as *Formal base change theorem* which is [13, Theorem 7.1.1] in the Noetherian case, and [24, XX, 4.4] in the non-Noetherian case. □

We will need the tilting invariance of (local) étale cohomology from [8, Theorem 2.2.7]. To state the theorem and establish a variant of it, we give some notations.

Definition 4.2. Let (A, I) and (B, J) be pairs such that there exists a ring isomorphism $\Phi : A/I \xrightarrow{\cong} B/J$. Then for any open subset $U \subseteq \mathrm{Spec}(B)$ containing $\mathrm{Spec}(B) \setminus V(J)$, we define an open subset $F_{A,\Phi}(U) \subseteq \mathrm{Spec}(A)$ as the complement of the closed subset $\mathrm{Spec}(\Phi)(\mathrm{Spec}(B) \setminus U) \subseteq \mathrm{Spec}(A)$.

One can define small tilts of Zariski-open subsets.

Definition 4.3. Let $(\{R_i\}_{i\geq 0}, \{t_i\}_{i\geq 0})$ be a perfectoid tower arising from some pair (R, I_0), and let $(\{R_i^{s.\flat}\}_{i\geq 0}, \{t_i^{s.\flat}\}_{i\geq 0})$ be the tilt associated to (R, I_0). Recall that we then have an isomorphism of rings $\overline{\Phi_0^{(i)}} : R_i^{s.\flat}/I_0^{s.\flat} R_i^{s.\flat} \xrightarrow{\cong} R_i/I_0 R_i$ for every $i \geq 0$. For every $i \geq 0$ and every open subset $U \subseteq \mathrm{Spec}(R_i)$ containing $\mathrm{Spec}(R_i) \setminus V(I_0 R_i)$, we define

$$U_{I_0}^{s.\flat} := F_{R_i^{s.\flat}, \overline{\Phi_0^{(i)}}}(U).$$

We also denote $U_{I_0}^{s.\flat}$ by $U^{s.\flat}$ as an abbreviated form.

Note that by the compatibility described in Lemma 3.39, the operation $U \rightsquigarrow U^{s.\flat}$ is compatible with the base extension along the transition maps of a perfectoid tower.

Let us give some examples of $U^{s.\flat}$.

Example 4.4 (punctured spectra of regular local rings). Keep the notation as in Example 3.62(1). In this situation, the isomorphism $\overline{\Phi_0^{(0)}} : R_0^{s.\flat}/I_0^{s.\flat} \xrightarrow{\cong} R_0/I_0$ in Definition 4.3 can be written as

$$k[\![x_1, \dots, x_d]\!]/(p^{s.\flat}) \xrightarrow{\cong} R/pR, \tag{4-1}$$

where $p^{s.\flat} \in k[\![x_1, \dots, x_d]\!]$ is some element. Set $U := \mathrm{Spec}(R) \setminus V(\mathfrak{m})$. Then, since the maximal ideal $\overline{\mathfrak{m}} \subseteq R/pR$ corresponds to the (unique) maximal ideal of $k[\![x_1, \dots, x_d]\!]/(p^{s.\flat})$, we have

$$U^{s.\flat} \cong \mathrm{Spec}(k[\![x_1, \dots, x_d]\!]) \setminus V((x_1, \dots, x_d)).$$

Example 4.5 (tilting for preperfectoid rings). Keep the notation as in Example 3.54. Then by Lemma 3.55, $\overline{\Phi_0^{(0)}} : R_0^{s.\flat}/I_0^{s.\flat} \xrightarrow{\cong} R_0/I_0$ is identified with the isomorphism

$$\overline{\theta}_{\hat{R}} : (\hat{R})^\flat/I_0^\flat(\hat{R})^\flat \xrightarrow{\cong} \hat{R}/I_0\hat{R} \tag{4-2}$$

which is induced by the bottom map in the diagram (3-17). In this case, we denote $F_{R^\flat, \overline{\Phi_0^{(0)}}}(U)$ by $U^\flat$ in distinction from $U^{s.\flat}$.

The comparison theorem we need, due to Česnavičius and Scholze, is stated as follows.

Theorem 4.6 [8, Theorem 2.2.7]. *Let A be a ϖ-adically Henselian ring with bounded ϖ-torsion for an element $\varpi \in A$ such that $p \in \varpi^p A$. Assume that the ϖ-adic completion of A is perfectoid. Let $U \subseteq \mathrm{Spec}(A)$ be a Zariski-open subset such that $\mathrm{Spec}(A) \setminus V(\varpi A) \subseteq U$, and let $U^\flat \subseteq \mathrm{Spec}(A^\flat)$ be its tilt (see Example 4.5).*

(1) *For every torsion abelian group G, we have $\boldsymbol{R}\Gamma(U_{\acute{e}t}, G) \cong \boldsymbol{R}\Gamma(U^\flat_{\acute{e}t}, G)$ in a functorial manner with respect to A, U, and G.*

(2) *Let Z be the complement of $U \subseteq \operatorname{Spec}(A)$. Then for a torsion abelian group G, we have*

$$\boldsymbol{R}\Gamma_Z(\operatorname{Spec}(A)_{\text{ét}}, G) \cong \boldsymbol{R}\Gamma_Z(\operatorname{Spec}(A^\flat)_{\text{ét}}, G).$$

Now we come to the main result on tilting étale cohomology groups. Recall that we have fixed a prime $p > 0$.

Proposition 4.7. *Let $(\{R_j\}_{j\geq 0}, \{t_j\}_{j\geq 0})$ be a perfectoid tower arising from some pair (R, I_0). Suppose that R_j is I_0-adically Henselian for every $j \geq 0$. Let ℓ be a prime different from p. Suppose further that for every $j \geq 0$, $t_j : R_j \to R_{j+1}$ is a module-finite extension of Noetherian normal domains whose generic extension is of p-power degree.*[13] *Fix a Zariski-open subset $U \subseteq \operatorname{Spec}(R)$ such that $\operatorname{Spec}(R) \setminus V(pR) \subseteq U$ and the corresponding open subset $U^{s.\flat} \subseteq \operatorname{Spec}(R^{s.\flat})$ (cf. Definition 4.3). Then, for any fixed $i, n \geq 0$ such that $|H^i(U^{s.\flat}_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})| < \infty$, one has*

$$|H^i(U_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})| \leq |H^i(U^{s.\flat}_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})|.$$

In particular, if $H^i(U^{s.\flat}_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) = 0$, then $H^i(U_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) = 0$.

Proof. Since each R_j is a p-adically Henselian normal domain, so is $R_\infty = \varinjlim_{j\geq 0} R_j$. Moreover, every prime ℓ different from p is a unit in R_j and R_∞. Attached to the tower $(\{R_j\}_{j\geq 0}, \{t_j\}_{j\geq 0})$, we get a tower of finite (not necessarily flat) maps of normal schemes:

$$U = U_0 \leftarrow \cdots \leftarrow U_j \leftarrow U_{j+1} \leftarrow \cdots . \tag{4-3}$$

More precisely, let $h_j : \operatorname{Spec}(R_{j+1}) \to \operatorname{Spec}(R_j)$ be the associated scheme map. Then the open set U_{j+1} is defined as the inverse image $h_j^{-1}(U_j)$, thus defining the map $U_{j+1} \to U_j$ in the tower (4-3). Since h_j is a finite morphism of normal schemes, Lemma 3.4 of [3] applies to yield a well-defined trace map $\operatorname{Tr} : h_{j*}h_j^*\mathbb{Z}/\ell^n\mathbb{Z} \to \mathbb{Z}/\ell^n\mathbb{Z}$ such that

$$\mathbb{Z}/\ell^n\mathbb{Z} \xrightarrow{h_j^*} h_{j*}h_j^*\mathbb{Z}/\ell^n\mathbb{Z} \xrightarrow{\operatorname{Tr}} \mathbb{Z}/\ell^n\mathbb{Z} \tag{4-4}$$

is multiplication by the generic degree of h_j (=p-power order). Then this is bijective, as the multiplication map by p on $\mathbb{Z}/\ell^n\mathbb{Z}$ is bijective. We have the natural map: $H^i(U_{j,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \to H^i(U_{j+1,\text{ét}}, h_j^*\mathbb{Z}/\ell^n\mathbb{Z})$. Since h_j is affine, the Leray spectral sequence gives $H^i(U_{j+1,\text{ét}}, h_j^*\mathbb{Z}/\ell^n\mathbb{Z}) \cong H^i(U_{j,\text{ét}}, h_{j*}h_j^*\mathbb{Z}/\ell^n\mathbb{Z})$. Composing these maps, the composite map (4-4) induces

$$H^i(U_{j,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \to H^i(U_{j+1,\text{ét}}, h_j^*\mathbb{Z}/\ell^n\mathbb{Z}) \xrightarrow{\cong} H^i(U_{j,\text{ét}}, h_{j*}h_j^*\mathbb{Z}/\ell^n\mathbb{Z}) \xrightarrow{\operatorname{Tr}} H^i(U_{j,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})$$

and the composition is bijective. Since $h_j^*\mathbb{Z}/\ell^n\mathbb{Z} \cong \mathbb{Z}/\ell^n\mathbb{Z}$, we get an injection

$$H^i(U_{j,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \hookrightarrow H^i(U_{j+1,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}). \tag{4-5}$$

[13] The existence of such towers is quite essential for applications to étale cohomology, because the extension degree of each $R_j \to R_{j+1}$ is controlled in such a way that the p-adic completion of its colimit is a perfectoid ring.

Set $U_\infty = \varprojlim_j U_j$. Since each morphism $U_{j+1} \to U_j$ is affine, by using (4-5) and [38, Tag 09YQ], we have

$$H^i(U_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \hookrightarrow \varinjlim_j H^i(U_{j,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \cong H^i(U_{\infty,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}).$$

Thus, it suffices to show that $|H^i(U_{\infty,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})| \le |H^i(U^{s.\flat}_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})|$. Hence by tilting étale cohomology using Theorem 4.6, we are reduced to showing

$$|H^i(U^{\flat}_{\infty,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})| \le |H^i(U^{s.\flat}_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z})|, \tag{4-6}$$

where $U^{\flat}_\infty$ is the open subset of $\operatorname{Spec}(R^{\flat}_\infty)$ that corresponds to $U_\infty \subseteq \operatorname{Spec}(R_\infty)$ in view of Example 4.5. On the other hand, considering the tilt of $(\{R_j\}_{j\ge0}, \{t_j\}_{j\ge0})$ associated to (R_0, I_0), we have a perfect $\mathbb{F}_p$-tower $(\{R^{s.\flat}_j\}_{j\ge0}, \{t^{s.\flat}_j\}_{j\ge0})$. Note that each $R^{s.\flat}_j$ is $I^{s.\flat}_0$-adically Henselian Noetherian ring[14] by Lemma 3.36 and Proposition 3.42(2), and $t^{s.\flat}_j$ is module-finite by Proposition 3.42(1). Considering the small tilts of the Zariski-open subsets appearing in (4-3) (see Definition 4.3), we get a tower of finite maps:

$$U^{s.\flat} = U^{s.\flat}_0 \leftarrow \cdots \leftarrow U^{s.\flat}_j \leftarrow U^{s.\flat}_{j+1} \leftarrow \cdots .$$

So let $U^{s.\flat}_\infty$ be the inverse image of $U^{s.\flat}$ under $\operatorname{Spec}(R^{s.\flat}_\infty) \to \operatorname{Spec}(R^{s.\flat})$. Since $U^{s.\flat}_\infty \to U^{s.\flat}$ is a universal homeomorphism, the preservation of the small étale sites [38, Tag 03SI] gives an isomorphism:

$$H^i(U^{s.\flat}_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \cong H^i(U^{s.\flat}_{\infty,\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}). \tag{4-7}$$

Now the combination of Lemma 3.55 and Theorem 4.1(2) together with the assumption finishes the proof of the theorem. □

Remark 4.8. One can formulate and prove the version of Proposition 4.7 for the étale cohomology with support in a closed subscheme of $\operatorname{Spec}(R)$, using Theorem 4.6. Then the resulting assertion gives a generalization of Česnavičius-Scholze's argument in [7, Theorem 3.1.3] which is a key part of their proof for the absolute cohomological purity theorem. One of the advantages of Proposition 4.7 is that it can be used to answer some cohomological questions on possibly singular Noetherian schemes (e.g. log-regular schemes) in mixed characteristic.

4B. ***Tilting the divisor class groups of local log-regular rings.*** We need a lemma of Grothendieck on the relationship between the divisor class group and the Picard group via direct limit. Its proof is found in [19, Proposition (21.6.12)] or [20, XI Proposition 3.7.1].

Lemma 4.9. *Let X be an integral Noetherian normal scheme, and let $\{U_i\}_{i\in I}$ be a family of open subsets of X. Consider the following conditions.*

(1) *$\{U_i\}_{i\in I}$ forms a filter base. In particular, one can define a partial order on I so that it is a directed set and $\{U_i\}_{i\in I}$ together with the inclusion maps forms an inverse system.*

14 It is not obvious whether $R^{s.\flat}_j$ is normal. However, the normality was used only in the trace argument and we do not need it in the following argument.

(2) *Let* $V_i := X \setminus U_i$ *for any* $i \in I$. *Then* $\operatorname{codim}_X(V_i) \geq 2$.

(3) *For any* $x \in \bigcap_{i \in I} U_i$, *the local ring* $\mathcal{O}_{X,x}$ *is factorial.*

If $\{U_i\}_{i \in I}$ *satisfies condition* (2), *then the natural map* $\operatorname{Pic}(U_i) \to \operatorname{Cl}(X)$ *is injective for any* $i \in I$. *If* $\{U_i\}_{i \in I}$ *satisfies conditions* (1), (2) *and* (3), *then* $\varinjlim_{i \in I} \operatorname{Pic}(U_i) \cong \operatorname{Cl}(X)$. *Thus, if* $U \subseteq X$ *is any open subset that is locally factorial with* $\operatorname{codim}_X(X \setminus U) \geq 2$, *then* $\operatorname{Pic}(U) \cong \operatorname{Cl}(X)$.

Next we establish two results on the torsion part of the divisor class group of a (Noetherian) normal domain; they are examples of numerous applications of Theorem 4.1 of independent interest.

Proposition 4.10. *Let* $(R, \mathfrak{m}, k)$ *be a strictly Henselian Noetherian local normal* $\mathbb{F}_p$*-domain of dimension* ≥ 2, *let* $X := \operatorname{Spec}(R)$ *and fix an ideal* $J \subseteq \mathfrak{m}$. *Let* $\{U_i\}_{i \in I}$ *be any family of open subsets of* X *satisfying* (1), (2) *and* (3) *as in the hypothesis of Lemma 4.9 and let* U_i^∞ *be the* $\mathbb{F}_p$*-scheme which is the perfection of* U_i.

(1) *For any prime* $\ell \neq p$,

$$\operatorname{Cl}(X)[\ell^n] \cong \varinjlim_{i \in I} H^1\big((U_i^\infty)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big).$$

(2) *Let* $\widehat{R}^{1/p^\infty}$ *denote the* J*-adic completion of* R^{1/p^∞}. *If each* U_i *has the property that* $X \setminus V(J) \subseteq U_i$, *then for any prime* $\ell \neq p$,

$$\operatorname{Cl}(X)[\ell^n] \cong \varinjlim_{i \in I} H^1\big((\widehat{U}_i^\infty)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big),$$

where $\widehat{U}_i^\infty$ *is inverse image of* U_i^∞ *via the scheme map* $\operatorname{Spec}(\widehat{R}^{1/p^\infty}) \to \operatorname{Spec}(R^{1/p^\infty})$.

Proof. Let us begin with a remark on the direct limit of étale cohomology groups. For the transition morphism $g : U_i^\infty \to U_j^\infty$ which is affine, there is a functorial map $H^1\big((U_j^\infty)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big) \to H^1\big((U_i^\infty)_{\text{ét}}, g^*(\mathbb{Z}/\ell^n\mathbb{Z})\big) \cong H^1\big((U_i^\infty)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big)$, which defines the direct system of cohomology groups.

(1) We prove that *for any* $n \in \mathbb{N}$, *there is an injection of abelian groups*

$$H^1(U_{et}, \mathbb{Z}/\ell^n\mathbb{Z}) \cong \operatorname{Pic}(U)[\ell^n] \subseteq \operatorname{Cl}(X)[\ell^n],$$

where $U \subseteq X$ *is an open subset whose complement is of codimension* ≥ 2. Indeed, consider the Kummer exact sequence

$$0 \to \mathbb{Z}/\ell^n\mathbb{Z} \cong \mu_{\ell^n} \to \mathbb{G}_m \xrightarrow{(\ )^{\ell^n}} \mathbb{G}_m \to 0,$$

where the identification of étale sheaves $\mu_{\ell^n} \cong \mathbb{Z}/\ell^n\mathbb{Z}$ follows from the fact that R is strict Henselian (one simply sends $1 \in \mathbb{Z}/\ell^n\mathbb{Z}$ to the primitive ℓ^n-th root of unity in R). Let $U \subseteq X$ be an open subset with its complement $V = X \setminus U$ having codimension ≥ 2. Then we have an exact sequence (see [31, Chapter III, Proposition 4.9])

$$\Gamma(U_{\text{ét}}, \mathbb{G}_m) \xrightarrow{(\ )^{\ell^n}} \Gamma(U_{\text{ét}}, \mathbb{G}_m) \to H^1(U_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \to \operatorname{Pic}(U) \xrightarrow{(\ )^{\ell^n}} \operatorname{Pic}(U).$$

Since R is strict local and $\ell \neq p$, Hensel's lemma yields that $R^\times = (R^\times)^{\ell^n}$. Since $\operatorname{codim}_X(V) \geq 2$ and X

is normal, we have $\Gamma(U_{\text{ét}}, \mathbb{G}_m) = R^\times$. Thus, $H^1(U_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \cong \mathrm{Pic}(U)[\ell^n]$. Note that $\mathrm{Pic}(U) \hookrightarrow \mathrm{Cl}(U)$ restricts to $\mathrm{Pic}(U)[\ell^n] \hookrightarrow \mathrm{Cl}(U)[\ell^n]$. Moreover, the natural homomorphism $\mathrm{Cl}(X) \to \mathrm{Cl}(U)$ is an isomorphism, thanks to $\mathrm{codim}_X(V) \geq 2$. Hence $H^1(U_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \cong \mathrm{Pic}(U)[\ell^n] \subseteq \mathrm{Cl}(X)[\ell^n]$, which proves the claim.

Since R is normal, the regular locus has complement with codimension ≥ 2. Using this fact, we can apply Lemma 4.9 to get an isomorphism $\mathrm{Cl}(X)[\ell^n] \cong \varinjlim_{i\in I} H^1\big((U_i)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big)$. By étale invariance of cohomology under taking perfection of $\mathbb{F}_p$-schemes [38, Tag 03SI], we get

$$\mathrm{Cl}(X)[\ell^n] \cong \varinjlim_{i\in I} H^1\big((U_i)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big) \cong \varinjlim_{i\in I} H^1\big((U_i^\infty)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big),$$

as desired.

(2) Since R is Henselian along $\mathfrak{m}$ and $J \subseteq \mathfrak{m}$, it is Henselian along J by [38, Tag 0DYD]. The perfect closure of R still preserves the Henselian property along J. Theorem 4.1 yields

$$H^1\big((U_i^\infty)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big) \cong H^1\big((\widehat{U}_i^\infty)_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}\big)$$

and the conclusion follows from (1). $\square$

Proposition 4.11. *Let A be a Noetherian ring with a regular element $t \in A$ such that A is t-adically Henselian and $A \to A/tA$ is the natural surjection between locally factorial domains. Pick an integer $n > 0$ that is invertible on A. Then if $\mathrm{Cl}(A)$ has no torsion element of order n, the same holds for $\mathrm{Cl}(A/tA)$. If moreover A is a $\mathbb{Q}$-algebra and $\mathrm{Cl}(A)$ is torsion-free, then so is $\mathrm{Cl}(A/tA)$.*

Proof. The Kummer exact sequence $0 \to \mu_n \to \mathbb{G}_m \xrightarrow{(\ )^n} \mathbb{G}_m \to 0$ induces the commutative diagram

$$\begin{tikzcd} H^1(\mathrm{Spec}(A)_{\text{ét}}, \mu_n) \arrow[r, "\delta_1"] \arrow[d, "\alpha"] & \mathrm{Pic}(A) \arrow[r, "(\ )^n"] \arrow[d] & \mathrm{Pic}(A) \arrow[d] \\ H^1(\mathrm{Spec}(A/tA)_{\text{ét}}, \mu_n) \arrow[r, "\delta_2"] & \mathrm{Pic}(A/tA) \arrow[r, "(\ )^n"] & \mathrm{Pic}(A/tA). \end{tikzcd}$$

By Theorem 4.1, the map α is an isomorphism. Then if $\mathrm{Pic}(A)$ has no torsion element of order n, δ_1 is the zero map. This implies that δ_2 is also the zero map and hence, $\mathrm{Pic}(A/tA)$ has no element of order n. Since both A and A/tA are locally factorial by assumption, we have $\mathrm{Cl}(A) \cong \mathrm{Pic}(A)$ and $\mathrm{Cl}(A/tA) \cong \mathrm{Pic}(A/tA)$. The assertion follows. $\square$

It is not necessarily true that δ_1 or δ_2 are injective, because we do not assume A to be strictly Henselian.

Lemma 4.12. *Let $(R, \mathcal{Q}, \alpha)$ be a log-regular ring. Then strict Henselization $(R^{\text{sh}}, \mathcal{Q}, \alpha^{\text{sh}})$ is also a log-regular ring, where $\alpha^{\text{sh}} : \mathcal{Q} \to R \to R^{\text{sh}}$ is the composition of homomorphisms.*

Proof. Since $R \to R^{\text{sh}}$ is a local ring map, $(R^{\text{sh}}, \mathcal{Q}, \alpha^{\text{sh}})$ is a local log ring by Lemma 2.18. Note that we have the equality $I_{\alpha^{\text{sh}}} = I_\alpha R^{\text{sh}}$. Since we have the isomorphism $R^{\text{sh}}/I_{\alpha^{\text{sh}}} \cong (R/I_\alpha)^{\text{sh}}$ by [38, Tag 05WS] and $(R/I_\alpha)^{\text{sh}}$ is a regular local ring by [38, Tag 06LN], $R^{\text{sh}}/I_{\alpha^{\text{sh}}}$ is a regular local ring. Since the dimension of R is equal to the dimension of a strict henselization R^{sh}, we obtain the equalities

$$\dim R^{\mathrm{sh}} - \dim(R^{\mathrm{sh}}/I_{\alpha^{\mathrm{sh}}}) = \dim R^{\mathrm{sh}} - \dim(R/I_\alpha)^{\mathrm{sh}} = \dim R - \dim(R/I_\alpha) = \dim \mathcal{Q}.$$

So the local log ring $(R^{\mathrm{sh}}, \mathcal{Q}, \alpha^{\mathrm{sh}})$ is log-regular. □

Now we can prove the following result on the divisor class groups of local log-regular rings, as an application of the theory of perfectoid towers.

Theorem 4.13. *Let $(R, \mathcal{Q}, \alpha)$ be a local log-regular ring of mixed characteristic with perfect residue field k of characteristic $p > 0$, and denote by $\mathrm{Cl}(R)$ the divisor class group with its torsion subgroup $\mathrm{Cl}(R)_{\mathrm{tor}}$.*

(1) *Assume that $R \cong W(k)[\![\mathcal{Q}]\!]$ for a fine, sharp, and saturated monoid $\mathcal{Q}$, where $W(k)$ is the ring of Witt vectors over k. Then $\mathrm{Cl}(R)_{\mathrm{tor}} \otimes \mathbb{Z}\big[\frac{1}{p}\big]$ is a finite group. In other words, the ℓ-primary subgroup of $\mathrm{Cl}(R)_{\mathrm{tor}}$ is finite for all primes $\ell \neq p$ and vanishes for almost all primes $\ell \neq p$.*

(2) *Assume that $\widehat{R^{\mathrm{sh}}}\big[\frac{1}{p}\big]$ is locally factorial, where $\widehat{R^{\mathrm{sh}}}$ is the completion of the strict Henselization R^{sh}. Then $\mathrm{Cl}(R)_{\mathrm{tor}} \otimes \mathbb{Z}\big[\frac{1}{p}\big]$ is a finite group. In other words, the ℓ-primary subgroup of $\mathrm{Cl}(R)_{\mathrm{tor}}$ is finite for all primes $\ell \neq p$ and vanishes for almost all primes $\ell \neq p$.*

Proof. Assertion (1) was already proved in Proposition 2.26. So let us prove assertion (2). We may assume that $\mathcal{Q}$ is fine, sharp, and saturated by Remark 2.20. The proof given below works for the first case under the assumption of local factoriality of $\widehat{R^{\mathrm{sh}}}\big[\frac{1}{p}\big]$.

Since $R \to \widehat{R^{\mathrm{sh}}}$ is a local flat ring map, the induced map $\mathrm{Cl}(R) \to \mathrm{Cl}(\widehat{R^{\mathrm{sh}}})$ is injective by Mori's theorem (cf. [11, Corollary 6.5.2]). Thus, it suffices to prove the theorem for $\widehat{R^{\mathrm{sh}}}$. Moreover, $\widehat{R^{\mathrm{sh}}}$ is log-regular with respect to the induced log ring structure $\alpha : \mathcal{Q} \to R \to \widehat{R^{\mathrm{sh}}}$ by Lemma 4.12. So without loss of generality, we may assume that the residue field of R is separably closed (hence algebraically closed in our case).

Henceforth, we denote $\widehat{R^{\mathrm{sh}}}$ by R for brevity and fix a prime ℓ that is different from p. By Lemma 4.9 and the local factoriality of $R\big[\frac{1}{p}\big]$, we claim that there is an open subset $U \subseteq X := \mathrm{Spec}(R)$ such that

$$\mathrm{Pic}(U) \cong \mathrm{Cl}(X), \quad X \setminus V(pR) \subseteq U \quad \text{and} \quad \mathrm{codim}_X(X \setminus U) \geq 2. \tag{4-8}$$

Indeed, X is a normal integral scheme by Kato's theorem (Theorem 2.21). Let U be the union of the regular locus of X and the open $\mathrm{Spec}\big(R\big[\frac{1}{p}\big]\big) \subseteq X$. Then by Serre's normality criterion, we see that $\mathrm{codim}_X(X \setminus U) \geq 2$. We fix such an open $U \subseteq X$ once and for all. Taking the cohomology sequence associated to the exact sequence

$$0 \to \mathbb{Z}/\ell^n\mathbb{Z} \to \mathbb{G}_m \xrightarrow{(\ )^{\ell^n}} \mathbb{G}_m \to 0$$

on the strict local scheme X and arguing as in the proof of Proposition 4.10, we have an isomorphism

$$H^1(U_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \cong \mathrm{Pic}(U)[\ell^n] \cong \mathrm{Cl}(X)[\ell^n]. \tag{4-9}$$

On the other hand, there is a perfectoid tower of module-finite extensions of local log-regular rings arising from $(R, (p))$:

$$(R, \mathcal{Q}, \alpha) = (R_0, \mathcal{Q}^{(0)}, \alpha_0) \to \cdots \to (R_j, \mathcal{Q}^{(j)}, \alpha_j) \to (R_{j+1}, \mathcal{Q}^{(j+1)}, \alpha_{j+1}) \to \cdots . \tag{4-10}$$

Each map is generically of p-power rank in view of Lemma 2.16(2) and Lemma 2.14(3). Moreover, the tilt of (4-10) (associated to $(R, (p))$) is given by

$$(R^{s.\flat}, \mathcal{Q}^{s.\flat}, \alpha^{s.\flat}) = ((R_0)^{s.\flat}_{(p)}, \mathcal{Q}^{s.\flat}_0, \alpha^{s.\flat}_0) \to \cdots \to ((R_j)^{s.\flat}_{(p)}, \mathcal{Q}^{s.\flat}_j, \alpha^{s.\flat}_j) \to ((R_{j+1})^{s.\flat}_{(p)}, \mathcal{Q}^{s.\flat}_{j+1}, \alpha^{s.\flat}_{j+1}) \to \cdots ,$$

where $((R_j)^{s.\flat}_{(p)}, \mathcal{Q}^{s.\flat}_j, \alpha^{s.\flat}_j)$ is a complete local log-regular ring of characteristic $p > 0$ in view of Theorem 3.61. The local ring $R^{s.\flat}$ is strictly Henselian and the complement of $U^{s.\flat}(= U^{s.\flat}_{(p)})$ has codimension ≥ 2 in $\operatorname{Spec}(R^{s.\flat})$. By repeating the proof of Proposition 4.10, we obtain an isomorphism

$$H^1(U^{s.\flat}_{\text{ét}}, \mathbb{Z}/\ell^n\mathbb{Z}) \cong \operatorname{Pic}(U^{s.\flat})[\ell^n]. \tag{4-11}$$

By Lemma 4.9, the map

$$\operatorname{Pic}(U^{s.\flat})[\ell^n] \to \operatorname{Cl}(R^{s.\flat})[\ell^n] \tag{4-12}$$

is injective. Combining (4-9), (4-11), (4-12) and Proposition 4.7, it is now sufficient to check that there exists an integer $N > 0$ depending only on $R^{s.\flat}$ such that

$$\operatorname{Cl}(R^{s.\flat})[\ell^N] = \bigcup_{n>0} \operatorname{Cl}(R^{s.\flat})[\ell^n], \text{ and } \operatorname{Cl}(R^{s.\flat})[\ell^N] \text{ is finite for all } \ell \text{ and zero for almost all } \ell \neq p.$$

Since we know that $R^{s.\flat}$ is strongly F-regular by Theorem 3.61 and Lemma 2.25, the aforementioned result of Polstra finishes the proof. □

Appendix: Construction of differential modules and maximality

The content of this appendix is taken from Gabber and Ramero's treatise [17], whose purpose is to supply a corrected version of Grothendieck's original presentation in EGA. So we give only a sketch of the constructions of relevant modules and maps. Readers are encouraged to look into [17] for more details as well as proofs. We are motivated by the following specific problem.

Problem A.1. *Let $(A, \mathfrak{m}_A)$ be a Noetherian regular local ring and fix a system of elements $f_1, \dots, f_n \in A$ and a system of integers $e_1, \dots, e_n$ with $e_i > 1$ for every $i = 1, \dots, n$. We set*

$$B := A[T_1, \dots, T_n]/(T_1^{e_1} - f_1, \dots, T_d^{e_n} - f_n).$$

Then find a sufficient condition that ensures that the localization B with respect to a maximal ideal $\mathfrak{n}$ with $\mathfrak{m}_A = A \cap \mathfrak{n}$ is regular.

From the construction, it is obvious that the induced ring map $A \to B$ is a flat finite injective extension. Let now $(A, \mathfrak{m}_A, k)$ be a Noetherian local ring with residue field $k_A := A/\mathfrak{m}_A$ of characteristic $p > 0$.

Following the presentation in [17, (9.6.15)], we define a certain $k_A^{1/p}$-vector space $\mathbf{\Omega}_A$ together with a map $\boldsymbol{d}_A : A \to \mathbf{\Omega}_A$ as follows.

Case I: $p \notin \mathfrak{m}_A^2$. Let $W_2(k_A)$ denote the p-typical ring of length 2 Witt vectors over k_A. Then there is the ghost component map $\overline{\omega}_0 : W_2(k_A) \to k_A$, and set $V_1(k_A) := \mathrm{Ker}(\overline{\omega}_0)$. More specifically, we have $W_2(k_A) = k_A \times k_A$ as sets with addition and multiplication given respectively by

$$(a, b) + (c, d) = \left(a + c, b + d + \frac{a^p + c^p - (a + c)^p}{p}\right) \quad \text{and} \quad (a, b)(c, d) = (ac, a^p d + c^p b).$$

Using this structure, we see that $V_1(k_A) = 0 \times k_A$ as sets, which is an ideal of $W_2(k_A)$ and $V_1(k_A)^2 = 0$. This makes $V_1(k_A)$ equipped with the structure as a k_A-vector space by letting $x(0, a) := (x, 0)(0, a)$ for $x \in k_A$. One can define the map of k_A-vector spaces

$$k_A^{1/p} \to V_1(k_A)\ ;\ a \mapsto (0, a^p), \tag{A-1}$$

which is a bijection. With this isomorphism, we may view $V_1(k_A)$ as a $k_A^{1/p}$-vector space. Next we form the fiber product ring:

$$A_2 := A \times_{k_A} W_2(k_A).$$

It gives rise to a short exact sequence of A_2-modules

$$0 \to V_1(k_A) \to A_2 \to A \to 0, \tag{A-2}$$

where $A_2 \to A$ is the natural projection, and the A_2-module structure of $V_1(k_A)$ is via the restriction of rings $A_2 \to W_2(k_A)$. From (A-2), we obtain an exact sequence of A-modules:

$$V_1(k_A) \to \overline{\Omega}_A \to \Omega^1_{A/\mathbb{Z}} \to 0,$$

where we put $\overline{\Omega}_A = \Omega^1_{A_2/\mathbb{Z}} \otimes_{A_2} A$. After applying $(\ ) \otimes_A k_A$ to this sequence, we have another sequence of k_A-vector spaces:

$$0 \to V_1(k_A) \xrightarrow{j_A} \overline{\Omega}_A \otimes_A k_A \to \Omega^1_{A/\mathbb{Z}} \otimes_A k_A \to 0. \tag{A-3}$$

Then this is right exact. Moreover, (A-1) yields a unique k_A-linear map $\psi_A : V_1(k_A) \otimes_{k_A} k_A^{1/p} \to V_1(k_A)$. Define $\mathbf{\Omega}_A$ as the push-out of the diagram:

$$V_1(k_A) \xleftarrow{\psi_A} V_1(k_A) \otimes_{k_A} k_A^{1/p} \xrightarrow{j_A \otimes k_A^{1/p}} \overline{\Omega}_A \otimes_A k_A^{1/p}.$$

More concretely, we have

$$\mathbf{\Omega}_A = \frac{V_1(k_A) \oplus (\overline{\Omega}_A \otimes_A k_A^{1/p})}{T},$$

where $T = \left\{(\psi(x), -(j_A \otimes k_A^{1/p})(x)) \mid \in V_1(k_A) \otimes_{k_A} k_A^{1/p}\right\}$. By the universality of push-outs, we get the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & V_1(k_A)\otimes_{k_A} k_A^{1/p} \arrow[r] \arrow[d, "\psi_A"] & \overline{\Omega}_A \otimes_A k_A^{1/p} \arrow[r] \arrow[d, "\boldsymbol{\psi}_A"] & \Omega^1_{A/\mathbb{Z}} \otimes_A k_A^{1/p} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & V_1(k_A) \arrow[r] & \boldsymbol{\Omega}_A \arrow[r] & \Omega^1_{A/\mathbb{Z}} \otimes_A k_A^{1/p} \arrow[r] & 0.
\end{tikzcd}$$

We define the map

$$\boldsymbol{d}_A : A \to \boldsymbol{\Omega}_A$$

as the composite mapping

$$A \xrightarrow{1\times \tau_{k_A}} A_2 = A \times_{k_A} W_2(k_A) \xrightarrow{d} \Omega^1_{A_2/\mathbb{Z}} \xrightarrow{\mathrm{id}\otimes 1} \overline{\Omega}_A = \Omega^1_{A_2/\mathbb{Z}} \otimes_A k_A^{1/p} \xrightarrow{\boldsymbol{\psi}_A} \boldsymbol{\Omega}_A.$$

Here, $d : A_2 \to \Omega^1_{A_2/\mathbb{Z}}$ is the universal derivation and $\tau_{k_A} : A \to k_A \to W_2(k_A)$, where the first map is the natural projection and the second one is the Teichmüller map.

Case II: $p \in \mathfrak{m}_A^2$. We just set $\boldsymbol{\Omega}_A := \Omega^1_{A/\mathbb{Z}} \otimes_A k_A^{1/p}$, and define $\boldsymbol{d}_A : A \to \boldsymbol{\Omega}_A$ as the map induced by the universal derivation $d_A : A \to \Omega^1_{A/\mathbb{Z}}$.

Combining Cases I and II, we have a map $\boldsymbol{d}_A : A \to \boldsymbol{\Omega}_A$. If $\phi : (A, \mathfrak{m}_A) \to (B, \mathfrak{m}_B)$ is a local ring map of local rings, this gives rise to the commutative diagram

$$\begin{tikzcd}
A \arrow[r, "\boldsymbol{d}_A"] \arrow[d, "\phi"] & \boldsymbol{\Omega}_A \arrow[d, "\boldsymbol{\Omega}_\phi"] \\
B \arrow[r, "\boldsymbol{d}_B"] & \boldsymbol{\Omega}_B.
\end{tikzcd}$$

With this in mind, one can consider the functor $A \mapsto \boldsymbol{\Omega}_A$ from the category of local rings $(A, \mathfrak{m}_A)$ of residual characteristic $p > 0$ to the category of the $k_A^{1/p}$-vector spaces $\boldsymbol{\Omega}_A$. Some distinguished features in this construction are as follows:

Proposition A.2 [17, Proposition 9.6.20]. *Let $\phi : (A, \mathfrak{m}_A) \to (B, \mathfrak{m}_B)$ be a local ring map of Noetherian local rings such that the residual characteristic of A is $p > 0$. Then*

(1) *Suppose that ϕ is formally smooth for the $\mathfrak{m}_A$-adic topology on A and the $\mathfrak{m}_B$-adic topology on B. Then the maps induced by ϕ and $\boldsymbol{\Omega}_\phi$, namely*

$$(\mathfrak{m}_A/\mathfrak{m}_A^2)\otimes_{k_A} k_B \to \mathfrak{m}_B/\mathfrak{m}_B^2 \quad \text{and} \quad \boldsymbol{\Omega}_A \otimes_{K_A^{1/p}} k_B^{1/p} \to \boldsymbol{\Omega}_B,$$

are injective.

(2) *Suppose that*

(a) $\mathfrak{m}_A B = \mathfrak{m}_B$,

(b) *the residue field extension $k_A \to k_B$ is separable algebraic,*

(c) *ϕ is flat.*

Then $\boldsymbol{\Omega}_\phi$ *induces an isomorphism of* $k_A^{1/p}$*-vector spaces*:

$$\boldsymbol{\Omega}_A \otimes_A B \cong \boldsymbol{\Omega}_B.$$

(3) *If* $B = A/\mathfrak{m}_A^2$ *and* $\phi : A \to B$ *is the natural map, then* $\boldsymbol{\Omega}_\phi$ *is an isomorphism.*

(4) *The functor* $\boldsymbol{\Omega}_\bullet$ *and the natural transformation* $\boldsymbol{d}_\bullet$ *commute with filtered colimits.*

We provide an answer to Problem A.1 as follows.

Theorem A.3 [17, Corollary 9.6.34]. *Let* $f_1, \dots, f_n$ *be a sequence of elements in* A, *and let* $e_1, \dots, e_n$ *be a system of integers with* $e_i > 1$ *for every* $i = 1, \dots, n$. *Set*

$$C := A[T_1, \dots, T_n]/(T_1^{e_1} - f_1, \dots, T_n^{e_n} - f_n).$$

Fix a prime ideal $\mathfrak{n} \subseteq C$ *such that* $\mathfrak{n} \cap A = \mathfrak{m}_A$, *and let* $B := C_{\mathfrak{n}}$. *Let* $E \subseteq \boldsymbol{\Omega}_A$ *be the* $k_A^{1/p}$*-vector space spanned by* $\boldsymbol{d}_A f_1, \dots, \boldsymbol{d}_A f_n$. *The following conditions are equivalent.*

(1) A *is a regular local ring*, *and* $\dim_{k_A^{1/p}} E = n$.

(2) B *is a regular local ring.*

In particular, in the situation of the above theorem, B is a regular local ring if A is a regular local ring and $f_1, \dots, f_n$ is *maximal* in the sense of the following definition.

Definition A.4. Let $(A, \mathfrak{m}_A, k_A)$ be a local ring with residual characteristic $p > 0$. Then we say that a sequence of elements $f_1, \dots, f_n$ in A is *maximal* if $\boldsymbol{d}_A f_1, \dots, \boldsymbol{d}_A f_n$ forms a basis of the $k_A^{1/p}$-vector space $\boldsymbol{\Omega}_A$.

In general, we have the following fact.

Lemma A.5. *Let* $(A, \mathfrak{m}_A, k_A)$ *be a regular local ring of mixed characteristic and assume that* $f_1, \dots, f_d$ *is a regular system of parameters of* A.

(1) $f_1, \dots, f_d$ *satisfies condition* (1) *of Theorem A.3.*

(2) *If the residue field* k_A *of* A *is perfect*, *then the sequence* $f_1, \dots, f_d$ *is maximal.*

Proof. (1) In the case that $p \notin \mathfrak{m}_A^2$, [17, Proposition 9.6.17] gives a short exact sequence

$$0 \to \mathfrak{m}_A/\mathfrak{m}_A^2 \otimes_{k_A} k_A^{1/p} \to \boldsymbol{\Omega}_A \to \Omega^1_{k_A/\mathbb{Z}} \otimes_{k_A} k_A^{1/p} \to 0. \tag{A-4}$$

Then the images $\overline{f}_1, \dots, \overline{f}_d$ form a basis of the $k_A^{1/p}$-vector space $\mathfrak{m}_A/\mathfrak{m}_A^2 \otimes_{k_A} k_A^{1/p}$. The desired claim follows from the left exactness of (A-4).

In the case that $p \in \mathfrak{m}_A^2$, [17, Lemma 9.6.6] gives a short exact sequence

$$0 \to \mathfrak{m}_A/(\mathfrak{m}_A^2 + p\mathfrak{m}_A) \to \Omega_A \to \Omega^1_{k_A/\mathbb{Z}} \to 0. \tag{A-5}$$

and we can argue as in the case $p \notin \mathfrak{m}_A^2$.

(2) If k_A is perfect, then $\Omega^1_{k_A/\mathbb{Z}} = 0$. Therefore, (A-4) and (A-5) (in the latter case, one tensors it with $k_A^{1/p}$ over k_A) gives the desired conclusion. ☐

Acknowledgements

The authors are grateful to Professor K. Fujiwara for his continued support and comments. Our gratitude also goes to Professor Shunsuke Takagi, Shunya Saito, Ryo Ishizuka, and Kazuki Hayashi for their advice on the content of this paper. Finally, we sincerely thank the referees for their careful reading and valuable suggestions.

Kei Nakazato was partially supported by JSPS Grant-in-Aid for Early-Career Scientists 23K12952. Kazuma Shimomoto was partially supported by JSPS Grant-in-Aid for Scientific Research(C) 18K03257 and Scientific Research(C) 23K03077.

References

[1] I. M. Aberbach, "Extension of weakly and strongly F-regular rings by flat maps", *J. Algebra* **241**:2 (2001), 799–807. MR

[2] Y. André, "La conjecture du facteur direct", *Publ. Math. Inst. Hautes Études Sci.* **127** (2018), 71–93. MR

[3] B. Bhatt, "On the non-existence of small Cohen–Macaulay algebras", *J. Algebra* **411** (2014), 1–11. MR

[4] B. Bhatt, "On the direct summand conjecture and its derived variant", *Invent. Math.* **212**:2 (2018), 297–317. MR

[5] B. Bhatt, M. Morrow, and P. Scholze, "Integral p-adic Hodge theory", *Publ. Math. Inst. Hautes Études Sci.* **128** (2018), 219–397. MR

[6] H. Cai, S. Lee, L. Ma, K. Schwede, and K. Tucker, "Perfectoid signature, perfectoid Hilbert–Kunz multiplicity, and an application to local fundamental groups", preprint, 2023. arXiv 2209.04046

[7] K. Česnavičius, "Purity for the Brauer group", *Duke Math. J.* **168**:8 (2019), 1461–1486. MR

[8] K. Česnavičius and P. Scholze, "Purity for flat cohomology", *Ann. of Math.* (2) **199**:1 (2024), 51–180. MR

[9] L. Claborn, "Every abelian group is a class group", *Pacific J. Math.* **18** (1966), 219–222. MR

[10] R. Datta and K. Tucker, "On some permanence properties of (derived) splinters", *Michigan Math. J.* **73**:2 (2023), 371–400. MR

[11] T. J. Ford, *Separable algebras*, Grad. Stud. Math. **183**, Amer. Math. Soc., Providence, RI, 2017. MR

[12] L. Fu, *Étale cohomology theory*, Nankai Tracts in Math. **13**, World Sci., Hackensack, NJ, 2011. MR

[13] K. Fujiwara, "Theory of tubular neighborhood in étale topology", *Duke Math. J.* **80**:1 (1995), 15–57. MR

[14] K. Fujiwara and F. Kato, *Foundations of rigid geometry, I*, Eur. Math. Soc., Zürich, 2018. MR

[15] K. Fujiwara, O. Gabber, and F. Kato, "On Hausdorff completions of commutative rings in rigid geometry", *J. Algebra* **332** (2011), 293–321. MR

[16] O. Gabber, "Affine analog of the proper base change theorem", *Israel J. Math.* **87**:1-3 (1994), 325–335. MR

[17] O. Gabber and L. Ramero, "Almost rings and perfectoid spaces", preprint, 2018, https://pro.univ-lille.fr/fileadmin/user_upload/pages_pros/lorenzo_ramero/hodge.pdf.

[18] P. Griffith and D. Weston, "Restrictions of torsion divisor classes to hypersurfaces", *J. Algebra* **167**:2 (1994), 473–487. MR

[19] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV", *Inst. Hautes Études Sci. Publ. Math.* **32** (1967), 5–361. MR

[20] A. Grothendieck, *Cohomologie locale des faisceaux cohérents et théorèmes de Lefschetz locaux et globaux* (Séminaire de Géométrie Algébrique du Bois Marie 1962), Doc. Math. (Paris) **4**, Soc. Math. France, Paris, 2005. MR arXiv math/0511279

[21] D. Hansen and K. S. Kedlaya, "Sheafiness criteria for Huber rings", preprint, 2021, https://kskedlaya.org/papers/criteria.pdf.

[22] M. Hochster, "Rings of invariants of tori, Cohen–Macaulay rings generated by monomials, and polytopes", *Ann. of Math.* (2) **96** (1972), 318–337. MR

[23] M. Hochster and C. Huneke, "Tight closure and strong F-regularity", pp. 119–133 in *Colloque en l'honneur de Pierre Samuel* (Orsay, France, 1987), Mém. Soc. Math. France (N.S.) **38**, Soc. Math. France, Paris, 1989. MR

[24] L. Illusie, Y. Laszlo, and F. Orgogozo (editors), *Travaux de Gabber sur l'uniformisation locale et la cohomologie étale des schémas quasi-excellents*, Astérisque **363-364**, Soc. Math. France, Paris, 2014. MR

[25] S. Ishiro, "Local log-regular rings vs. toric rings", *Comm. Algebra* **53**:4 (2025), 1640–1655. MR

[26] K. Kato, "Logarithmic structures of Fontaine–Illusie", pp. 191–224 in *Algebraic analysis, geometry, and number theory* (Baltimore, MD, 1988), edited by J.-I. Igusa, Johns Hopkins Univ. Press, Baltimore, MD, 1989. MR

[27] K. Kato, "Toric singularities", *Amer. J. Math.* **116**:5 (1994), 1073–1099. MR

[28] L. Ma and T. Polstra, "F-singularities: a commutative algebra approach", preprint, 2023, https://www.math.purdue.edu/~ma326/F-singularitiesBook.pdf.

[29] H. Matsumura, *Commutative ring theory*, Cambridge Stud. Adv. Math. **8**, Cambridge Univ. Press, 1986. MR

[30] E. Miller and B. Sturmfels, *Combinatorial commutative algebra*, Grad. Texts in Math. **227**, Springer, 2005. MR

[31] J. S. Milne, *Étale cohomology*, Princeton Math. Ser. **33**, Princeton Univ. Press, 1980. MR

[32] T. Murayama, "Relative vanishing theorems for $\mathbb{Q}$-schemes", *Algebr. Geom.* **12**:1 (2025), 84–144. MR

[33] K. Nakazato and K. Shimomoto, "Finite étale extensions of Tate rings and decompletion of perfectoid algebras", *J. Algebra* **589** (2022), 114–158. MR

[34] A. Ogus, *Lectures on logarithmic algebraic geometry*, Cambridge Stud. Adv. Math. **178**, Cambridge Univ. Press, 2018. MR

[35] T. Polstra, "A theorem about maximal Cohen–Macaulay modules", *Int. Math. Res. Not.* **2022**:3 (2022), 2086–2094. MR

[36] P. Schenzel, "A criterion for I-adic completeness", *Arch. Math. (Basel)* **102**:1 (2014), 25–33. MR

[37] K. Shimomoto, "Almost Cohen–Macaulay algebras in mixed characteristic via Fontaine rings", *Illinois J. Math.* **55**:1 (2011), 107–125. MR

[38] "The Stacks project", electronic reference, 2005–, http://stacks.math.columbia.edu.

Communicated by Bhargav Bhatt
Received 2023-03-09 Revised 2024-08-08 Accepted 2024-12-23

shinnosukeishiro@gmail.com *National Institute of Technology, Gunma College, Gunma 371-8530, Japan*

keinakazato31@gmail.com *Proxima Technology Inc., Tokyo 111-0005, Japan*

shimomotokazuma@gmail.com *Department of Mathematics, Institute of Science Tokyo, Tokyo 152-8551, Japan*

ALGEBRA AND NUMBER THEORY 19:12 (2025)
https://doi.org/10.2140/ant.2025.19.2359

Reduction modulo p of Noether's problem

Emiliano Ambrosi and Domenico Valloni

Let R be a complete valuation ring of mixed characteristic $(0, p)$ with algebraically closed fraction field K and residue field k. Let X/R be a smooth projective morphism. We show that if X_k is stably rational, then $H^3_{\text{ét}}(X_K, \hat{\mathbb{Z}})$ is torsion-free. The proof uses the integral p-adic Hodge theory of Bhatt, Morrow and Scholze and the study of differential forms in positive characteristic. We then apply this result to study the Noether problem for finite p-groups.

1. Introduction

Many techniques to study the birational geometry of varieties are based on specialization methods in equicharacteristic or in mixed characteristic (see [Voi15; CTP16], for instance), often exploiting pathological phenomena in positive characteristic [Kol95; Tot16; ABBGvB21].

In this paper we study birational invariants in mixed characteristic. Let us fix a complete valuation ring R of mixed characteristic $(0, p)$ with algebraically closed fraction field K and residue field k (necessarily algebraically closed) and a smooth proper morphism X/R. Motivated by the different behavior of the Noether problem in positive and zero characteristic (see Section 1.2 for details) we study the following question:

Question 1.1. What can one say about the generic fiber X_K knowing that the special fiber X_k is stably rational? (The condition means that $X_k \times \mathbb{P}^n_k$ is birational to $\mathbb{P}^N_k$ for some $n, N \in \mathbb{N}$.)

In general, X_K need not be stably rational, as we show for instance in Section 3, where we follow a suggestion of Colliot-Thélène and adapt the constructions of [HPT18] to construct smooth proper morphisms X/R with X_K stably irrational and X_k rational. On the other hand, we show that the following vanishing holds:

Theorem 1.2. *Let X/R be a smooth proper morphism, and assume that the special fiber X_k is stably rational. Then the p-torsion of $H^3_{\text{ét}}(X_K, \mathbb{Z}_p)$ vanishes.*

MSC2020: 14E08, 14F30, 14M20.
Keywords: Noether's problem, crystalline cohomology, Artin–Mumford invariant.

If $K \subset \mathbb{C}$ is an embedding, then the vanishing of $H^3_{\text{ét}}(X_K, \mathbb{Z}_p)[p]$ together with smooth-proper base change in étale cohomology (with ℓ-adic coefficients, $\ell \neq p$) implies that $H^3_{\text{sing}}(X(\mathbb{C}), \mathbb{Z})$ is torsion free whenever X_k is stably rational. We recall that the torsion of $H^3_{\text{sing}}(X(\mathbb{C}), \mathbb{Z})$ is a stably birational invariant of smooth proper varieties over $\mathbb{C}$, which is naturally isomorphic to $\text{Br}(X)/\text{Br}(X)_{\text{div}}$, where $\text{Br}(X) := H^2_{\text{ét}}(X, \mathbb{G}_m)$. This group was used by Artin and Mumford in their seminal paper [AM72] to give the first elementary example of a unirational threefold which was not rational.

1.1. ***Strategy.*** To explain the strategy, we begin by showing an analogous statement concerning global differential forms. We retain the notation from the previous section. By the Künneth formula and Hartogs' lemma, the vector spaces $H^0(X_K, \Omega^i_{X_K})$ are stably birational invariants of smooth proper K-varieties, and therefore they vanish if X_K is stably rational. If X/R is a smooth proper morphism, the semicontinuity theorem yields the inequality

$$\dim_k(H^0(X_k, \Omega^i_{X_k})) \geq \dim_K(H^0(X_K, \Omega^i_{X_K})), \tag{1.1.3}$$

from which it follows that if X_k is stably rational, then $H^0(X_K, \Omega^i_{X_K}) = 0$ necessarily.

Concerning étale cohomology, as already mentioned, the proper smooth base change asserts that for a prime $\ell \neq p$ one has

$$H^3_{\text{ét}}(X_K, \mathbb{Z}_\ell)[\ell] \simeq H^3_{\text{ét}}(X_k, \mathbb{Z}_\ell)[\ell] = 0,$$

where the last equality follows from the stable birational invariance of $H^3_{\text{ét}}(X_k, \mathbb{Z}_\ell)[\ell]$ and the fact that X_k is stably rational. For p-adic coefficients, we can replace the smooth proper base-change with the integral p-adic Hodge theory of Bhatt, Morrow and Scholze, which will play the role of the semicontinuity theorem (1.1.3). By [BMS18, Theorem 1.1 (ii)], one has the inequality

$$\dim_k(H^3_{\text{crys}}(X_k/W))[p]) \geq \dim_{\mathbb{F}_p}(H^3_{\text{ét}}(X_K, \mathbb{Z}_p)[p]), \tag{1.1.4}$$

where $H^3_{\text{crys}}(X_k/W)$ is the third crystalline cohomology group with integral coefficients of X_k. Thus, it would be enough to show that $\dim_k(H^3_{\text{crys}}(X_k)[p])$ is a stably birational invariant of smooth proper varieties. It is unclear how to prove this without assuming resolution of singularities, also because crystalline cohomology behaves badly for open or singular varieties. We prove instead the following vanishing, which is enough to deduce Theorem 1.2:

Theorem 1.1.5. *Let k be an algebraically closed field of characteristic p and X be a smooth proper k-variety. Assume that*

(1) $H^i(X, \mathcal{O}_X) = 0$ *for* $i = 2, 3$;

(2) $H^0(X, \Omega^2_X) = 0$;

(3) $\text{Br}(X)[p] = 0$.

Then $H^3_{\text{crys}}(X/W)[p] = 0$.

The assumptions of Theorem 1.1.5 are satisfied if X is stably rational, since all the conditions are stably birational invariants of smooth proper varieties. (See [CR11, Theorem 1] for (1); point (2) follows from Hartog's lemma; [CTS21, Corollary 6.2.11, p. 170] gives (3).) Hence, Theorem 1.2 follows from Theorem 1.1.5 and (1.1.4).

1.2. ***Applications to the Noether problem.*** As mentioned, our original motivation was the Noether problem, which we now briefly recall. For a finite group G and a field K, the Noether problem asks whether $\mathbb{P}^n_K/G$ is a stably rational variety, where G acts on $\mathbb{P}^n_K$ in a linear and faithful way. The problem is well-posed since the stably birational class of the quotient does not depend on the chosen representation [BK85, Lemma 1.3]. For $K = \mathbb{C}$ we can then define the Artin–Mumford invariant $\mathrm{AM}(G)$ of G as $\mathrm{Tors}(H^3_{\mathrm{sing}}(X, \mathbb{Z}))$ where $X/\mathbb{C}$ is any smooth proper birational model of $\mathbb{P}^n_K/G$.

The first counterexample of the Noether problem over $\mathbb{C}$ was given by Saltman in [Sal84], who for any prime p produced a finite p-group G for which the quotient $\mathbb{P}^n_{\mathbb{C}}/G$ is not stably rational, by showing that $\mathrm{AM}(G) \neq 0$. A general formula for $\mathrm{AM}(G)$ was later given by Bogomolov [Bog88, Theorem 3.1], and this group is now known as the Bogomolov multiplier of G.

The connection to Theorem 1.2 comes from the classical observation that if G is a p-group and K has characteristic p, then $\mathbb{P}^n_K/G$ is always rational; see, e.g., [Kun54] and [Gas59].

Keeping the notation as in the previous section, we can fix a finite p-group G, and linear faithful actions of G on $\mathbb{P}^n_k$ and $\mathbb{P}^n_K$. Theorem 1.2 together with the observation above implies the following:

Corollary 1.2.1. *If* $\mathrm{AM}(G) \neq 0$, *there does not exists a smooth proper* X/R *such that* X_k *is stably birational to* $\mathbb{P}^n_k/G$ *and* X_K *is stably birational to* $\mathbb{P}^n_K/G$.

It follows that all the examples constructed in [Sal84; Bog88] cannot have good stably rational reduction. This also implies that if $\mathrm{AM}(G) \neq 0$ and $G \to \mathrm{PGL}_n(R)$ is a representation such that the reduction map $G \to \mathrm{PGL}_n(k)$ is injective, then one cannot resolve the singularities of $\mathbb{P}^n_R/G$ relatively to R, i.e., there cannot be a smooth projective X/R with a R-morphism $\pi : X \to \mathbb{P}^n_R/G$ which induces a resolution of singularities on both fibers.

Peyre [Pey08] constructed groups G such that $\mathbb{P}^n_{\mathbb{C}}/G$ is not rational but $\mathrm{AM}(G) = 0$. It is an interesting question at this point whether there is a p-group G for which a resolution X of $\mathbb{P}^n_R/G$ like in the corollary above exists, but $\mathbb{P}^n_K/G$ is not stably rational.

Remark 1.2.2. A related phenomenon appears in the recent work of Lazda and Skorobogatov in [LS23]. They prove that, if $p = 2$ and $Y \to \mathrm{Spec}(R)$ is an abelian surface such that X_k is not supersingular, then one can resolve the singularities of $Y/\{\pm 1\}$ to obtain a smooth proper morphism $X \to \mathrm{Spec}(R)$ such that the generic fiber is the Kummer variety $\mathrm{Kum}(Y_k)$ of Y_k and the generic fiber is the Kummer variety $\mathrm{Kum}(Y_K)$ of Y_K.

On the other hand, if Y_k is surpersingular, then $\mathrm{Kum}(Y_k)$ is a rational surface due to [Kat78]. Since $H^0(\mathrm{Kum}(Y_K), \Omega^2_{Y_K}) \neq 0$, the argument in the beginning of Section 1.1 applies, and such Y does not exist. Since $H^0(\mathrm{Kum}(Y_k), \Omega^2_{X_k}) = 0$ if and only if X_k is supersingular, the group $H^0(\mathrm{Kum}(Y_K), \Omega^2_{X_K})$ is the only obstruction to the construction of such Y in this case.

2. Proof of Theorem 1.1.5

2.1. ***Notation.*** Let k be an algebraically closed field of characteristic $p > 0$ and let X/k be a smooth proper variety such that

$$H^2(X, \mathcal{O}_X) = H^3(X, \mathcal{O}_X) = H^0(X, \Omega^2_{X/k}) = \mathrm{Br}(X)[p] = 0.$$

We let $\Omega^\bullet_X$ be the de Rham complex of X and we define the following sheaves over X:

$$Z^i_X := \mathrm{Ker}(d : \Omega^i_X \to \Omega^{i+1}_X); \quad B^i_X := \mathrm{Im}(d : \Omega^{i-1}_X \to \Omega^i_X); \quad \mathcal{H}^i_X = \frac{Z^i_X}{B^i_X}.$$

For every complex of sheaves $\mathcal{F}^\bullet$ over X and every $i \in \mathbb{N}$, we let $\tau_{\geq i}\mathcal{F}^\bullet$ (resp. $\tau_{\leq i}\mathcal{F}^\bullet$) be the upper (resp. lower) canonical truncation of $\mathcal{F}^\bullet$ and $\mathcal{F}^{\geq i}$ (resp. $\mathcal{F}^{\leq i}$) the upper (resp. lower) naive filtration of $\mathcal{F}^\bullet$. Recall that, for every $i \in \mathbb{N}$ there exists an exact triangle (see, e.g., [Stacks, Remark 08J5]):

$$\tau_{\leq i}\mathcal{F}^\bullet \to \mathcal{F}^\bullet \to \tau_{\geq i+1}\mathcal{F}^\bullet.$$

2.2. ***Preliminary reductions.*** The universal coefficient theorem for crystalline cohomology (see, e.g., [Ill79, (4.9.1), p. 623]), gives us an exact sequence

$$0 \to H^2_{\mathrm{crys}}(X/W) \otimes k \to H^2_{\mathrm{dR}}(X) \to H^3_{\mathrm{crys}}(X/W)[p] \to 0,$$

so that, to prove Theorem 1.1.5, it is enough to show that the natural map

$$H^2_{\mathrm{crys}}(X/W) \otimes k \to H^2_{\mathrm{dR}}(X)$$

is surjective. In fact, we shall prove that the cycle class map $\mathrm{cl}_{\mathrm{dR}} : \mathrm{NS}(X) \otimes k \to H^2_{\mathrm{dR}}(X/k)$ is surjective, which is enough due to the commutative diagram

(2.2.1)
$$\begin{array}{ccc} \mathrm{NS}(X) \otimes W & \xrightarrow{\mathrm{cl}_{\mathrm{crys}}} & H^2_{\mathrm{crys}}(X/W) \\ \downarrow\!\!\downarrow & & \downarrow \\ \mathrm{NS}(X) \otimes k & \xrightarrow{\mathrm{cl}_{\mathrm{dR}}} & H^2_{\mathrm{dR}}(X), \end{array}$$

where the first horizontal arrow is the crystalline cycle class map.

2.3. ***Factorization of the cycle class map.*** To study $\mathrm{cl}_{\mathrm{dR}} : \mathrm{NS}(X) \otimes k \to H^2_{\mathrm{dR}}(X/k)$, we factorize it in three arrows. Let $\iota : Z^1_X[-1] \to \Omega^\bullet$ be the natural inclusion.

Recall that, by construction, the cycle class map $\mathrm{cl}_{\mathrm{dR}} : \mathrm{NS}(X) \otimes k \to H^2_{\mathrm{dR}}(X/k)$ is induced by the dlog map

$$\mathrm{dlog} : \mathcal{O}^*_X[-1] \to \Omega^\bullet_X$$

(see, e.g., [Stacks, Section 0FLE]), which factors through

$$\mathcal{O}^*_X[-1] \xrightarrow{\mathrm{dlog}} Z^1_X[-1] \xrightarrow{\iota} \Omega^\bullet_X.$$

Hence, the cycle class map factors trough the induced map $\iota : H^1(X, Z^1_X) \to H^2_{\mathrm{dR}}(X)$, giving a first factorization

$$\mathrm{cl}_{\mathrm{dR}} : \mathrm{NS}(X) \otimes k \to H^1(X, Z^1_X) \xrightarrow{\iota} H^2_{\mathrm{dR}}(X/k).$$

To go further, recall from [Ill79, (5.1.4), p. 626], that for every i, there is a canonical isomorphism $H^i_{\mathrm{flat}}(X, \mu_p) \xrightarrow{\simeq} H^{i-1}(X, \mathcal{O}_X/(\mathcal{O}^*_X)^p)$ hence a canonical map $\alpha : H^2(X, \mu_p) \to H^1(X, Z^1_X)$ induced again by $\mathrm{dlog} : \mathcal{O}_X/(\mathcal{O}^*_X)^p \to Z^1_X$. In conclusion, the map $\mathrm{cl}_{\mathrm{dR}} : \mathrm{NS}(X) \otimes k \to H^2_{\mathrm{dR}}(X)$ factorises further as

$$\mathrm{NS}(X) \otimes k \xrightarrow{\mathrm{cl}_{\mathrm{flat}} \otimes k} H^2_{\mathrm{flat}}(X, \mu_p) \otimes k \xrightarrow{\alpha} H^1(X, Z^1_X) \xrightarrow{\iota} H^2_{\mathrm{dR}}(X),$$

where $\mathrm{cl}_{\mathrm{flat}} : \mathrm{NS}(X) \to H^2_{\mathrm{flat}}(X, \mu_p)$ is the cycle class map in flat cohomology.

2.4. ***Studying the factorization.*** To show that $\mathrm{cl}_{\mathrm{dR}} : \mathrm{NS}(X) \otimes k \to H^2_{\mathrm{dR}}(X)$ is surjective, we show that:

(1) the map $\mathrm{cl}_{\mathrm{flat}} : \mathrm{NS}(X) \to H^2_{\mathrm{flat}}(X, \mu_p)$ is surjective;

(2) the map $\alpha : H^2_{\mathrm{flat}}(X, \mu_p) \otimes k \to H^1(X, Z^1_X)$ is an isomorphism;

(3) the map $\iota : H^1(X, Z^1_X) \to H^2_{\mathrm{dR}}(X)$ is surjective.

Proof of (1). The map $\mathrm{cl}_{\mathrm{flat}} : \mathrm{NS}(X) \to H^2_{\mathrm{flat}}(X, \mu_p)$ is induced by the connecting map

$$H^1(X, \mathcal{O}^*_X) = \mathrm{Pic}(X) \to H^2_{\mathrm{flat}}(X, \mu_p)$$

in the long exact sequence associated to the short exact sequence

$$0 \to \mu_p \to \mathcal{O}^*_X \xrightarrow{(-)^p} \mathcal{O}^*_X \to 0$$

of sheaves in flat site of X. Hence the surjectivity of $\mathrm{cl}_{\mathrm{flat}} : \mathrm{NS}(X) \to H^2_{\mathrm{flat}}(X, \mu_p)$ follows directly from the assumption $H^2_{\mathrm{flat}}(X, \mathcal{O}^*_X) = H^2_{\text{ét}}(X, \mathcal{O}^*_X) = \mathrm{Br}(X) = 0$.

Proof of (2). Recall from [Ill79, (2.1.23), p. 518] the exact sequence of étale sheaves

$$0 \to \mathcal{O}^*_X/(\mathcal{O}^*_X)^p \xrightarrow{\mathrm{dlog}} Z^1_X \xrightarrow{i-C} \Omega^1_X \to 0,$$

where $i : Z^1_X \to \Omega^1_X$ is the natural inclusion and $C : Z^1_X \to \Omega^1_X$ is the Cartier operator. Since étale cohomology and Zariski cohomology agree for coherent sheaves, taking the associated long exact sequence in cohomology, we get an exact sequence

$$H^0(X, Z^1_X) \xrightarrow{i-C} H^0(X, \Omega^1_X) \to H^2_{\mathrm{fppf}}(X, \mu_p) \to H^1(X, Z^1_X) \xrightarrow{i-C} H^1(X, \Omega^1_X)$$

where i is a linear morphism and C is a Frobenius-linear map.

To prove (2), it is then enough to show that $H^0(X, Z^1_X) \xrightarrow{i-C} H^0(X, \Omega^1_X)$ is surjective and that

$$\mathrm{Ker}(H^1(X, Z^1_X) \xrightarrow{i-C} H^1(X, \Omega^1_X)) \otimes k \simeq H^1(X, Z^1_X).$$

Since $H^i(X, Z^1_X)$ and $H^i(X, \Omega^1_X)$ are finite-dimensional vector spaces, we can use the following classical lemma (see, e.g., [Mil16, Lemma 4.13, p. 128]).

Lemma 2.4.1. *Let k be an algebraically closed field of characteristic $p > 0$ and let V be a finite-dimensional vector space over k. Let $f : V \to V$ be a k-linear isomorphism and let $C : V \to V$ be any Frobenius-linear map. Then*:

(a) *$f - C$ is surjective.*

(b) *If C is bijective, then $V = \ker(f - C) \otimes k$.*

Hence to prove (2) it is enough to show that the following maps are isomorphisms:

$$i : H^0(X, Z^1_X) \to H^0(X, \Omega^1_X), \quad i : H^1(X, Z^1_X) \to H^1(X, \Omega^1_X), \quad C : H^1(X, Z^1_X) \to H^1(X, \Omega^1_X).$$

That the map $H^0(X, Z^1_X) \to H^0(X, \Omega^1_X)$ is an isomorphism follows directly from the exact sequence of sheaves

$$0 \to Z^1_X \to \Omega^1_X \xrightarrow{d} \Omega^2_X$$

and the assumption $H^0(X, \Omega^2_X) = 0$.

As for i and C, by comparing dimensions we see that it is enough to prove that

(i) $i : H^1(X, Z^1_X) \to H^1(X, \Omega^1_X)$ is injective, and

(ii) $C : H^1(X, Z^1_X) \to H^1(X, \Omega^1_X)$ is surjective.

Proof of (i). Since $Z^1_X[-1] = \tau_{\leq 1}\Omega^{\geq 1}_X$, there is an exact triangle

$$Z^1_X[-1] \to \Omega^{\geq 1}_X \to \tau_{\geq 2}\Omega^{\geq 1}_X,$$

so that the map $H^1(X, Z^1_X) \to H^2(X, \Omega^{\geq 1}_X)$ is injective, since $H^1(X, \tau_{\geq 2}\Omega^{\geq 1}_X) = 0$, because $\tau_{\geq 2}\Omega^{\geq 1}_X$ is concentrated in degrees ≥ 2. So it is enough to show that the natural map $H^2(X, \Omega^{\geq 1}_X) \to H^2(X, \Omega^1_X[-1]) = H^1(X, \Omega^1_X)$ induced by the map $\Omega^{\geq 1}_X \to \Omega^1_X[-1]$ is injective. But the latter fits in the short exact sequence of complexes

$$0 \to \Omega^{\geq 2}_X \to \Omega^{\geq 1}_X \to \Omega^1_X[-1] \to 0,$$

so that we just need to prove that $H^2(X, \Omega^{\geq 2}_X) = 0$. This follows from the inclusion

$$H^2(X, \Omega^{\geq 2}_X) = \mathrm{Ker}(H^0(X, \Omega^2_X) \xrightarrow{d} H^0(X, \Omega^3_X)) \subseteq H^0(X, \Omega^2_X) = 0$$

by our assumption on X. Hence $i : H^1(X, Z^1_X) \to H^1(X, \Omega^1_X)$ is injective.

Proof of (ii). Since k is perfect, from [Ill79, (2.1.22)] we have a short exact sequence of sheaves

$$0 \to B^1_X \to Z^1_X \xrightarrow{C} \Omega^1_X \to 0,$$

where C is the Cartier operator. Hence it is enough to show that $H^2(X, B^1_X)$ is zero. But this follows from the short exact sequence

$$0 \to \mathcal{O}_X \xrightarrow{(-)^p} \mathcal{O}_X \xrightarrow{d} B^1_X \to 0$$

and the assumption $H^2(X, \mathcal{O}_X) = H^3(X, \mathcal{O}_X) = 0$.

Proof of (3). Since $H^2(X, \mathcal{O}_X)=0$ by assumption, the short exact sequence

$$0 \to Z^1_X[-1] \to \tau_{\leq 1}\Omega^\bullet_X \to \mathcal{O}_X \to 0$$

shows that the natural map $H^1(X, Z^1_X) \to H^2(X, \tau_{\leq 1}\Omega^\bullet_X)$ is surjective. So, it is enough to show that the natural map $H^2(X, \tau_{\leq 1}\Omega^\bullet_X) \to H^2(X, \Omega^\bullet_X) = H^2_{\mathrm{dR}}(X)$ is surjective. Since there is an exact triangle

$$\tau_{\leq 1}\Omega^\bullet_X \to \Omega^\bullet_X \to \tau_{\geq 2}\Omega^\bullet_X,$$

it is then enough to show that $H^2(X, \tau_{\geq 2}\Omega^\bullet_X)=0$. But

$$H^2(X, \tau_{\geq 2}\Omega^\bullet_X) = \ker(H^0(X, \tau_{\geq 2}\Omega^\bullet_X[2]) \to H^0(X, \Omega^3_X)) = H^0(X, \ker(\Omega^2_X/B^2_X \to \Omega^3_X)) = H^0(X, \mathcal{H}^2_X).$$

Again, since k is perfect, the inverse Cartier operator [Ill79, (2.1.22)] gives an isomorphism $\mathcal{H}^2(\Omega^\bullet_X) \cong \Omega^2_X$, so

$$H^2(X, \tau_{\geq 2}\Omega^\bullet_X) \simeq H^0(X, \Omega^2_X) = 0.$$

Hence the natural map $H^1(X, Z^1_X) \to H^2_{\mathrm{dR}}(X/k)$ is surjective. This concludes the proof of (3) and the proof of Theorem 1.1.5.

3. A stably irrational variety reducing to a rational variety

Let R be the ring of integers of $K := \mathbb{C}_p$ and k its residue field. In this last section, we show how to construct, for every $p \gg 0$, examples of smooth proper schemes $X \to \mathrm{Spec}(R)$ such that X_K is not stably rational and such that X_k is rational, as suggested to us by Colliot-Thélène. The construction uses and is based on the analogous construction in [HPT18] of a family of proper smooth varieties over the complex number with stably irrational general fiber but with some rational fiber.

3.1. ***A general lemma.*** We begin by reducing the construction of examples to the construction of mixed characteristic families with properties that are easier to check. Let B/R be smooth with geometrically integral fibers and $X \to B$ a smooth proper family of varieties.

Lemma 3.1.1. *Assume that there exists a point $b \in B(\mathbb{C}_p)$ such that X_b is not stably rational. Then, for every $a \in B(k)$ there exists a lift $b' \in B(R)$ of a such that $X_{b'}$ is not stably rational.*

Proof. By [NO21, Corollary 4.1.2], the set

$$B(\mathbb{C}_p)_r := \{b \in B(\mathbb{C}_p) : X_b \text{ is stably rational}\}$$

is a countable union of closed subvarieties. Define now $B(\mathbb{C}_p)_{nr} := B(\mathbb{C}_p) \setminus B(\mathbb{C}_p)_r$. By the assumption on b, the set $B(\mathbb{C}_p)_r$ is the countable union of *proper* closed subvarieties.

Since, by the Hensel lemma, the map $\pi : B(R) \to B(k)$ is surjective, we can choose a lift b'' of a. The set $\pi^{-1}(a) \subseteq B(R) \subseteq B(K)$ is an open neighborhood of b'' in $B(K)$. Since $B(\mathbb{C}_p)_r$ is the countable union of proper closed subvarieties, we can apply [MP12, Lemma 4.29] to deduce that there exists a $b' \in B(\mathbb{C}_p)_{nr} \cap \pi^{-1}(a)$. This concludes the proof. □

3.2. ***An example.*** By Lemma 3.1.1, to construct a smooth proper scheme $X \to \mathrm{Spec}(R)$ such that X_K is not stably rational and such that X_k is rational, it is enough to construct a family $X \to B$ over R such that there exist points $b \in B(\mathbb{C}_p)$ and $a \in B(k)$ such that X_b is stably irrational and X_a is rational. Such a family can be constructed directly using the example of Hassett, Pirutka, and Tschinkel [HPT18]; see also [CTS21, Section 12.2.2] and [Sch19]. We give some details.

Let $X' \to Z$ be the universal family of quadric bundles over $\mathbb{P}^2_{\mathbb{Q}}$ given in $\mathbb{P}^2_{\mathbb{Q}} \times \mathbb{P}^3_{\mathbb{Q}}$ by a bihomogeneous form of bidegree $(2, 2)$. After choosing coordinates x, y, z and U, V, W, T on $\mathbb{P}^2_{\mathbb{Q}} \times \mathbb{P}^3_{\mathbb{Q}}$, the variety Z identifies with the space of bihomogeneous forms $F = F(x, y, z, U, V, W, T)$ of bidegree $(2, 2)$ in $\mathbb{P}^2_{\mathbb{Q}} \times \mathbb{P}^3_{\mathbb{Q}}$ that are symmetric quadratic forms in the variables U, V, W, T, since any such F determines a quadric bundle over $\mathbb{P}^2_{\mathbb{Q}}$ via the projection $\mathbb{P}^2_{\mathbb{Q}} \times \mathbb{P}^3_{\mathbb{Q}} \to \mathbb{P}^2_{\mathbb{Q}}$. In turn, these forms are given by a 4×4 symmetric matrix $A = (a_{i,j})_{1 \le i,j \le 4}$ where each entry $a_{i,j} = a_{i,j}(x, y, z)$ is a homogeneous polynomial of degree 2.

By the arguments in [Sch19] and Bertini's theorem, there exists a dense open Zariski $B_{\mathbb{Q}} \subset Z$ such that the restriction of the family $X_{\mathbb{Q}} \to B_{\mathbb{Q}}$ to $B_{\mathbb{Q}}$ parametrizes quadric bundle flat over $\mathbb{P}^2_{\mathbb{Q}}$ and with smooth total space.

By spreading out, this construction extends to give a smooth family $X \to B$ over $\mathbb{Z}[1/n]$ for n large enough whose base change to $\mathbb{Q}$ identifies with $X_{\mathbb{Q}} \to B_{\mathbb{Q}}$. By the main result of [HPT18], the general fiber of $X(\mathbb{C}) \to B(\mathbb{C})$ is not stably rational, hence, for every p, there exists a $b \in B(\mathbb{C}_p)$ such that Y_b is not stably rational. So, we are left to show that for $p \gg 0$, there exists a $a \in B(\overline{\mathbb{F}}_p)$ such that X_a is rational.

Using Bertini, there exists a rational point $r \in B(\overline{\mathbb{Q}})$ such that the corresponding 4×4 symmetric matrix $A = (a_{i,j})_{1 \le i,j \le 4}$ has $a_{1,1} = 0$. By spreading out, we can choose $p \gg 0$ such that $r \in B(\overline{\mathbb{Q}})$ extends to a point $\tilde{a} \in B(\overline{\mathbb{Z}}_p)$ whose reduction a modulo p defines a flat quadric bundle $X_a \to \mathbb{P}^2_{\overline{\mathbb{F}}_p}$ with smooth total space, whose associated matrix has $a_{1,1} = 0$. Since $a_{1,1} = 0$, for every $x \in \mathbb{P}^2_{\overline{\mathbb{F}}_p}$, the point $[1:0:0:0]$ is $k(x)$-rational point of the fiber of $X_a \to \mathbb{P}^2_{\overline{\mathbb{F}}_p}$ in x. In particular the morphism $X_a \to \mathbb{P}^2_{\overline{\mathbb{F}}_p}$ has a rational section, hence X_a is rational. This concludes the construction of a proper smooth scheme over R with rational special fiber and stably irrational generic fiber.

Acknowledgements

The authors are grateful to Stefan Schreieder for having introduced Valloni to the Noether problem, to Giuseppe Ancona for many interesting discussions and comments on a preliminary version of the article, and to Colliot-Thélène for pointing out to us the example in Section 3. The authors also thank the referees for their comments and for having greatly improved both the statement and proof of our main result.

References

[ABBGvB21] A. Auel, A. Bigazzi, C. Böhning, and H.-C. Graf von Bothmer, "Unramified Brauer groups of conic bundle threefolds in characteristic two", *Amer. J. Math.* **143**:5 (2021), 1601–1631. MR

[AM72] M. Artin and D. Mumford, "Some elementary examples of unirational varieties which are not rational", *Proc. Lond. Math. Soc.* (3) **25** (1972), 75–95. MR

[BK85] F. A. Bogomolov and P. I. Katsylo, "Rationality of some quotient varieties", *Mat. Sb.* (*N.S.*) **126(168)**:4 (1985), 584–589. In Russian; translated in *Math. USSR-Sb.* **54**:2 (1986), 571–576. MR

[BMS18] B. Bhatt, M. Morrow, and P. Scholze, "Integral p-adic Hodge theory", *Publ. Math. Inst. Hautes Études Sci.* **128** (2018), 219–397. MR

[Bog88] F. A. Bogomolov, "The Brauer group of quotient spaces by linear group actions", *Izv. Akad. Nauk SSSR Ser. Mat.* **51**:3 (1987), 485–516. In Russian; translated in *Math. USSR-Izv.* **30**:3 (1988), 455–485. MR

[CR11] A. Chatzistamatiou and K. Rülling, "Higher direct images of the structure sheaf in positive characteristic", *Algebra Number Theory* **5**:6 (2011), 693–775. MR

[CTP16] J.-L. Colliot-Thélène and A. Pirutka, "Hypersurfaces quartiques de dimension 3: non-rationalité stable", *Ann. Sci. École Norm. Sup.* (4) **49**:2 (2016), 371–397. MR

[CTS21] J.-L. Colliot-Thélène and A. N. Skorobogatov, *The Brauer–Grothendieck group*, Ergebnisse der Math. (3) **71**, Springer, 2021. MR

[Gas59] W. Gaschütz, "Fixkörper von p-Automorphismengruppen rein-transzendener Körpererweiterungen von p-Charakteristik", *Math. Z.* **71** (1959), 466–468. MR

[HPT18] B. Hassett, A. Pirutka, and Y. Tschinkel, "Stable rationality of quadric surface bundles over surfaces", *Acta Math.* **220**:2 (2018), 341–365. MR

[Ill79] L. Illusie, "Complexe de de Rham–Witt et cohomologie cristalline", *Ann. Sci. École Norm. Sup.* (4) **12**:4 (1979), 501–661. MR

[Kat78] 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, Tokyo, 1978. MR

[Kol95] J. Kollár, "Nonrational hypersurfaces", *J. Amer. Math. Soc.* **8**:1 (1995), 241–249. MR

[Kun54] H. Kuniyoshi, "On purely-transcendency of a certain field", *Tohoku Math. J.* (2) **6** (1954), 101–108. MR

[LS23] C. D. Lazda and A. N. Skorobogatov, "Reduction of Kummer surfaces modulo 2 in the non-supersingular case", *Épijournal Géom. Algébrique* **7** (2023), art. id. 10. MR

[Mil16] J. S. Milne, *Étale cohomology*, Princeton Math. Ser. **33**, Princeton Univ. Press, 1980. MR

[MP12] D. Maulik and B. Poonen, "Néron–Severi groups under specialization", *Duke Math. J.* **161**:11 (2012), 2167–2206. MR

[NO21] J. Nicaise and J. C. Ottem, "A refinement of the motivic volume, and specialization of birational types", pp. 291–322 in *Rationality of varieties*, edited by G. Farkas et al., Progr. Math. **342**, Birkhäuser, Cham, 2021. MR

[Pey08] E. Peyre, "Unramified cohomology of degree 3 and Noether's problem", *Invent. Math.* **171**:1 (2008), 191–225. MR

[Sal84] D. J. Saltman, "Noether's problem over an algebraically closed field", *Invent. Math.* **77**:1 (1984), 71–84. MR

[Sch19] S. Schreieder, "On the rationality problem for quadric bundles", *Duke Math. J.* **168**:2 (2019), 187–223. MR

[Stacks] "The Stacks project", electronic reference, 2005–, available at http://stacks.math.columbia.edu.

[Tot16] B. Totaro, "Hypersurfaces that are not stably rational", *J. Amer. Math. Soc.* **29**:3 (2016), 883–891. MR

[Voi15] C. Voisin, "Unirational threefolds with no universal codimension 2 cycle", *Invent. Math.* **201**:1 (2015), 207–237. MR

Communicated by Jean-Louis Colliot-Thélène
Received 2023-04-02 Revised 2024-11-20 Accepted 2024-12-23

eambrosi@unistra.fr *IRMA (Université de Strasbourg), 67084 Strasbourg, France*

domenico.valloni@epfl.ch *EPFL, 1015 Lausanne, Switzerland*

ALGEBRA AND NUMBER THEORY 19:12 (2025)
https://doi.org/10.2140/ant.2025.19.2369

On the Grothendieck ring of a quasireductive Lie superalgebra

Maria Gorelik, Vera Serganova and Alexander Sherman

Given a Lie superalgebra $\mathfrak{g}$ and a maximal quasitoral subalgebra $\mathfrak{h}$, we consider properties of restrictions of $\mathfrak{g}$-modules to $\mathfrak{h}$. This is a natural generalization of the study of characters in the case when $\mathfrak{h}$ is an even maximal torus. We study the case of $\mathfrak{g} = \mathfrak{q}_n$ with $\mathfrak{h}$ a Cartan subalgebra, and prove several special properties of the restriction in this case, including an explicit realization of the $\mathfrak{h}$-supercharacter ring.

1. Introduction

1.1. ***Maximal toral subalgebras and restriction.*** Let $\mathfrak{g}$ be a Lie superalgebra, not necessarily finite-dimensional. Assume that $\mathfrak{g}$ contains a subalgebra $\mathfrak{h}$ such that

(1) $\mathfrak{t} := \mathfrak{h}_{\bar{0}}$ has diagonalizable adjoint action on $\mathfrak{g}$ with finite-dimensional weight spaces; and,

(2) $\mathfrak{h} = \mathfrak{g}^{\mathfrak{t}}$, where $\mathfrak{g}^{\mathfrak{t}}$ denotes the centralizer of $\mathfrak{t}$ in $\mathfrak{g}$;

We will call subalgebras with the above properties *maximal quasitoral*. Maximal quasitoral subalgebras play an analogous role to maximal toral subalgebras of Lie algebras. In the purely even setting, the restriction of a representation to a maximal toral subalgebra is exactly the data of its character. The character of a representation is a powerful invariant, and provides (nice) formulas for characters of irreducible representations is a central problem in representation theory.

For many Lie superalgebras of interest, maximal quasitoral subalgebras are even, i.e., $\mathfrak{h} = \mathfrak{t}$ (e.g., for $\mathfrak{gl}_{m|n}$, $\mathfrak{osp}_{m|2n}$, $\mathfrak{p}_n$, ... See [16] for the definition of these Lie superalgebras). In this case the restriction

MSC2020: 17B10, 17B20, 17B55, 17D10.
Keywords: Lie superalgebra, Duflo–Serganova functor, queer superalgebra, representation theory, Grothendieck group.

of an irreducible highest weight representation to $\mathfrak{t}$ completely determines it. But when $\mathfrak{h} \neq \mathfrak{t}$, as in the case of the queer Lie superalgebra $\mathfrak{q}_n$, this longer holds, as for certain simple modules L one has $\operatorname{Res}^{\mathfrak{g}}_{\mathfrak{t}} L \cong \operatorname{Res}^{\mathfrak{g}}_{\mathfrak{t}} \Pi L$ and so the restriction map of Grothendieck groups $\mathscr{K}(\mathfrak{g}) \to \mathscr{K}(\mathfrak{t})$ is not injective.

On the other hand the restriction to $\mathfrak{h}$ is fine enough to make such distinctions. Choose a Borel subalgebra $\mathfrak{b}$ of $\mathfrak{g}$ containing $\mathfrak{h}$ (from a triangular decomposition as explained in Section 2.1), and let $\mathcal{C}$ be the full subcategory of $\mathfrak{g}$-modules such that each module in $\mathcal{C}$ is of finite length, has a diagonal action of $\mathfrak{t}$, and is locally finite over $\mathfrak{b}$ (these conditions can be slightly weakened; see Corollary 5.2). Write $\mathcal{M}^s(\mathfrak{h})$ for the category of $\mathfrak{h}$-modules with diagonal $\mathfrak{t}$ action and finite-dimensional weight spaces. Then the restriction functor $\operatorname{Res}^{\mathfrak{g}}_{\mathfrak{h}}$ induces an injective map $\mathscr{K}(\mathcal{C}) \to \mathscr{K}(\mathcal{M}^s(\mathfrak{h}))$, see Corollary 5.2.

It is therefore of interest to understand the $\mathfrak{h}$-character of irreducible $\mathfrak{g}$-modules. A difficulty that arises is that the Grothendieck ring of finite-dimensional $\mathfrak{h}$-modules does not have a simple manifestation as in the case of $\mathfrak{t}$-modules. Nevertheless a description can be given, and its structure is interesting in its own right.

1.2. ***$\mathscr{K}_-(\mathfrak{g})$ and $\mathscr{K}_+(\mathfrak{g})$.*** Let $\mathscr{K}(\mathfrak{g})$ denote the Grothendieck group of finite-dimensional $\mathfrak{g}$-modules, and write M_{gr} for the natural image of a finite-dimensional module M in $\mathscr{K}(\mathfrak{g})$. This ring admits two natural quotients: $\mathscr{K}_+(\mathfrak{g})$, which is obtained by identifying $M_{\mathrm{gr}} = (\Pi M)_{\mathrm{gr}}$, and $\mathscr{K}_-(\mathfrak{g})$, obtained by identifying $M_{\mathrm{gr}} = -(\Pi M)_{\mathrm{gr}}$. There is a natural embedding of $\mathscr{K}(\mathfrak{g})$ into $\mathscr{K}_+(\mathfrak{g}) \times \mathscr{K}_-(\mathfrak{g})$, and this embedding becomes an isomorphism over $\mathbb{Q}$; thus a proper understanding of $\mathscr{K}_+(\mathfrak{g})$ and $\mathscr{K}_-(\mathfrak{g})$ suffices for the understanding of $\mathscr{K}(\mathfrak{g})$ (see Section 5 for a precise relationship between $\mathscr{K}(\mathfrak{g})$ and $\mathscr{K}_+(\mathfrak{g}) \times \mathscr{K}_-(\mathfrak{g})$).

The ring $\mathscr{K}_+(\mathfrak{g})$ behaves like a character ring, and in fact embeds into $\mathscr{K}_+(\mathfrak{t})$ under the restriction map; thus the information it carries is less interesting from our standpoint. On the other hand, $\mathscr{K}_-(\mathfrak{g})$, the *reduced Grothendieck ring*, has a nontrivial, even exotic structure as a ring, and will be our main object of study. The restriction map $\mathscr{K}_-(\mathfrak{g}) \to \mathscr{K}_-(\mathfrak{t})$ is sometimes very far from being an embedding; for instance if $\mathfrak{g} = \mathfrak{q}_n$ the image of any nontrivial finite-dimensional irreducible module is zero by [3]. Thus is it necessary to study instead $\mathscr{K}_-(\mathfrak{g}) \to \mathscr{K}_-(\mathfrak{h})$.

A further advantage of using $\mathscr{K}_-(\mathfrak{g})$ is that the Duflo–Serganova functor, while not being exact, always induces a morphism $ds_x : \mathscr{K}_-(\mathfrak{g}) \to \mathscr{K}_-(\mathfrak{g}_x)$ for appropriate $x \in \mathfrak{g}_{\bar{1}}$. It has been known for some time (see [13]) that for Kac–Moody superalgebras, the map induced by ds_x on supercharacters is given by restriction to $\mathfrak{t}_x$, the Cartan subalgebra of $\mathfrak{g}_x$. This is a reflection of a more general property of ds_x, discussed in Section 8, which shows that ds_x can always be thought of as a restriction map to $\mathfrak{g}_x$. Therefore in our setting the induced map $ds_x : \mathscr{K}_-(\mathfrak{h}) \to \mathscr{K}_-(\mathfrak{h}_x)$ is also given by restriction of modules from $\mathfrak{h}$ to $\mathfrak{h}_x$, where $\mathfrak{h}_x$ is a maximal quasitoral subalgebra of $\mathfrak{g}_x$.

1.3. ***Results for $\mathfrak{g} = \mathfrak{q}_n$.*** Let $\mathfrak{g} = \mathfrak{q}_n$; for two weights λ, μ, write $\lambda \sim \mu$ if $L(\lambda)$ and $L(\mu)$ lie in the same block of $\mathcal{F}in(\mathfrak{q}_n)$. Then for a $\mathfrak{g}$-module M, write $\operatorname{sch}_{\mathfrak{h}} M$ for the natural image of M in $\mathscr{K}_-(\mathcal{F}in\, \mathfrak{h})$.

Theorem 1.1.
$$\operatorname{sch}_{\mathfrak{h}} L(\lambda) = \sum_{\mu \sim \lambda} \operatorname{sch}_{\mathfrak{h}} L(\lambda)_\mu \tag{1}$$

Equivalently, if $\mu \not\sim \lambda$ *then* $[L(\lambda) : C(\mu)] = [L(\lambda) : \Pi C(\mu)]$, *where* $C(\mu)$ *is an irreducible* $\mathfrak{h}$*-module of weight* μ*.*

See Theorem 9.2 for a proof of this statement, and a stronger result. In Section 6 we present a version of the above result which holds for a more general Lie superalgebra, but only for particular blocks, those which are the closest to being typical.

1.3.1. *Supercharacter isomorphism.* Let $\mathcal{C}$ denote the subcategory of $\mathfrak{h}$-modules with weights which appear in some finite-dimensional representation of $\mathfrak{gl}_n = (\mathfrak{q}_n)_{\bar{0}}$. The Weyl group $W = S_n$ has a natural action on $\mathscr{K}_-(\mathcal{C})$ in this case, and we may consider the invariant subalgebra. We prove that $\mathscr{K}_-(\mathcal{C})^W$ has a basis given by $\{a_\lambda\}_{\lambda \in P^+(\mathfrak{g})}$, where $P^+(\mathfrak{g})$ are the dominant weights for $\mathfrak{g} = \mathfrak{q}_n$, and

$$a_\lambda = \sum_{w \in W/\operatorname{Stab}_W(\lambda)} [C(\lambda)^w],$$

where $(-)^w$ denotes the twisting functor.

It is clear that the supercharacter map induces an embedding $\mathrm{sch}_{\mathfrak{h}} : \mathscr{K}_-(\mathfrak{g}) \to \mathscr{K}_-(\mathcal{C})^W$; such a result holds for any quasireductive Lie superalgebra (see Section 7). For $\mathfrak{g} = \mathfrak{q}_n$ we have (see Corollary 9.5):

Theorem 1.2. *The map* $\mathrm{sch}_{\mathfrak{h}} : \mathscr{K}_-(\mathfrak{g})_{\mathbb{Q}} \to \mathscr{K}_-(\mathcal{C})^W_{\mathbb{Q}}$ *is an isomorphism of rings.*

Here the subscript $\mathbb{Q}$ means that we extend scalars to $\mathbb{Q}$. To obtain an isomorphism we only need to invert 2, in fact. This is a consequence of the work in Section 9.

Remark 1.1. The ring $\mathscr{K}_-(\mathfrak{g})$ has a natural basis given by irreducible modules, and an important question is to understand the relation between this basis and the basis $\{a_\lambda\}$.

1.3.2. *Realization of* $\mathscr{K}_-(\mathcal{F}(\mathfrak{q}_n)_{\mathrm{int}})$. Using the above isomorphism, we are able to provide another realization of $\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\mathrm{int}})$, where $\mathcal{F}(\mathfrak{g})_{\mathrm{int}}$ denotes the category of finite-dimensional $\mathfrak{q}_n$-modules with integral weights. To simplify the explanation for the introduction, we will explain this realization over $\mathbb{C}$.

Let $V := \mathbb{C}^{\mathbb{Z}\setminus\{0\}}$ denote the complex vector space with basis $\{v_i\}_{i \in \mathbb{Z}\setminus\{0\}}$. Then write

$$A = \bigwedge V := \bigoplus_{n \in \mathbb{N}} \bigwedge\nolimits^n V.$$

This is a superalgebra, and we write $A_{\bar{0}} = \bigoplus_{n \in \mathbb{N}} \bigwedge^{2n} V$ for the even part. Let $J_k \subseteq A$ be the ideal generated by $\bigwedge^{k+1} V$. For the following, see Theorem 9.3.

Theorem 1.3. *We have an explicit isomorphism of algebras*

$$\mathscr{K}_-(\mathcal{F}(\mathfrak{q}_n)_{\mathrm{int}}) \otimes_{\mathbb{Z}} \mathbb{C} \to (A/J_n)_{\bar{0}}.$$

where, up to a scalar, a_λ *is mapped to* $v_\lambda := v_{j_1} \wedge \cdots \wedge v_{j_k}$, *where* $j_1, \ldots, j_k$ *are the nonzero coordinates of* λ*.*

1.3.3. *Relation to the Duflo–Serganova functor.* For $\mathfrak{q}_n$, the maps ds_x depend only on the *rank* of x, which is a nonnegative half-integer $s \in \frac{1}{2}\mathbb{N}$. Thus we write $ds_x = ds_s$, where $s = \operatorname{rank}(x)$, and this gives a map $ds_s : \mathscr{K}_-(\mathcal{F}(\mathfrak{q}_n)) \to \mathscr{K}_-(\mathcal{F}(\mathfrak{q}_{n-2s}))$. We have the following simple formula for ds_s in terms of the basis $\{a_\lambda\}$, using the previously noted fact that ds_x is given by restriction:

$$\mathrm{ds}_s(a_\mu) = \begin{cases} 0 & \text{if } \operatorname{zero}\mu < 2s \\ a_{\mu'} & \text{if } \operatorname{zero}\mu \geq 2s \end{cases}$$

where $\operatorname{zero}\mu$ is the number of zero coordinates that μ has, and $\mu' \in P^+(\mathfrak{g}_x)$ is such that μ' and μ have the same nonzero coordinates. In [11], we compute ds_s on the basis $\{[L(\lambda)]\}_{\lambda\in P^+(\mathfrak{g})}$ of irreducible modules; remarkably it admits a similarly simple expression:

$$\mathrm{ds}_s([L(\lambda)]) = \begin{cases} 0 & \text{if } \operatorname{zero}\lambda < 2s, \\ [L(\lambda')] & \text{if } \operatorname{zero}\lambda \geq 2s. \end{cases}$$

1.4. *List of notation.*

symbol	§	symbol	§	symbol	§	symbol	§	symbol	§
$\mathcal{M}(\mathfrak{h})$	3.1	$L(\lambda)$	4.1	$M_{\mathrm{gr}}, M_{\mathrm{gr},\pm}$	5.1	Ξ	9.2	$P(\mathfrak{g})$	7
$C(\lambda)$	3.2	$\mathscr{K}(\mathcal{C})$	5.1	I_0, I_1	3.2	Core	9.2	$P(\mathfrak{g})'$	7.1
K_λ	3.1.1	$\mathscr{K}_+(\mathcal{C})$	5.1	$\mathfrak{t}_G$	5.6	$\mathcal{F}(\mathfrak{g})$	7.1	smult	9.5
F_λ	3.1.1	$\mathscr{K}_-(\mathcal{C})$	5.1	$\mathrm{sch}_{\mathfrak{h}}$	5.2	$P^+(\mathfrak{g})$	7	ds_s	9.5

2. Preliminaries: maximal quasitoral subalgebras and Clifford algebras

We work over the field over of complex numbers, $\mathbb{C}$, and denote by $\mathbb{N}$ the set of nonnegative integers. For a super vector space V we write $V = V_{\bar{0}} \oplus V_{\bar{1}}$ for its parity decomposition. Then ΠV will denote the parity-shifted super vector space obtained from V. Let $\delta_V := \delta$ denote the endomorphism of V given by $\delta(v) = (-1)^{p(v)}v$.

We work with Lie superalgebras which are not necessarily finite-dimensional. For a Lie superalgebra $\mathfrak{g}$ we denote by $\mathcal{F}in(\mathfrak{g})$ the category of finite-dimensional $\mathfrak{g}$-modules.

2.1. *Maximal (quasi)toral subalgebras.*

2.1.1. *Definition.* Let $\mathfrak{g}$ be Lie superalgebra. We say that a finite-dimensional subalgebra $\mathfrak{t} \subseteq \mathfrak{g}_{\bar{0}}$ is a *maximal toral* subalgebra if it is commutative, acts diagonally on $\mathfrak{g}$ under the adjoint representation, and we have $\mathfrak{g}_{\bar{0}}^{\mathfrak{t}} = \mathfrak{t}$. In this case we set $\mathfrak{h} := \mathfrak{g}^{\mathfrak{t}}$, and we refer to $\mathfrak{h}$ as a *maximal quasitoral* subalgebra of $\mathfrak{g}$. Observe that $\mathfrak{h}_{\bar{0}} = \mathfrak{t}$.

Denote by $\Delta(\mathfrak{g}) := \Delta \subseteq \mathfrak{t}^*$ the nonzero eigenvalues of $\mathfrak{t}$ in $\operatorname{Ad}\mathfrak{g}$, and write $Q = \mathbb{Z}\Delta$. We will assume throughout that

$$\text{all eigenspaces } \mathfrak{g}_\nu \ (\nu \in \Delta \cup \{0\}) \text{ are finite-dimensional.} \qquad (*)$$

In particular we assume that $\mathfrak{h}$ is finite-dimensional. We have

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Delta} \mathfrak{g}_\alpha.$$

2.1.2. *Triangular decomposition.* We choose a group homomorphism $\gamma : \mathbb{Z}\Delta \to \mathbb{R}$ such that $\gamma(\alpha) \neq 0$ for all $\alpha \in \Delta$. Such a homomorphism exists because $\mathfrak{h} \cong \mathbb{R}$ as a $\mathbb{Q}$-vector space. We introduce the triangular decomposition $\Delta(\mathfrak{g}) = \Delta^+(\mathfrak{g}) \coprod \Delta^-(\mathfrak{g})$, with

$$\Delta^{\pm}(\mathfrak{g}) := \{\alpha \in \Delta(\mathfrak{g}) \mid \pm\gamma(\alpha) > 0\},$$

and define a partial order on $\mathfrak{t}^*$ by

$$\lambda \geq \nu \ \text{ if } \ \nu - \lambda \in \mathbb{N}\Delta^-.$$

We set $\mathfrak{n}^{\pm} := \oplus_{\alpha\in\Delta^{\pm}}\mathfrak{g}_\alpha$ and call a subalgebra of the form $\mathfrak{b} := \mathfrak{h} \oplus \mathfrak{n}^+$ a Borel subalgebra. We further assume throughout that

$$\mathcal{U}(\mathfrak{n}^-)_\nu \text{ is finite-dimensional for all } \nu \in Q. \tag{**}$$

Remark 2.1. Several notions of triangular decompositions for Lie superalgebras have appeared in the literature. In [18] and [20] for example, a notion of positive roots arose from the choice of a generic hyperplane in $\mathfrak{h}^*$. Our approach generalizes these approaches and admits a more flexible definition. One can construct finite-dimensional Lie superalgebras for which our definition gives rise to more triangular decompositions as compared to [18] and [20]; for example, consider $\mathfrak{g}$ with $\mathfrak{g}_{\bar{0}}$ one-dimensional acting by real, $\mathbb{Q}$-linearly independent characters on $\mathfrak{g}_{\bar{1}}$.

On the other hand, for simple, finite-dimensional Lie superalgebras our notion of triangular decomposition agrees with that of [20].

2.1.3. *Examples.*

- If $\mathfrak{g}$ is a Kac–Moody superalgebra, then, by [21] any maximal toral subalgebra $\mathfrak{t}$ satisfying $(*)$ is Cartan subalgebra of $\mathfrak{g}_{\bar{0}}$; one has $\mathfrak{h} = \mathfrak{t}$.
- If $\mathfrak{g}$ is a quasireductive Lie superalgebra ($\dim \mathfrak{g} < \infty$, $\mathfrak{g}_{\bar{0}}$ is reductive and $\mathfrak{g}_{\bar{1}}$ is a semisimple $\mathfrak{g}$-module), then a maximal toral (resp. quasitoral) subalgebra $\mathfrak{t}$ is a Cartan subalgebra of $\mathfrak{g}_{\bar{0}}$ (resp. $\mathfrak{g}$). In both this example and the former, $\mathfrak{t}$ and $\mathfrak{h}$ are unique up to a conjugation by inner automorphism, see [21].
- If we fix an invariant form on a quasireductive Lie superalgebra $\mathfrak{g}$ (which can be the zero form), we can construct the affinization $\hat{\mathfrak{g}}$ with $\hat{\mathfrak{t}} = \mathfrak{t} + \mathbb{C}K + \mathbb{C}d$ and $\hat{\mathfrak{h}} = \mathfrak{h} + \mathbb{C}K + \mathbb{C}d$.
- The cases when $\mathfrak{t} \neq \mathfrak{h}$ include the queer Lie superalgebras and their affinizations.

2.2. ***Clifford algebras.*** For a vector space V with a symmetric (not necessarily nondegenerate) bilinear form F, let $\mathcal{C}\ell(V, F)$ denote the corresponding Clifford algebra. This is a superalgebra where the elements of V are declared to be odd. Write $K \subseteq V$ for the kernel of F, so that F induces a nondegenerate form on V/K, which we also write as F. We have an isomorphism of superalgebras

$$\mathcal{C}\ell(V, F) \cong \mathcal{C}\ell(V/K, F) \otimes \textstyle\bigwedge^{\bullet} K.$$

The superalgebra $\mathcal{C}\ell(V, F)$ is semisimple if and only if F is nondegenerate.

2.2.1. Suppose F is nondegenerate, and write m for the dimension of V. We have

$$\mathcal{C}\ell(V, F) \cong Q(2^{n+1}) \quad \text{if } m = 2n+1, \quad \mathcal{C}\ell(V, F) \cong \operatorname{End}(\mathbb{C}^{2^{n-1}|2^{n-1}}) \quad \text{if } m = 2n.$$

Here $Q(r)$ is the queer superalgebra, i.e., the associative subalgebra of $\operatorname{End}(\mathbb{C}^r|\mathbb{C}^r)$ consisting of the matrices of the form $\left(\begin{smallmatrix} A & B \\ B & A \end{smallmatrix}\right)$.

It follows that $\mathcal{C}\ell(V, F)$ is a simple superalgebra and admits a unique, parity invariant, irreducible module when m is odd, while if m is even there are two irreducible modules that differ by parity. Moreover, all ($\mathbb{Z}_2$-graded) $\mathcal{C}\ell(V, F)$-modules are completely reducible. For $m \neq 0$, if E is an irreducible representation of $\mathcal{C}\ell(V, F)$ one has

$$\dim E_{\bar{0}} = \dim E_{\bar{1}} = 2^{\lfloor \frac{m-1}{2} \rfloor}.$$

2.2.2. For a nonnegative integer m, we write $\mathcal{C}\ell(m)$ for the Clifford algebra $\mathcal{C}\ell(\mathbb{C}^m, F)$, where F is the standard nondegenerate symmetric bilinear form on $\mathbb{C}^m$. Clearly $\mathcal{C}\ell(V, F) \cong \mathcal{C}\ell(m)$ if $\dim V = m$ and F is nondegenerate.

2.2.3. Let A_1, A_2 be associative superalgebras and V_i be A_i-modules; we define the outer tensor product $V_1 \boxtimes V_2$ as the space $V_1 \otimes V_2$ endowed by the $A_1 \otimes A_2$-action

$$(a_1, a_2)(v_1 \otimes v_2) := (-1)^{p(a_2)p(v_1)} a_1 v_1 \otimes a_2 v_2.$$

One has $\mathcal{C}\ell(m) \otimes \mathcal{C}\ell(n) \cong \mathcal{C}\ell(m+n)$; if V_1 (resp., V_2) are simple modules over $\mathcal{C}\ell(m_1)$ (resp., $\mathcal{C}\ell(m_2)$), then $V_1 \boxtimes V_2$ is a simple if either m_1 or m_2 is even, and if m_1 and m_2 are both odd then $V_1 \otimes V_2 = L \oplus \Pi L$, where L is simple over $\mathcal{C}\ell(n_1 + m_1)$.

2.3. ***Realization of the irreducible representation of $\mathcal{C}\ell(2n)$.*** Consider $\mathbb{C}^{2n}$ with standard basis $e_1, \dots, e_n$, $f_1, \dots, f_n$, equipped with the symmetric nondegenerate form $(-, -)$ satisfying:

$$(e_i, f_j) = \delta_{ij}, \quad (e_i, e_j) = (f_i, f_j) = 0.$$

Consider the polynomial superalgebra $L = \mathbb{C}[\xi_1, \dots, \xi_n]$ with odd generators $\xi_1, \dots, \xi_n$. Then we may realize L as an irreducible representation of $\mathcal{C}\ell(2n)$ via $e_i \mapsto \xi_i$ and $f_j \mapsto \partial_{\xi_j}$, i.e., e_i acts by multiplication by ξ_i and f_j acts by the derivation sending $\xi_i \mapsto \delta_{ij}$. In this way we have defined a surjective morphism $\mathcal{C}\ell(2n) \to \operatorname{End}(L)$ (in fact it is an isomorphism). Every irreducible representation of $\mathcal{C}\ell(2n)$ is isomorphic to L or ΠL.

2.3.1. Continuing with the setup from Section 2.3, if we choose W a maximal isotropic subspace of $\mathbb{C}^{2n}$, then $\bigwedge^\bullet W$ acts on L, and under this action L is isomorphic to the exterior algebra of W under left multiplication. Thus given two arbitrary irreducible representations of $\mathcal{C}\ell(V)$, they are isomorphic if and only if the parities of the one-dimensional W-invariant subspaces of each are the same.

2.4. ***The case $\mathcal{C}\ell(2n+1)$.*** Next we look at $\mathbb{C}^{2n+1}$ with nondegenerate symmetric form $(-,-)$ and basis $e_1,\dots,e_n, f_1,\dots,f_n, g$. Here the inner product relations for the e_i's and f_i's are the same as in Section 2.3, along with

$$(g,e_i)=(g,f_j)=0,\quad (g,g)=2.$$

Consider the exterior algebra $L=\mathbb{C}[\xi_1,\dots,\xi_{n+1}]$. Then we may realize L as the unique irreducible representation of $\mathcal{C}\ell(2n+1)$ by $e_i\mapsto\xi_i$, $f_j\mapsto\partial_{\xi_j}$, and $g\mapsto\xi_{n+1}+\partial_{\xi_{n+1}}$.

2.5. ***The operator T.*** Choose an orthonormal basis $H_1,\dots,H_m$ of $\mathbb{C}^m$, and let $T=H_1\cdots H_m\in\mathcal{C}\ell(m)$. This operator is an eigenvector of $O(m)$ with weight given by the determinant. Thus it is well-defined up to an orientation on V.

Let V be an irreducible representation of $\mathcal{C}\ell(m)$. If m is odd, then $\Pi V\cong V$, and $\mathrm{End}_{\mathbb{C}}(V)=\mathcal{C}\ell(m)\oplus\mathcal{C}\ell(m)\delta_V$ (here we consider all endomorphisms), where $\delta_V(v)=(-1)^{p(v)}v$. In this case, $T=\phi\delta_V$, where ϕ is an odd $\mathcal{C}\ell(m)$-equivariant automorphism of V. If m is even, then $V\ncong\Pi(V)$ and $\mathcal{C}\ell(m)=\mathrm{End}(V)$. In this case, $T=(-1)^n\delta_V\in\mathrm{End}(V)$, where $\dim V=2n$.

3. Representation theory of quasitoral Lie superalgebras

3.1. ***$\mathfrak{h}$-modules.*** Take $\mathfrak{h}$ as in Section 2.1: $\mathfrak{h}$ is a finite-dimensional Lie superalgebra with

$$[\mathfrak{t},\mathfrak{h}]=0,$$

where $\mathfrak{t}=\mathfrak{h}_{\bar{0}}$. We call Lie superalgebras of this form quasitoral. For a semisimple $\mathfrak{t}$-module N, and $\nu\in\mathfrak{t}^*$, write N_ν for the ν-weight space in N.

Denote by $\mathcal{M}(\mathfrak{h})$ the full subcategory of $\mathfrak{h}$-modules N with diagonal action of $\mathfrak{t}$ and finite-dimensional weight spaces N_ν. We set $\mathcal{F}(\mathfrak{h})$ to be the full subcategory of $\mathcal{M}(\mathfrak{h})$ consisting of those modules which are finite-dimensional. The simple modules in $\mathcal{M}(\mathfrak{h})$ and $\mathcal{F}(\mathfrak{h})$ coincide. In this section we study the questions of restriction, tensor product, and extensions of simple modules in $\mathcal{M}(\mathfrak{h})$.

We denote by σ the antiautomorphism of $\mathcal{U}(\mathfrak{h})$ induced by the antipode $-\mathrm{Id}|_{\mathfrak{h}}$ (recall that antiautomorphism means $\sigma(ab)=(-1)^{p(a)p(b)}\sigma(b)\sigma(a)$). This map induces the standard duality $*$ on $\mathcal{F}(\mathfrak{h})$.

3.1.1. View $\mathcal{U}(\mathfrak{h})$ as a Clifford algebra over the polynomial algebra $\mathcal{S}(\mathfrak{t})$; the corresponding symmetric bilinear form is given on $\mathfrak{h}_{\bar{1}}$ by the formula $F(H,H')=[H,H']$ (see Appendix in [9] for details).

For each $\lambda\in\mathfrak{t}^*$, the evaluation of F at λ gives a symmetric form $F_\lambda:(H,H')\mapsto\lambda([H,H'])$. We denote by $\mathrm{rk}\,F_\lambda$ the rank of this form. For each $\lambda\in\mathfrak{t}^*$ we consider the Clifford algebra

$$\mathcal{C}\ell(\lambda):=\mathcal{C}\ell(\mathfrak{h}_{\bar{1}},F_\lambda)=\mathcal{U}(\mathfrak{h})/\mathcal{U}(\mathfrak{h})I(\lambda),$$

where $I(\lambda)$ stands for the kernel of the algebra homomorphism $\mathcal{S}(\mathfrak{t})\to\mathbb{C}$ induced by λ. We will write $K_\lambda\subseteq\mathfrak{h}_{\bar{1}}$ for the kernel of F_λ. Then we have an isomorphism of superalgebras

$$\mathcal{C}\ell(\lambda)\cong\mathcal{C}\ell(\mathrm{rank}\,F_\lambda)\otimes\bigwedge K_\lambda. \tag{2}$$

Denote by $\phi_\lambda : \mathcal{U}(\mathfrak{h}) \to \mathcal{C}\ell(\lambda)$ the canonical epimorphism, and $p_\lambda : \mathcal{C}\ell(\lambda) \to \mathcal{C}\ell(\lambda)/(K_\lambda)$ for the projection. Let $\varphi : \mathfrak{h} \to \mathfrak{h}$ be an automorphism of Lie superalgebras, and write also φ for the corresponding automorphism of $\mathcal{U}(\mathfrak{h})$. Then for every λ, φ induces isomorphisms of algebras

$$\varphi_\lambda : \mathcal{C}\ell(\lambda) \to \mathcal{C}\ell(\varphi^{-1}(\lambda)), \quad \overline{\varphi_\lambda} : \mathcal{C}\ell(\lambda)/K_\lambda \to \mathcal{C}\ell(\varphi^{-1}(\lambda))/(K_{\varphi^{-1}(\lambda)}).$$

We have the commutative diagram

$$\begin{tikzcd}
\mathcal{U}(\mathfrak{h}) \arrow[r, "\varphi"] \arrow[d, "\phi_\lambda"] & \mathcal{U}(\mathfrak{h}) \arrow[d, "\phi_{\varphi^{-1}(\lambda)}"] \\
\mathcal{C}\ell(\lambda) \arrow[r, "\varphi_\lambda"] \arrow[d, "p_\lambda"] & \mathcal{C}\ell(\varphi^{-1}(\lambda)) \arrow[d, "p_{\varphi^{-1}(\lambda)}"] \\
\mathcal{C}\ell(\lambda)/(K_\lambda) \arrow[r, "\overline{\varphi_\lambda}"] & \mathcal{C}\ell(\varphi^{-1}(\lambda))/(K_{\varphi^{-1}(\lambda)})
\end{tikzcd} \tag{3}$$

For the anti-involution σ we also have the same diagram as above, where the induced maps $\sigma_\lambda, \overline{\sigma_\lambda}$ are anti-algebra isomorphisms.

Lemma 3.1. (1) *A $\mathcal{C}\ell(\lambda)$-module N is semisimple if and only if* $\operatorname{Ann}_{\mathfrak{h}_{\bar{1}}} N = K_\lambda$.

(2) *If N is an indecomposable $\mathcal{C}\ell(\lambda)$-module of length* 2, *then* $[N : C(\lambda)] = [N : \Pi(C(\lambda))]$.

Proof. These follow from formula (2). ☐

3.1.2. *Examples.*

- If $\mathfrak{h}$ is quasitoral such that $\mathfrak{h}_{\bar{1}}$ is commutative, then F is the zero form.
- For $\mathfrak{g} = \mathfrak{q}_n$ and $\mathfrak{h}$ a maximal quasitoral subalgebra, one has $\dim \mathfrak{h} = (n|n)$ and $\mathfrak{h} \cong \mathfrak{q}_1 \times \cdots \times \mathfrak{q}_1$. Thus in this case the form F is diagonal, and $\operatorname{rk} F_\lambda$ is the number of nonzero entries of λ under the decomposition $\mathfrak{t} \cong (\mathfrak{q}_1)_{\bar{0}} \times \cdots \times (\mathfrak{q}_1)_{\bar{0}}$.
- For $\mathfrak{g} = \mathfrak{sq}_n$ and $\mathfrak{h}$ a maximal quasitoral subalgebra, one has $\dim \mathfrak{h} = (n|n-1)$. In this case F is not diagonal.

3.2. ***Irreducible $\mathfrak{h}$-modules.*** The irreducible $\mathfrak{h}$-modules all arise from irreducible modules over $\mathcal{C}\ell(\lambda)$ for some $\lambda \in \mathfrak{t}^*$. We denote by $C(\lambda)$ a simple $\mathcal{C}\ell(\lambda)$-module and also view it as an $\mathfrak{h}$-module. For $\lambda = 0$ we fix the grading by taking $C_0 = \mathbb{C}$; for all other values of λ we fix a grading in an arbitrary way until further notice. By the above,

$$\begin{aligned} &\dim C(\lambda) = 2^{n_\lambda}, \quad \text{where } n_\lambda := \left\lfloor \frac{\operatorname{rank} F_\lambda + 1}{2} \right\rfloor, \\ &\{u \in \mathfrak{h}_{\bar{1}} \mid u C(\lambda) = 0\} = K_\lambda. \end{aligned} \tag{4}$$

Set

$$I_i = \{\lambda \in \mathfrak{t}^* : \operatorname{rank} F_\lambda \equiv i \bmod 2\} \text{ for } i = 0, 1. \tag{5}$$

Then $C(\lambda) \cong \Pi C(\lambda)$ if and only if $\lambda \in I_1$; we will often use the notation $\Pi^{(\operatorname{rank} F_\lambda)/2}C(\lambda)$, where $\Pi^{(\operatorname{rank} F_\lambda)/2}C(\lambda) \cong C(\lambda) \cong \Pi C(\lambda)$ whenever $\operatorname{rank} F_\lambda$ is odd, and if $\operatorname{rank} F_\lambda$ is even it has the obvious meaning.

3.3. ***Blocks of $\mathcal{M}(\mathfrak{h})$ and $\mathcal{F}(\mathfrak{h})$.*** The blocks of both $\mathcal{M}(\mathfrak{h})$ and $\mathcal{F}(\mathfrak{h})$ are parametrized up to parity shift by $\lambda \in \mathfrak{t}^*$, as follows: If $\operatorname{corank} F_\lambda > 0$, then there is one block of both $\mathcal{M}(\mathfrak{h})$ and $\mathcal{F}(\mathfrak{h})$ on which $\mathfrak{t}$ acts by the character λ; in both cases this block is equivalent to the category of finite-dimensional modules over $\bigwedge K_\lambda$. If $\operatorname{corank} F_\lambda = 0$ then for both categories there is one (resp., two) block(s) of $\mathfrak{h}$ on which $\mathfrak{t}$ acts by λ when $\lambda \in I_1$ (resp., $\lambda \in I_0$), and the block(s) is (are) semisimple.

3.3.1. *Remark.* One can think of $\operatorname{corank} F_\lambda$ as the "atypicality" of its corresponding block. In particular $\operatorname{corank} F_\lambda = 0$ if and only if the block is semisimple and thus its objects are projective in $\mathcal{M}(\mathfrak{h})$, and in general the block corresponding to λ is equivalent to modules over a Grassmann algebra on $\operatorname{corank}(F_\lambda)$-many variables.

3.4. ***The operator $T_\mathfrak{h}$.*** Let $H_1, \dots, H_n$ be a basis of $\mathfrak{h}_{\bar{1}}$, and define

$$T_\mathfrak{h} = \{H_1, \{\cdots\{H_n, 1\}\cdots\} \in \mathcal{U}(\mathfrak{h})$$

where $\{x, y\} = xy + (-1)^{\overline{xy}}yx$ denotes the super anticommutator in $\mathcal{U}(\mathfrak{h})$. It is known that up to a scalar, $T_\mathfrak{h}$ does not depend on the choice of a basis; see [8]. This operator anticommutes with $\mathfrak{h}_{\bar{1}}$, so the image of a submodule under $T_\mathfrak{h}$ remains a submodule.

3.4.1. *Action of $T_\mathfrak{h}$ on simples.* The action of $T_\mathfrak{h}$ on simple $\mathfrak{h}$-modules is deduced from Section 2.5, and is as follows. If $\operatorname{corank} F_\lambda > 0$, then $T_\mathfrak{h}$ acts by 0. If $\operatorname{corank} F_\lambda = 0$, then $T_\mathfrak{h}$ acts by an automorphism, although not $\mathfrak{h}$-equivariantly. If n is odd, then $T_\mathfrak{h}$ is a nonzero multiple of $\delta\phi$, where ϕ is an $\mathfrak{h}$-equivariant odd automorphism. If n is even, $T_\mathfrak{h}$ is a nonzero multiple of δ.

In particular when n is even, T acts on $C(\lambda)$ by an operator of the form

$$a(\lambda)\operatorname{Id}_{(C(\lambda))_{\bar{0}}} \oplus (-a(\lambda))\operatorname{Id}_{(C(\lambda))_{\bar{1}}},$$

for a scalar $a(\lambda)$. Thus T distinguishes between $C(\lambda)$ and its parity shift for projective irreducible modules.

3.4.2. *Action of $T_\mathfrak{h}$ on all of $\mathcal{M}(\mathfrak{h})$.* Let $\operatorname{corank} F_\lambda > 0$. Then the injective hull of $C(\lambda)$ is given by the $\mathcal{C}\ell(\lambda)$-module $I(C(\lambda)) = C(\lambda) \otimes \bigwedge^\bullet K_\lambda$. We claim that

(1) $T_\mathfrak{h}$ annihilates the radical of $I(C(\lambda))$;

(2) $\operatorname{Im} T_\mathfrak{h} = C(\lambda) = \operatorname{socle} I(C(\lambda))$.

In other words, $T_\mathfrak{h}$ acts by taking the head of this module to its socle. It follows that we understand completely the action of $T_\mathfrak{h}$ on every module in $\mathcal{M}(\mathfrak{h})$.

To prove our claim, choose a basis $f_1, \dots, f_r$ of K_λ and extend it to a basis $e_1, \dots, e_s, f_1, \dots, f_r$ of $\mathfrak{h}_{\bar{1}}$ so that $F_\lambda(e_i, e_j) = \delta_{ij}$. Then the image of $T_\mathfrak{h}$ in $\mathcal{C}\ell(\mathfrak{h})$, up to scalar, is given by $e_1 \cdots e_s f_1 \cdots f_r$. By considering the action of this operator, the statement is clear.

3.4.3. *Remark.* It is possible to uniformly choose the parity of irreducible $\mathfrak{h}$-modules as follows. We consider a linear order on $\mathbb{C}$ given by $c_2 > c_1$ if the real part of $c_2 - c_1$ is positive or the real part is zero and the imaginary part is positive. We fix any function

$$t : \mathfrak{t}^* \to \{c \in \mathbb{C} \mid c > 0\}.$$

Retain the notation from Section 3.1.1. For each λ we choose $T_\lambda \in \mathcal{U}(\mathfrak{h})$ in such a way that

$$\overline{T}_\lambda = p_\lambda \circ \phi_\lambda(T_\lambda) \in \mathcal{C\ell}(\lambda)/K_\lambda$$

is an anticentral element satisfying $\overline{T}_\lambda^2 = t(\lambda)^2$ (the element $\overline{T}_\lambda$ is unique up to sign).

If $\lambda \in I_0$, then T_λ is even and it acts on $C(\lambda)$ by a nonzero superconstant $\pm t(\lambda)$ and we fix a grading on $C(\lambda)$ by taking

$$(C(\lambda))_{\overline{0}} := \{v \in C(\lambda) \mid T_\lambda v = t(\lambda)v\}.$$

3.4.4. *Dualities in $\mathcal{F}(\mathfrak{h})$.* Fix $\lambda \in \mathfrak{t}^*$. The category $\mathcal{F}(\mathfrak{h})$ has the duality $*$ induced by σ, and another contragredient involution $(-)^\# : \mathcal{F}(\mathfrak{h}) \to \mathcal{F}(\mathfrak{h})$ induced by the antiautomorphism $\sigma'(a) = a$ for $a \in \mathfrak{h}_{\overline{0}}$ and $\sigma'(a) = \sqrt{-1}a$ for $a \in \mathfrak{h}_{\overline{1}}$. Note that σ' induces an anti-involution on $\mathcal{C\ell}(\lambda)$.

The element $\overline{T}_\lambda$ can be written as the product $H'_1 \dots H'_k$, where $H'_1, \dots, H'_k$ is a lift of a basis of $\mathcal{C\ell}(\lambda)/K_\lambda$ satisfying $[H'_i, H'_j] = 0$ for $i \neq j$. Therefore $\overline{T}_\lambda = (-1)^{(\operatorname{rank} F_\lambda)/2}\sigma'(\overline{T}_\lambda)$ for $\lambda \in I_0$; this gives the following useful formula

$$C(\lambda)^\# \cong \Pi^{(\operatorname{rank} F_\lambda)/2} C(\lambda); \tag{6}$$

which was first established in [7], Lemma 7.

Since $\overline{\sigma_\lambda} : \mathcal{C\ell}(\lambda)/K_\lambda \to \mathcal{C\ell}(-\lambda)/K_{-\lambda}$ is an anti-isomorphism we have $\dfrac{\overline{\sigma_\lambda}(\overline{T}_\lambda)}{t(\lambda)} = (-1)^i \dfrac{\overline{T}_{-\lambda}}{t(-\lambda)}$ for some $i \in \{0, 1\}$, and correspondingly $C(\lambda)^* \cong \Pi^i C(-\lambda)$.

3.5. ***Restriction to quasitoral subalgebra.*** Given a quasitoral subalgebra $\mathfrak{h}' \subseteq \mathfrak{h}$, we have $\mathfrak{t}' = \mathfrak{h}'_{\overline{0}} \subseteq \mathfrak{h}_{\overline{0}} = \mathfrak{t}$. Thus we have a natural restriction $\mathfrak{t}^* \to (\mathfrak{t}')^*$, and therefore we consider weights $\lambda \in \mathfrak{t}^*$ as defining weights in $(\mathfrak{t}')^*$ naturally. We write F'_λ for the bilinear form induced on $\mathfrak{h}'_{\overline{1}}$ by a given weight $\lambda \in \mathfrak{t}^*$, which is exactly the restriction of F_λ to $\mathfrak{h}'_{\overline{1}}$.

Let $\mathcal{C\ell}'(\lambda)$ be the subalgebra of $\mathcal{C\ell}(\lambda)$ which is generated by $\mathfrak{h}'_{\overline{1}}$; clearly, $\mathcal{C\ell}'(\lambda) = \mathcal{C\ell}(\mathfrak{h}'_{\overline{1}}, F'_\lambda)$. Denote by $E'(\lambda)$ a simple $\mathcal{C\ell}'(\lambda)$-module.

Proposition 3.1. *Write $V' = \mathfrak{h}'_{\overline{1}}$ and $V = \mathfrak{h}_{\overline{1}}$.*

(i) *$C(\lambda)$ is simple over $\mathcal{C\ell}'(\lambda)$ if and only if* $\left\lfloor \frac{\operatorname{rank} F_\lambda + 1}{2} \right\rfloor = \left\lfloor \frac{\operatorname{rank} F'_\lambda + 1}{2} \right\rfloor$.

(ii) *$C(\lambda)$ is semisimple over $\mathcal{C\ell}'(\lambda)$ if and only if* $\operatorname{Ker} F'_\lambda = V' \cap \operatorname{Ker} F_\lambda$.

(iii) *If* $\operatorname{rank} F_\lambda \neq \operatorname{rank} F'_\lambda$, *then* $[C(\lambda) : E'_\lambda] = [C(\lambda) : \Pi E'_\lambda]$.

Proof. Case (i) follows from (4) and case (ii) follows from part (ii) of Lemma 3.1.

For (iii) assume that $\operatorname{rank} F_\lambda \neq \operatorname{rank} F'_\lambda$. Substituting $\mathfrak{h}$ by $\mathfrak{h}/\operatorname{Ker} F_\lambda$ we may assume that $\operatorname{Ker} F_\lambda = 0 \neq \operatorname{Ker} F'_\lambda$. Then $\mathfrak{h}_{\bar{1}}$ admits a basis $H'_1, \ldots, H'_m$ such that V' is spanned by $H_1, \ldots, H'_p$ and the matrix of F_λ takes the form

$$\begin{pmatrix} \mathrm{Id}_{p-s} & 0 & 0 & 0 \\ 0 & 0 & \mathrm{Id}_s & 0 \\ 0 & \mathrm{Id}_s & 0 & 0 \\ 0 & 0 & 0 & \mathrm{Id}_k \end{pmatrix}$$

with $k+s+p=m$, $k, s \geq 0$ and $k+s>0$ (since $V' \neq V$). Note that $\operatorname{Ker} F'_\lambda$ is spanned by $H_{p+1}, \ldots, H'_{p+s}$ and so E'_λ is a simple $\mathcal{C}\ell(p-s)$-module. If $k \neq 0$, then the action of H'_k to $C(\lambda)$ is an odd involutive $\mathcal{C}\ell'(\lambda)$-homomorphism, so $\operatorname{Res}_{\mathcal{C}\ell'(\lambda)} C(\lambda)$ is Π-invariant. Consider the remaining case $k=0$. Then $s \neq 0$ and $\mathcal{C}\ell(\lambda) = \mathcal{C}\ell(p-s) \otimes \mathcal{C}\ell(2s)$. Using 2.2.3 we get $C(\lambda) \cong E'_\lambda \boxtimes E''$, where E'' is a simple $\mathcal{C}\ell(2s)$-module. By the above, $\dim E''_{\bar{0}} = \dim E''_{\bar{1}}$. Hence $\operatorname{Res}_{\mathcal{C}\ell(p-s)} C(\lambda)$ is Π-invariant, that is

$$[C(\lambda) : E'_\lambda] = [C(\lambda) : \Pi E'_\lambda]$$

as required. This establishes (iii). □

3.6. ***Tensor product of irreducible $\mathfrak{h}$-modules.*** Let $\lambda, \mu \in \mathfrak{t}^*$. We compute $C(\lambda) \otimes C(\mu)$. Observe that $C(\lambda) \otimes C(\mu)$ is naturally a module over $\mathcal{C}\ell(\mathfrak{h}_{\bar{1}}/K_\lambda \cap K_\mu, F_{\lambda+\mu})$ and $K_\lambda \cap K_\mu \subseteq K_{\lambda+\mu}$. Set

$$K_{\lambda,\mu} := K_{\lambda+\mu}/(K_\lambda \cap K_\mu).$$

We have an isomorphism of superalgebras

$$\mathcal{C}\ell(\mathfrak{h}_{\bar{1}}/K_\lambda \cap K_\mu, F_{\lambda+\mu}) \cong \mathcal{C}\ell(\mathfrak{h}_{\bar{1}}/K_{\lambda+\mu}, F_{\lambda+\mu}) \otimes \bigwedge K_{\lambda,\mu}.$$

Lemma 3.2. *$C(\lambda) \otimes C(\mu)$ is projective over $\mathcal{C}\ell(\mathfrak{h}_{\bar{1}}/(K_\lambda \cap K_\mu), F_{\lambda+\mu})$.*

Proof. It suffices to show that $\bigwedge K_{\lambda,\mu}$ acts freely. Let $v \in K_{\lambda,\mu}$. Then without loss of generality $v \notin K_\lambda$, so the subalgebra generated by v acts projectively on $C(\lambda)$, and thus also on the tensor product. The statement now follows from facts about the representation theory of exterior algebras (see [1]). □

3.6.1. Notice that the unique (up to parity) indecomposable projective module P over $\bigwedge K_{\lambda,\mu}$ is the free module of rank 1. Thus we have shown that $C(\lambda) \otimes C(\mu)$ is a sum of modules of the form $(\Pi)C(\lambda+\mu) \otimes \bigwedge K_{\lambda,\mu}$.

If the rank of either F_λ or F_μ is odd, or the rank of $F_{\lambda+\mu}$ is odd, then the tensor product $C(\lambda+\mu) \otimes P$ is parity invariant, so the explicit decomposition of $C(\lambda) \otimes C(\mu)$ is the appropriate number of copies of $C(\lambda+\mu) \otimes P$ and its parity shift, according to a dimension count.

3.6.2. Thus let us suppose that $\operatorname{rank} F_\lambda$, $\operatorname{rank} F_\mu$, and $\operatorname{rank} F_{\lambda+\mu}$ are all even and we have $\operatorname{rank} F_\nu = 2n_\nu$ for $\nu = \lambda, \mu, \lambda+\mu$.

By Section 2.3, we may realize $C(\lambda)$ as $k[\xi_1, \ldots, \xi_n]$ and $C(\mu)$ as $k[\eta_1, \ldots, \eta_m]$, so that

$$C(\lambda) \otimes C(\mu) = k[\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_m].$$

Now choose a maximal isotropic subspace U for $F_{\lambda+\mu}$ in $\mathfrak{h}_{\bar{1}}$. Then the parity decomposition of $C(\lambda)\otimes C(\mu)$ is described by the U-invariants on this space. Let $u\in U$. Then because u^2 acts trivially on this module, its action on the tensor product is given by (using Section 2.3)

$$u\mapsto\sum_{i=1}^{n}a_iX_i+\sum_{j=1}^{m}b_jY_j,$$

where $a_i,b_j\in\mathbb{C}$, and $X_i\in\{\xi_i,\partial_{\xi_i}\}$ and $Y_j\in\{\eta_j,\partial_{\eta_j}\}$. Further, because $[U,U]$ acts trivially, we can choose $X_1,\dots,X_n$ and $Y_1,\dots,Y_m$ uniformly so that every element of U acts in the way described for u, with potentially different coefficients a_i and b_j. Now, suppose that $X_i=\xi_i$ for some i. Define an odd, linear automorphism s_i of $k[\xi_1,\dots,\xi_n]$ as follows. For $J=\{i_1,\dots,i_{|J|}\}\subseteq\{1,\dots,n\}\setminus\{i\}$, write $\xi_J=\xi_{i_1}\cdots\xi_{i_{|j|}}$, and then set

$$s_i(\xi_J)=\xi_i\xi_J,\qquad s_i(\xi_i\xi_J)=\xi_J.$$

Then under this automorphism, multiplication by ξ_i becomes ∂_{ξ_i} and vice versa, while for $i\neq j$, ∂_{ξ_j} and multiplication by ξ_j become negative themselves. Using this automorphism, we may instead assume that $X_i=\partial_{\xi_i}$, and in this way we may assume that $X_i=\partial_{\xi_i}$ for all i, and $Y_j=\partial_{\eta_j}$ for all j. Thus U acts by a subspace of constant coefficient vector fields on $k[\xi_1,\dots,\xi_n,\eta_1,\dots,\eta_m]$. Write $Z\subseteq\langle\xi_1,\dots,\xi_n,\eta_1,\dots,\eta_m\rangle$ for the invariants of U in this subspace. Then

$$(C(\lambda)\otimes C(\mu))^U=\textstyle\bigwedge^\bullet Z.$$

Thus we have shown the following, still with $n_\lambda:=\left\lfloor\frac{\operatorname{rank}F_\lambda+1}{2}\right\rfloor$:

Theorem 3.1. *If $n_{\lambda+\mu}+\dim K_{\lambda,\mu}=n_\lambda+n_\mu$, then up to parity $C(\lambda)\otimes C(\mu)\cong C(\lambda+\mu)\otimes\bigwedge K_{\lambda,\mu}$. Otherwise*

$$C(\lambda)\otimes C(\mu)=(C(\lambda+\mu)\otimes\textstyle\bigwedge K_{\lambda,\mu})\otimes\mathbb{C}^{2^a|2^a}.$$

where $a=(n_\lambda+n_\mu-n_{\lambda+\mu})/2-\dim K_{\lambda,\mu}$.

Corollary 3.1. *The module $C(\lambda)\otimes C(\mu)$ is Π-invariant except for the case when K_λ, K_μ, $K_{\lambda+\mu}$ have even codimensions in $\mathfrak{h}_{\bar{1}}$ and*

$$\mathfrak{h}_{\bar{1}}=K_\lambda+K_\mu+K_{\lambda+\mu}.$$

Proof. Note that $\Pi(C(\lambda))=C(\lambda)\otimes\Pi(\mathbb{C})\cong C(\lambda)$ implies $\Pi(C(\lambda)\otimes C(\mu))\cong C(\lambda)\otimes C(\mu)$. On the other hand, if $\operatorname{codim}\operatorname{Ker}F_{\lambda+\mu}$ is odd, then any $\mathfrak{g}_{\bar{1}}/K_{\lambda+\mu}$-module is Π-invariant. Therefore $C(\lambda)\otimes C(\mu)$ is Π-invariant if at least one of the numbers $\operatorname{codim}K_\lambda$, $\operatorname{codim}K_\mu$, $\operatorname{codim}K_{\lambda+\mu}$ is odd. Now assume that these numbers are even. Note that $K_\lambda\cap K_\mu=K_{\lambda+\mu}\cap K_\mu=K_\lambda\cap K_{\lambda+\mu}$. and set

$$m_{\lambda,\mu}:=\dim(K_\lambda\cap K_\mu),\quad r_\lambda:=\dim K_\lambda/(K_\lambda\cap K_\mu),\quad r_\mu:=\dim K_\mu/(K_\lambda\cap K_\mu).$$

Assume that $C(\lambda)\otimes C(\mu)$ is not Π-invariant. By Theorem 3.1 in this case

$$n_{\lambda+\mu}+\dim K_{\lambda,\mu}=n_\lambda+n_\mu.$$

One has

$$n_\lambda = \left\lfloor \frac{\operatorname{codim} K_\lambda + 1}{2} \right\rfloor = \frac{\operatorname{codim} K_\lambda}{2} = \frac{\dim \mathfrak{h}_{\bar{1}} - r_\lambda - m_{\lambda,\mu}}{2}$$

with similar formulae for n_μ and $n_{\lambda+\mu} = \dfrac{n - \dim K_{\lambda,\mu} - m_{\lambda,\mu}}{2}$. This gives

$$\dim K_{\lambda,\mu} + m_{\lambda,\mu} + r_\lambda + r_\mu = \dim \mathfrak{h}_{\bar{1}},$$

as required. □

4. The irreducible modules $L(\lambda)$ of $\mathfrak{g}$

We now return to the setting of Section 2.1, so that $\mathfrak{g}$ denotes a Lie superalgebra containing a finite-dimensional quasitoral subalgebra $\mathfrak{h}$. Choose a triangular decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}$ as in Section 2.1.2.

4.1. ***Highest weight modules.*** We call a $\mathfrak{g}$-module N a module of highest weight λ if $N_\lambda \neq 0$ and $N_\nu \neq 0$ implies $\nu \leq \lambda$.

View $C(\lambda)$ as a $\mathfrak{b}$-module with the zero action of $\mathfrak{n}$ and set

$$M(\lambda) := \operatorname{Ind}_{\mathfrak{b}}^{\mathfrak{g}} C(\lambda);$$

the module $M(\lambda)$ has a unique simple quotient, which we denote by $L(\lambda)$. Each simple module of highest weight λ is isomorphic to $L(\lambda)$ if rank F_λ is odd (i.e., if $\lambda \in I_1$); if rank F_λ is even (i.e., if $\lambda \in I_0$), each simple module of highest weight λ is isomorphic to either $L(\lambda)$ or to $\Pi L(\lambda)$, and these modules are not isomorphic.

Note that $\mathfrak{t}$ acts diagonally on $M(\lambda)$ and all weight spaces $M(\lambda)_\nu$ are finite-dimensional (since we assume all weight spaces $\mathcal{U}(\mathfrak{n}^-)_\nu$ are finite-dimensional); in particular, $\operatorname{Res}_{\mathfrak{h}}^{\mathfrak{g}} M(\lambda)$ lies in $\mathcal{M}(\mathfrak{h})$.

4.2. ***Duality.*** In many cases the antiautomorphism $(-)^\#$ introduced in Section 3.4.4 can be extended to an antiautomorphism of $\mathfrak{g}$ which satisfies $(a^\#)^\# = (-1)^{p(a)} a$. Using this antiautomorphism we can introduce a contragredient duality on $\mathfrak{g}$-modules N satisfying $\operatorname{Res}_{\mathfrak{h}}^{\mathfrak{g}} N \in \mathcal{M}(\mathfrak{h})$, in such a way that $\operatorname{Res}_{\mathfrak{h}}^{\mathfrak{g}} N^\# = (\operatorname{Res}_{\mathfrak{h}}^{\mathfrak{g}} N)^\#$. The map $N \xrightarrow{\sim} (N^\#)^\#$ is given by $v \mapsto (-1)^{p(v)} v$. By (6), $L(\lambda)^\# \cong L(\lambda) \cong \Pi L(\lambda)$ for $\lambda \in I_1$ and

$$L(\lambda)^\# \cong \Pi^{(\operatorname{rank} F_\lambda)/2} L(\lambda) \quad \text{for } \lambda \in I_0. \tag{7}$$

The antiautomorphism $(-)^\#$ exists for Kac–Moody superalgebras. For $\mathfrak{gl}(m|n)$ the antiautomorphism $(-)^\#$ can be given by the formula $a^\# := a^t$ for $a \in \mathfrak{g}_{\bar{0}}$ and $a^\# := \sqrt{-1} a^t$ for $a \in \mathfrak{g}_{\bar{1}}$ (where a^t stands for the transposed matrix); this antiautomorphism on $\mathfrak{gl}(n|n)$ induces $(-)^\#$ for the queer superalgebras $\mathfrak{q}_n, \mathfrak{sq}_n, \mathfrak{pq}_n, \mathfrak{psq}_n$.

4.2.1. *Remark.* The duality $(-)^\#$ can be defined using a "naive antiautomorphism", i.e., an invertible map $\sigma' : \mathfrak{g} \to \mathfrak{g}$ satisfying $\sigma'([a, b]) = [\sigma'(b), \sigma'(a)]$ via the formula $g.f(v) := f(\sigma(g)v)$ for $g \in \mathfrak{g}$, $f \in N^*$ and $v \in N$. This was done in [9] and [7].

Consider the map $\theta' : \mathfrak{g} \to \mathfrak{g}$ given by $\theta'(g) = (\sqrt{-1})^m g$, where $m = 0$ for $g \in \mathfrak{g}_{\bar{0}}$ and $m = 1$ for $g \in \mathfrak{g}_{\bar{1}}$. If σ' is a "naive antiautomorphism", then $\sigma'\theta'$ is an antiautomorphism.

5. Grothendieck rings and (super)character morphisms

Let $\mathcal{C}$ be a full subcategory of the category of $\mathfrak{g}$-modules. In this paper we always assume that all modules in $\mathcal{C}$ are of finite length and that $\mathcal{C}$ is a dense subcategory, i.e., for every short exact sequence $0 \to M' \to M \to M'' \to 0$ the module M lies in $\mathcal{C}$ if and only if M', M'' lie in $\mathcal{C}$. In addition, we usually assume that $\Pi\mathcal{C} = \mathcal{C}$.

If B is a super ring, we set $B_{\mathbb{Q}} := B \otimes_{\mathbb{Z}} \mathbb{Q}$.

5.1. ***Grothendieck groups.*** We denote by $\mathscr{K}(\mathcal{C})$ the Grothendieck group of $\mathcal{C}$, which is the free abelian group generated by N_{gr} for each module N in $\mathcal{C}$, modulo the relation that $[N] = [N'] + [N'']$ whenever $0 \to N' \to N \to N'' \to 0$ is a short exact sequence in $\mathcal{C}$. When $\Pi\mathcal{C} = \mathcal{C}$, we define the structure of a $\mathbb{Z}[\xi]/(\xi^2 - 1)$-module on $\mathscr{K}(\mathcal{C})$ by setting $\xi N_{\mathrm{gr}} := \Pi(N)_{\mathrm{gr}}$ for $N \in \mathcal{C}$. Set $\mathscr{K}_{\pm}(\mathcal{C}) := \mathscr{K}(\mathcal{C})/(\xi \mp 1)$.

We call the group $\mathscr{K}_{-}(\mathcal{C}) := \mathscr{K}(\mathcal{C})/(\xi + 1)$ the *reduced Grothendieck group*. We denote by $N_{\mathrm{gr},\pm}$ the image of N_{gr} in $\mathscr{K}_{\pm}(\mathcal{C})$; later we will use $[N]$ for $N_{\mathrm{gr},-}$.

Note that $\mathscr{K}(\mathcal{C})$ is an abelian group, so $\mathscr{K}_{\pm}(\mathcal{C})$ are also abelian groups.

5.1.1. We denote by $\mathrm{Irr}(\mathcal{C})$ the set of isomorphism classes of irreducible modules in $\mathcal{C}$ modulo Π and write

$$\mathrm{Irr}(\mathcal{C}) = \mathrm{Irr}(\mathcal{C})_{\bar{0}} \coprod \mathrm{Irr}(\mathcal{C})_{\bar{1}},$$

where $L \in \mathrm{Irr}(\mathcal{C})_{\bar{0}}$ if $\Pi(L) \ncong L$ and $L \in \mathrm{Irr}(\mathcal{C})_{\bar{1}}$ if $\Pi(L) \cong L$.

By our assumptions on $\mathcal{C}$, $\mathscr{K}(\mathcal{C})$ a free $\mathbb{Z}$-module with a basis

$$\{L_{\mathrm{gr}}, \xi L_{\mathrm{gr}} \mid L \in \mathrm{Irr}(\mathcal{C})_{\bar{0}}\} \coprod \{L_{\mathrm{gr}} \mid L \in \mathrm{Irr}(\mathcal{C})_{\bar{1}}\}.$$

5.1.2. The group $\mathscr{K}_{+}(\mathcal{C})$ is a free $\mathbb{Z}$-module with a basis $\{L_{\mathrm{gr},+} | L \in \mathrm{Irr}(\mathcal{C})\}$.

5.1.3. *The reduced Grothendieck group.* One has

$$\mathscr{K}_{-}(\mathcal{C}) = \mathscr{K}_{-}(\mathcal{C})_{\mathrm{free}} \oplus \mathscr{K}_{-}(\mathcal{C})_{2\text{-tor}}, \tag{8}$$

where $\mathscr{K}_{-}(\mathcal{C})_{\mathrm{free}}$ is a free $\mathbb{Z}$-module with a basis $\{L_{\mathrm{gr},-} \mid L \in \mathrm{Irr}(\mathcal{C})_{\bar{0}}\}$ and $\mathscr{K}_{-}(\mathcal{C})_{2\text{-tor}}$ is a free $\mathbb{Z}/2\mathbb{Z}$-module with a basis $\{L_{\mathrm{gr},-} \mid L \in \mathrm{Irr}(\mathcal{C})_{\bar{1}}\}$.

Proposition 5.1. (1) *The natural map $\psi : \mathscr{K}(\mathcal{C}) \to \mathscr{K}_{-}(\mathcal{C}) \times \mathscr{K}_{+}(\mathcal{C})$ is an embedding.*

(2) *The image of ψ is the subgroup consisting of the pairs*

$$\Big(\sum_{L \in \mathrm{Irr}(\mathcal{C})} m_L L_{\mathrm{gr},+}, \sum_{L \in \mathrm{Irr}(\mathcal{C})} n_L L_{\mathrm{gr},-}\Big),$$

where $m_L, n_L \in \mathbb{Z}$ with $m_L \equiv n_L \bmod 2$ for all L and $n_L \in \{0, 1\}$ for $L \in \mathrm{Irr}(\mathcal{C})_{\bar{1}}$.

(3) *The map ψ induces an isomorphism*

$$\mathscr{K}(\mathcal{C})_{\mathbb{Q}} \xrightarrow{\sim} \mathscr{K}_{-}(\mathcal{C})_{\mathbb{Q}} \times \mathscr{K}_{+}(\mathcal{C})_{\mathbb{Q}}.$$

Proof. For (i) take $a \in \mathscr{K}(\mathcal{C})(\xi - 1) \cap \mathscr{K}(\mathcal{C})(\xi + 1)$. Then $a = a_+(\xi - 1) = a_-(\xi + 1)$ for $a_\pm \in \mathscr{K}(\mathcal{C})$. Using $\xi^2 = 1$ we obtain

$$2a_+(1-\xi) = a_+(\xi - 1)^2 = a_-(\xi^2 - 1) = 0.$$

Since $\mathscr{K}_+(\mathcal{C})$ is a free $\mathbb{Z}$-module this gives $a_+(\xi - 1) = 0$, so $a = 0$. This establishes (i).

Assertion (ii) follows from the fact that the subgroup generated by the pairs $\psi(L_{\mathrm{gr}}) = (L_{\mathrm{gr},+}, L_{\mathrm{gr},-})$ and $\psi(\Pi L_{\mathrm{gr}}) = (L_{\mathrm{gr},+}, -L_{\mathrm{gr},-})$ for all $L \in \mathrm{Irr}(\mathcal{C})$ coincides with the subgroup described in (ii).

Finally, (iii) follows from (ii). $\square$

5.1.4. *Grothendieck rings.* If $\mathcal{C}$ is closed under $\otimes$, then $\mathscr{K}(\mathcal{C})$, $\mathscr{K}_\pm(\mathcal{C})$ are commutative rings with unity and ψ in Proposition 5.1 is a ring homomorphism. In this case $\mathscr{K}_-(\mathcal{C})_{2\text{-tor}}$ is an ideal in $\mathscr{K}_-(\mathcal{C})$.

If $\mathcal{C}$ is rigid, $\mathscr{K}(\mathcal{C})$ is equipped by an involution $*$ and $\mathscr{K}_\pm(\mathcal{C})$ and $\mathscr{K}_-(\mathcal{C})_{\mathrm{free}}$, $\mathscr{K}_-(\mathcal{C})_{2\text{-tor}}$ are $*$-invariant.

5.2. ***The map*** $\mathbf{ch}_{\mathfrak{h},\xi}$**.** Let $\mathfrak{g}'$ be a subalgebra of $\mathfrak{g}$ and let $\mathcal{C}'$ be a category of $\mathfrak{g}'$ modules such that restriction induces a functor $\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}'} : \mathcal{C} \to \mathcal{C}'$. For a suitable category $\mathcal{C}'$ for $\mathfrak{g}'$-modules, this functor induces a map $\mathrm{res}_{\mathfrak{g}'} : \mathscr{K}(\mathcal{C}) \to \mathscr{K}(\mathcal{C}')$ which is very useful if $\mathscr{K}(\mathcal{C}')$ is simple enough. Below we consider this map for the cases when $\mathfrak{g}' = \mathfrak{t}$ is a maximal toral subalgebra and for $\mathfrak{g}' = \mathfrak{g}^{\mathfrak{t}} = \mathfrak{h}$, a maximal quasitoral subalgebra.

As we will see below, $\mathrm{res}_{\mathfrak{h}}$ is an embedding if $\mathcal{C}$ is "nice enough"; in this case $\mathrm{res}_{\mathfrak{t}}$ induces an embedding $\mathscr{K}_+(\mathcal{C}) \to \mathscr{K}_+(\mathcal{M}(\mathfrak{t}))$ and this map is given by the usual (nongraded) characters.

5.2.1. Let $\mathfrak{h}$ be quasitoral, and let $\tilde{R}(\mathfrak{h})$ be the $\mathbb{Z}[\xi]$-module consisting of the sums

$$\sum_{\nu \in I_0} (m_\nu + k_\nu \xi)[C(\nu)] + \sum_{\nu \in I_1} m_\nu [C(\nu)], \quad m_\nu, k_\nu \in \mathbb{Z},$$

with the ξ-action given by

$$\xi\Big(\sum_{\nu \in I_0} (m_\nu + k_\nu \xi)[C(\nu)] + \sum_{\nu \in I_1} m_\nu [C(\nu)]\Big) = \sum_{\nu \in I_0} (m_\nu \xi + k_\nu)[C(\nu)] + \sum_{\nu \in I_1} m_\nu [C(\nu)]. \tag{9}$$

For $N \in \mathcal{M}(\mathfrak{h})$ we introduce

$$\mathrm{ch}_{\mathfrak{h},\xi}(N) := \sum_{\nu \in I_0} (m_\nu + k_\nu \xi)[C(\nu)] + \sum_{\nu \in I_1} m_\nu [C(\nu)] \in \tilde{R}(\mathfrak{h})$$

where $m_\nu := [N_\nu : C(\nu)]$ and $k_\nu := [N_\nu : \Pi(C(\nu))]$.

This defines a linear map $\mathrm{ch}_{\mathfrak{h},\xi} : \mathscr{K}(\mathcal{M}(\mathfrak{h})) \to \tilde{R}(\mathfrak{h})$, which we refer to as the graded $\mathfrak{h}$-character of N. We denote by $\mathrm{sch}\, N$ the image of $\mathrm{ch}_{\mathfrak{h},\xi}$ in $\tilde{R}(\mathfrak{h})/\tilde{R}(\mathfrak{h})(\xi + 1)$. Then

$$\mathrm{sch}_{\mathfrak{h}}(N) := \sum_{\nu \in I_0} (m_\nu - k_\nu)[C(\nu)] + \sum_{\nu \in I_1} m'_\nu [C(\nu)]$$

where $m'_\nu = 0$ if m_ν is even and $m'_\nu = 1$ if m_ν is odd.

Lemma 5.1. *The maps* $[N] \to \mathrm{ch}_{\mathfrak{h},\xi} N$ *and* $[N] \to \mathrm{sch}_{\mathfrak{h}} N$ *define isomorphisms* $\mathscr{K}(\mathcal{M}(\mathfrak{h})) \to \tilde{R}(\mathfrak{h})$ *and* $\mathscr{K}_-(\mathcal{M}(\mathfrak{h})) \to \tilde{R}(\mathfrak{h})/\tilde{R}(\mathfrak{h})(\xi + 1)$, *respectively. These maps are compatible with* $(-)^\#$.

5.2.2. *Note:* Because of this lemma, we will subsequently do away with the notation $\tilde{R}(\mathfrak{h})$ and instead directly identify $\mathscr{K}(\mathcal{M}(\mathfrak{h}))$ (and $\mathscr{K}_-(\mathcal{M}(\mathfrak{h}))$) with the corresponding spaces $\tilde{R}(\mathfrak{h})$ (and $\tilde{R}(\mathfrak{h})/\tilde{R}(\mathfrak{h})(1+\xi)$) as presented above.

Proof of Lemma 5.1. We show that $\mathrm{ch}_{\mathfrak{h},\xi}$ is an isomorphism, with the result for $\mathrm{sch}_{\mathfrak{h}}$ following easily. Clearly $\mathrm{ch}_{\mathfrak{h},\xi}$ is surjective, so it suffices to prove injectivity. First we observe that for N in $\mathcal{M}(\mathfrak{h})$, we have the following equality in the Grothendieck ring $\mathscr{K}(\mathcal{M}(\mathfrak{h}))$:

$$[N]=\Big[\bigoplus_{\nu\in I_0}\big(C(\nu)^{\oplus m_\nu}\oplus\Pi C(\nu)^{\oplus k_\nu}\big)\oplus\bigoplus_{\nu\in I_1}C(\nu)^{m_\nu}\Big].$$

This simply follows from the fact that N has finite Loewy length, since this is true for the algebras $\mathcal{C}\ell(\lambda)$. It is not difficult to see that a basis of $\mathscr{K}(\mathcal{M}(\mathfrak{h}))$ is given by elements of the form

$$\Big[\bigoplus_{\nu\in I_0}\big(C(\nu)^{\oplus m_\nu}\oplus\Pi C(\nu)^{\oplus k_\nu}\big)\oplus\bigoplus_{\nu\in I_1}C(\nu)^{m_\nu}\Big],$$

where $m_\nu, k_\nu\in\mathbb{N}$. From this the isomorphism easily follows. Compatibility with $(-)^{\#}$ is obvious. □

Corollary 5.1. (1) *One has* $\mathrm{ch}_{\mathfrak{h},\xi}\,\Pi N=\xi\,\mathrm{ch}_{\mathfrak{h},\xi}\,N$, $\;\mathrm{sch}_{\mathfrak{h}}\,\Pi N=-\,\mathrm{sch}_{\mathfrak{h}}\,N$.

(2) *For* $\lambda\in I_1$ *one has* $\mathrm{sch}_{\mathfrak{h}}\,L(\lambda)=\sum_{\mu\in I_1}m_\mu[C(\mu)]$.

(3) *If* $\mathfrak{g}$ *admits* $(-)^{\#}$ *as in Section 4.2, then the coefficient of* $[C(\nu)]$ *in* $\mathrm{sch}_{\mathfrak{h}}\,L(\lambda)$ *is zero if* $\operatorname{rank}F_\lambda$, $\operatorname{rank}F_\nu$ *are even and* $\operatorname{rank}F_\nu\not\equiv\operatorname{rank}F_\lambda \bmod 4$.

Proof. The assertions follow from (5) and (7). □

The next corollary is a direct generalization of [25], Proposition 4.2.

Corollary 5.2. *Let* $\mathcal{C}$ *be a full subcategory of* $\mathfrak{g}$*-modules with the following properties: each module in* $\mathcal{C}$ *is of finite length and is locally finite over* $\mathfrak{b}$, *and the restriction to* $\mathfrak{h}$ *lies in* $\mathcal{M}(\mathfrak{h})$. *Then the map* $\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{h}}$ *induces injective maps* $\mathrm{ch}_{\mathfrak{h},\xi}:\mathscr{K}(\mathcal{C})\hookrightarrow\mathscr{K}(\mathcal{M}(\mathfrak{h}))$ *and* $\mathrm{sch}_{\mathfrak{h}}:\mathscr{K}_-(\mathcal{C})\hookrightarrow\mathscr{K}_-(\mathcal{M}(\mathfrak{h}))$.

Proof. Let us check the injectivity of the first map $\mathrm{ch}_{\mathfrak{h},\xi}:\mathscr{K}(\mathcal{C})\to\mathscr{K}(\mathcal{M}(\mathfrak{h}))$. Any simple module in $\mathcal{C}$ is $L(\lambda)$ or $\Pi(L(\lambda))$ for some $\lambda\in\mathfrak{t}^*$. Since every module in $\mathcal{C}$ has finite length, $\mathscr{K}(\mathcal{C})$ is a free $\mathbb{Z}$-module spanned by $[L(\lambda)]$, $\xi[L(\lambda)]$ for $\lambda\in I_0$ and $[L(\lambda)]$ for $\lambda\in I_1$. Assume that

$$\mathrm{ch}_{\mathfrak{h},\xi}\Big(\sum_{i=1}^{s}(m_i+k_i\xi)[L(\lambda_i)]\Big)=0,$$

where $k_i=0$ for $\lambda_i\in I_1$ and $\gamma(\lambda_1)$ is maximal among $\gamma(\lambda_i)$ for $i=1,\dots,s$. Then for $i=2,\dots,s$ one has $L(\lambda_i)_{\lambda_1}=0$, so the coefficient of $[C(\lambda_1)]$ in $\mathrm{ch}_{\mathfrak{t}}([L(\lambda_i)]$ is zero. Hence $(m_1+k_1\xi)[L(\lambda_1)]=0$. This gives $m_1=k_1=0$ and implies the injectivity of $\mathrm{ch}_{\mathfrak{h},\xi}$. The injectivity of $\mathrm{sch}_{\mathfrak{h}}$ easily follows. □

5.3. *Example.* The category of finite-dimensional $\mathfrak{g}$-modules $\mathcal{F}in(\mathfrak{g})$ is a rigid tensor category with the duality $N \mapsto N^*$ given by the antiautomorphism $-\mathrm{Id}_{\mathfrak{g}}$. Note that the Grothendieck ring $\mathscr{K}(\mathcal{F}in(\mathfrak{g}))$ is a commutative ring with a basis $\{[L]\}$, where L runs through isomorphism classes of finite-dimensional simple modules. Corollary 5.2 gives an embedding $\mathrm{ch}_{\mathfrak{h},\xi} : \mathscr{K}(\mathcal{F}in(\mathfrak{g})) \hookrightarrow \mathscr{K}(\mathcal{F}in(\mathfrak{h})))$ ($[M] \mapsto \mathrm{ch}_{\mathfrak{h},\xi} M$) which is a ring homomorphism. By abuse of notation we will also denote the image of this homomorphism by $\mathscr{K}(\mathcal{F}in(\mathfrak{g}))$.

The duality induces an involution on $\mathscr{K}(\mathcal{F}in(\mathfrak{g}))$, which we also denote by $*$. One has $\xi^* = \xi$. The homomorphism $\mathrm{ch}_{\mathfrak{h},\xi} : \mathscr{K}(\mathcal{F}in(\mathfrak{g})) \hookrightarrow \mathscr{K}(\mathcal{F}in(\mathfrak{h}))$ is compatible with $*$, so $\mathscr{K}(\mathcal{F}in(\mathfrak{g}))$ is a $*$-stable subring of $\mathscr{K}(\mathcal{F}in(\mathfrak{h}))$.

5.3.1. *Remark.* Let $\mathfrak{g}$ be a Kac–Moody superalgebra (so $\mathfrak{t} = \mathfrak{h}$) and let $\Lambda_{\mathrm{int}} \subset \mathfrak{h}^*$ be a lattice containing $\Delta(\mathfrak{g})$ such that the parity $p : \Delta \to \mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ can be extended to $p : \Lambda_{\mathrm{int}} \to \mathbb{Z}_2$. Assume, in addition, that for the category $\mathcal{C}$, each $N \in \mathcal{C}$ has $N_\nu = 0$ if $\nu \notin \Lambda_{\mathrm{int}}$.

Then $\mathcal{C} = \mathcal{C}_+ \oplus \Pi(\mathcal{C}_+)$, where $N \in \mathcal{C}$ lies in $\mathcal{C}_+$ if and only if $N_\nu \subset N_{p(\nu)}$. We have that

$$\mathscr{K}(\mathcal{C}) = \mathscr{K}(\mathcal{C}_+) \times \mathbb{Z}[\xi]/(\xi^2 - 1).$$

and thus one can recover $\mathscr{K}(\mathcal{C})$ from $\mathscr{K}(\mathcal{C}_+)$; however we note that $\mathcal{C}_+$ is not Π-invariant.

Further, we have in this case that $\mathscr{K}_-(\mathcal{C}) \cong \mathscr{K}_+(\mathcal{C}) \cong \mathscr{K}(\mathcal{C}_+)$. If $\mathcal{C}$ is a tensor category, then $\mathcal{C}_+$ is also a tensor category (but $\Pi(\mathcal{C}_+)$ is not). In [25], Sergeev and Veselov described the ring $\mathscr{K}(\mathcal{C}_+)$ for the finite-dimensional Kac–Moody superalgebras.

5.4. *The ring $\mathscr{K}(\mathfrak{h})$.* Let $\mathfrak{h}$ be quasitoral. We write $\mathscr{K}(\mathfrak{h}) := \mathscr{K}(\mathcal{F}in(\mathfrak{h}))$. The map $[N] \mapsto \mathrm{ch}_{\mathfrak{h},\xi} N$ introduced in Section 5.2 gives an isomorphism of $\mathscr{K}(\mathfrak{h})$ and the free $\mathbb{Z}$-module spanned by $[C(\nu)]$, $\xi[C(\nu)]$ for $i \in I_0$ and $[C(\nu)]$ for $i \in I_1$. We view $\mathscr{K}(\mathfrak{h})$ as a commutative algebra endowed with the involutions $(-)*$ and $(-)^{\#}$. One has $C(\lambda) \in \mathrm{Irr}(\mathcal{C})_{\bar{i}}$ if and only if $\lambda \in I_i$.

One has $[E_0] = 1$, $[\Pi(E_0)] = \xi$ and

$$\begin{array}{ll} [C(\lambda)]^* \in \{[C(-\lambda)], \xi[C(-\lambda)]\}, \quad [C(\lambda)]^{\#} = \xi^{(\mathrm{rank}\, F_\lambda)/2}[C(\lambda)], & \text{for } \lambda \in I_0 \\ \xi[C(\lambda)] = [C(\lambda)], \quad [C(\lambda)]^* = [C(-\lambda)], \quad [C(\lambda)]^{\#} = [C(\lambda)], & \text{for } \lambda \in I_1. \end{array} \tag{10}$$

5.4.1. The multiplication in $\mathscr{K}_+(\mathfrak{h})$ is given by

$$C(\lambda)_{\mathrm{gr},+} C(\nu)_{\mathrm{gr},+} = \frac{\dim C(\lambda) \dim C(\nu)}{\dim C(\lambda + \nu)} C(\lambda + \nu)_{\mathrm{gr},+}.$$

Let $\mathbb{Z}[e^\nu, \nu \in \mathfrak{t}^*]$ be the group ring of $\mathfrak{t}^*$. For $N \in \mathcal{M}(\mathfrak{h})$ we set

$$\mathrm{ch}_{\mathfrak{t}} N := \sum_{\nu \in \mathfrak{t}^*} \dim N_\nu e^\nu.$$

The map $N \mapsto \mathrm{ch}_{\mathfrak{t}} N$ induces an embedding $\mathscr{K}_+(\mathfrak{h}) \hookrightarrow \mathbb{Z}[e^\nu, \nu \in \mathfrak{t}^*]$. The image is the subring of elements $\sum_\nu m_\nu e^\nu$ where m_ν is divisible by $\dim C(\nu)$, and we have an isomorphism $\mathscr{K}_+(\mathfrak{h})_{\mathbb{Q}} \xrightarrow{\sim} \mathbb{Q}[e^\nu, \nu \in \mathfrak{t}^*]$.

5.5. ***Spoiled superalgebras.*** We call a $\mathbb{Z}$-superalgebra $A = A_{\bar 0} \oplus A_{\bar 1}$ *spoiled* if $2A_{\bar 1} = A_{\bar 1}^2 = 0$. Given a superalgebra B, we make it into a spoiled superalgebra B^{spoil} by setting, as a ring,

$$B^{\mathrm{spoil}} = \big(B \otimes_{\mathbb{Z}} \mathbb{Z}[\varepsilon]/(\varepsilon^2, 2\varepsilon)\big)_{\bar 0}.$$

Here ε is an element of degree 1. The $\mathbb{Z}_2$-grading on B^{spoil} is declared to be $B_{\bar 0}^{\mathrm{spoil}} = B_{\bar 0}$ and $B_{\bar 1}^{\mathrm{spoil}} = B_{\bar 1}\varepsilon$. We observe that

$$B_{\mathbb{Q}}^{\mathrm{spoil}} := (B_{\bar 0})_{\mathbb{Q}}.$$

5.5.1. *The algebra $\mathscr{K}_-(\mathfrak{h})$.* From now on we will write $[M]$ in place of $M_{\mathrm{gr},-}$. Using Corollary 3.1 we obtain that in $\mathscr{K}_-(\mathfrak{h})$ we have the following multiplication law:

$$[C(\nu)][C(\lambda)] = \begin{cases} \pm[C(\lambda+\nu)] & \text{if } \operatorname{rank} F_\lambda + \operatorname{rank} F_\nu = \operatorname{rank} F_{\lambda+\nu} \text{and} \operatorname{rank} F_\lambda \cdot \operatorname{rank} F_\nu \equiv 0 \bmod 2, \\ 0 & \text{otherwise.} \end{cases}$$

As a result, $\mathscr{K}_-(\mathfrak{h})$ is $\mathbb{Z}$-graded algebra

$$\mathscr{K}_-(\mathfrak{h}) = \bigoplus_{i=0}^{\infty} \mathscr{K}_-(\mathfrak{h})_i \tag{11}$$

where $\mathscr{K}_-(\mathfrak{h})_i$ is spanned by $[C(\nu)]$ with $\operatorname{rank} F_\nu = i$. We consider the corresponding $\mathbb{Z}_2$-grading

$$\mathscr{K}_-(\mathfrak{h})_{\bar 0} := \bigoplus_{i=0}^{\infty} \mathscr{K}_-(\mathfrak{h})_{2i}, \quad \mathscr{K}_-(\mathfrak{h})_{\bar 1} := \bigoplus_{i=0}^{\infty} \mathscr{K}_-(\mathfrak{h})_{2i+1}.$$

One has $\mathscr{K}_-(\mathfrak{h})_{\bar 1} \cdot \mathscr{K}_-(\mathfrak{h})_{\bar 1} = 0$ and $2\mathscr{K}_-(\mathfrak{h})_{\bar 1} = 0$, so that $\mathscr{K}_-(\mathfrak{h})$ is a spoiled algebra. The following corollary is clear.

Corollary 5.3. (1) *The algebra $\mathscr{K}_-(\mathfrak{h})$ is a spoiled superalgebra with*

$$\mathscr{K}_-(\mathfrak{h})_{\bar 0} = \bigoplus_{i=0}^{\infty} \mathscr{K}_-(\mathfrak{h})_{2i} = \mathscr{K}_-(\mathfrak{h})_{\mathrm{free}}, \quad \mathscr{K}_-(\mathfrak{h})_{\bar 1} = \bigoplus_{i=0}^{\infty} \mathscr{K}_-(\mathfrak{h})_{2i+1} = \mathscr{K}_-(\mathfrak{h})_{2\text{-tor}}.$$

(2) *The algebra $\mathscr{K}(\mathfrak{h})$ is isomorphic to a subalgebra of $\mathbb{Z}[e^\nu, \nu \in \mathfrak{t}^*] \times \mathscr{K}_-(\mathfrak{h})$ consisting of*

$$\Big(\sum_{\nu\in\mathfrak{t}^*} m_\nu^+ \dim C(\nu) e^\nu,\ \sum_{\nu\in I_0} m_\nu^- [C(\nu)] + \varepsilon \sum_{\nu\in I_1} m_\nu^- [C(\nu)]\Big)$$

where $m_\nu^\pm \in \mathbb{Z}$ are equal to zero except for finitely many values of ν, $m_\nu^+ \equiv m_\nu^-$ modulo 2 and $m_\nu^- \in \{0, 1\}$ for $\nu \in I_1$.

(3) *The algebra $\mathscr{K}(\mathfrak{h})_{\mathbb{Q}}$ is isomorphic to $\mathbb{Q}[e^\nu, \nu \in \mathfrak{t}^*] \times \mathscr{K}_-(\mathfrak{h})_{\mathbb{Q}}$.*

5.5.2. For the rest of this section we set $\mathscr{K}(\mathfrak{g}) := \mathscr{K}(\mathcal{F}in(\mathfrak{g}))$, where $\mathfrak{g}$ is as in Section 4. Let $\psi_\pm(\mathfrak{g}) : \mathscr{K}(\mathfrak{g}) \to \mathscr{K}_\pm(\mathfrak{g})$ be the canonical epimorphisms. By Proposition 5.1, $\psi_+ \times \psi_-$ gives an embedding $\mathscr{K}(\mathfrak{g}) \hookrightarrow \mathscr{K}_+(\mathfrak{g}) \times \mathscr{K}_-(\mathfrak{g})$.

We will use the following construction: for any subsets $A_\pm \subset \mathscr{K}_\pm(\mathfrak{h})$ we introduce

$$A_+ \underset{\mathscr{K}(\mathfrak{h})}{\times} A_- := \{a \in \mathscr{K}(\mathfrak{h}) \mid \psi_\pm(a) \in A_\pm\}.$$

Note that $A_+ \underset{\mathscr{K}(\mathfrak{h})}{\times} A_-$ is a subring of $\mathscr{K}(\mathfrak{h})$ if $A_\pm$ are rings.

Lemma 5.2. *Let A be a $\mathbb{Z}[\xi]$-submodule of $\mathscr{K}(\mathfrak{h})$ with the following property: if $a \in \mathscr{K}(\mathfrak{h})$ and $2a \in A$, then $a \in A$. Then*
$$A = \psi_+(A) \underset{\mathscr{K}(\mathfrak{h})}{\times} \psi_-(A).$$
Proof. Take $a \in \mathscr{K}(\mathfrak{h})$ such that $\psi_\pm(a) \in \psi_\pm(A)$. Then A contains $a - c(1-\xi)$ for some $c \in \mathscr{K}(\mathfrak{h})$. Since A is a $\mathbb{Z}[\xi]$-submodule, A contains $(1+\xi)(a - c(1-\xi)) = (1+\xi)a$. Similarly, A contains $(1-\xi)a$, so $2a \in A$. Then the assumption gives $a \in A$ as required. $\square$

Corollary 5.4. $\mathscr{K}(\mathfrak{g}) = \mathscr{K}_+(\mathfrak{g}) \underset{\mathscr{K}(\mathfrak{h})}{\times} \mathscr{K}_-(\mathfrak{g})$.

Proof. Recall by Corollary 5.2 that $\mathscr{K}(\mathfrak{g})$ is a $\mathbb{Z}[\xi]$-subring of $\mathscr{K}(\mathfrak{h})$. Take $a \in \mathscr{K}(\mathfrak{h})$ with $2a \in \mathscr{K}(\mathfrak{g})$. Write
$$2a = \sum_{\lambda \in P^+(\mathfrak{g})} m_\lambda \operatorname{ch}_{\mathfrak{h},\xi} L(\lambda) = \sum_\nu 2m'_\nu [C(\nu)]$$
where $m_\lambda, m'_\nu \in \mathbb{Z}[\xi]$, and $P^+(\mathfrak{g})$ is the set of dominant weights of $\mathfrak{g}$. Let λ be maximal such that $m_\lambda \neq 0$. Then $m_\lambda = 2m'_\lambda$, so we may subtract $2L(\lambda)_{\mathrm{gr}}$, and conclude by induction. $\square$

5.6. *Equivariant setting.* Suppose that G is a finite group which acts on a quasitoral superalgebra $\mathfrak{h}$ by automorphisms. Our main example of this setup is when we consider quasireductive Lie superalgebras in Section 7.1, and G is the Weyl group.

The group G then acts naturally on $\mathscr{K}(\mathfrak{h})$ by twisting, i.e., $g \cdot V_{\mathrm{gr}} = V_{\mathrm{gr}}^g$ for an $\mathfrak{h}$-module V. This descends to a natural action of G on $\mathscr{K}_-(\mathfrak{h})$. Suppose that $\nu \in \mathfrak{t}^*$ and $g \in G$ such that $g\nu = \nu$. Then we have either $g \cdot C(\nu)_{\mathrm{gr}} = C(\nu)_{\mathrm{gr}}$ or $g \cdot C(\nu)_{\mathrm{gr}} = \xi C(\nu)_{\mathrm{gr}}$. It follows that on the reduced Grothendieck ring we have $g \cdot [C(\nu)] = \pm[C(\nu)]$. Thus $\operatorname{Stab}_G[C(\nu)] \subseteq \operatorname{Stab}_G \nu$. Define
$$\mathfrak{t}_G^* = \{\nu \in \mathfrak{t}^* : \operatorname{Stab}_G[C(\nu)] = \operatorname{Stab}_G \nu\}.$$
We observe that $\mathfrak{t}_G^*$ is a nonempty, G-stable cone in $\mathfrak{t}^*$. It need not be open or closed, and it may consist only of 0. It is clear that $I_1 \subseteq \mathfrak{t}_G^*$.

Let $\nu \in \mathfrak{t}^*$. We introduce the grading on $C(\nu)$ in the way described in Section 3.4.3. Since $\overline{T}_\nu$ is proportional to a product of a basis elements in $\mathfrak{h}_{\bar{1}}/K_\lambda$, one has
$$g(\overline{T}_\nu) = \det(g|_{\mathfrak{h}_{\bar{1}}/K_\nu})\overline{T}_\nu \quad \text{for each } g \in \operatorname{Stab}_W \nu. \tag{12}$$
Since g is acting by an orthogonal transformation on $\mathfrak{h}_{\bar{1}}/K_\nu$ we have $\det(g|_{\mathfrak{h}_{\bar{1}}/K_\nu}) = \pm 1$. Therefore for $\nu \in I_0$ and $g \in \operatorname{Stab}_G \nu$ we have
$$g[C(\nu)] = \det(g|_{\mathfrak{h}_{\bar{1}}/K_\nu})[C(\nu)].$$
Corollary 5.5. (1) *We have* $\mathfrak{t}_G^* \cap I_0 = \{\nu \in I_0 : \det(g|_{\mathfrak{h}_{\bar{1}}/K_\nu}) = 1$ *for all* $g \in \operatorname{Stab}_G \nu\}$.

(2) *For $\nu \in \mathfrak{t}_G^*$, the element*
$$a_\nu := \sum_{g \in G/\operatorname{Stab}_G \nu} g[C(\nu)]$$

is nonzero and well-defined.

(3) *The algebra $\mathscr{K}_-(\mathfrak{h})^G$ is naturally a spoiled superalgebra; the even part has a $\mathbb{Z}$-basis given by a_ν for a choice of coset representatives $\nu \in \mathfrak{t}^*_G \cap I_0/G$. The odd part has a $\mathbb{Z}_2$-basis given by a_ν for a choice of coset representatives $\nu \in I_1/G$.*

6. $\mathfrak{h}$-supercharacters of some highest weight $\mathfrak{g}$-modules

We continue to write $[N]$ for the image a module N in the corresponding reduced Grothendieck group.

Fix a triangular decomposition $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+$ coming from $\gamma : \mathfrak{t}^* \to \mathbb{R}$ as in Section 4, and consider the corresponding category $\mathcal{O}$ with respect to $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+$. To be precise, $\mathcal{O}$ consists of all finitely generated $\mathfrak{g}$-modules which are weight modules and $\mathfrak{n}$-locally finite. We take $M(\lambda)$ as in Section 4.1. For weights $\lambda, \mu \in \mathfrak{t}^*$ we write $\lambda \sim_0 \mu$ if $L(\lambda)$ is a composition factor of $M(\mu)$; then we let $\sim$ be the equivalence relation on $\mathfrak{t}^*$ generated by $\sim_0$.

The goal of this section is to prove Theorem 6.1, which in some sense is the best version of Theorem 1.1 that holds in great generality. The idea is to enforce assumptions that guarantee that all Verma modules with highest weights lying in a fixed equivalence class of interest have an especially nice $\mathfrak{h}$-supercharacter.

6.1. *Notation.* Let $\mathfrak{a}' \subseteq \mathfrak{a}$ be Lie superalgebras, L' a simple $\mathfrak{a}'$-module, and N an $\mathfrak{a}$-module such that $\operatorname{Res}^{\mathfrak{a}}_{\mathfrak{a}'} N$ has a finite length. Set

$$\operatorname{smult}(N : L') := \begin{cases} [\operatorname{Res}^{\mathfrak{a}}_{\mathfrak{a}'} N : L'] - [\operatorname{Res}^{\mathfrak{a}}_{\mathfrak{a}'} N : \Pi L'] & \text{if } L' \ncong \Pi L' \\ [\operatorname{Res}^{\mathfrak{a}}_{\mathfrak{a}'} N : L'] \bmod 2 & \text{if } L' \cong \Pi L'. \end{cases} \tag{13}$$

If $\operatorname{Res}^{\mathfrak{g}}_{\mathfrak{h}} N \in \mathcal{M}(\mathfrak{h})$ we set

$$\Omega(N) := \{\mu \in \mathfrak{t}^* | N_\mu \neq 0\}, \qquad \mathrm{s}\,\Omega(N) := \{\mu \in \mathfrak{t}^* | \operatorname{smult}(N_\mu : C(\mu)) \neq 0\}. \tag{14}$$

Then

$$\operatorname{sch}_{\mathfrak{h}} N = \sum_{\nu \in \mathrm{s}\,\Omega(N)} \operatorname{smult}(N_\nu : C(\nu))[C(\nu)].$$

6.2. *On* $\operatorname{sch}_{\mathfrak{h}} M(\lambda)$. For $\nu \in Q^-$ we have an isomorphism of $\mathfrak{h}$-modules:

$$M(\lambda)_{\lambda+\nu} \cong \mathcal{U}(\mathfrak{n}^-)_\nu \otimes C(\lambda).$$

By Theorem 3.1 , $\operatorname{sch}_{\mathfrak{h}}(C(\nu) \otimes C(\lambda)) \neq 0$ implies $\dim C(\nu) \cdot \dim C(\lambda) = \dim C(\lambda + \nu)$. Therefore

$$\mathrm{s}\,\Omega(M(\lambda)) \subset \big\{\nu \in \mathfrak{t}^* \mid \dim C(\nu) \cdot \dim C(\lambda) = \dim C(\lambda + \nu)\big\}. \tag{15}$$

Corollary 6.1. *Assume that $F_\nu \neq 0$ for $\nu \in Q^- \setminus \{0\}$. If* $\operatorname{corank} F_\lambda \leq 1$ *and* $\dim \mathfrak{h}_{\bar{1}}$ *is even, or* $\operatorname{corank} F_\lambda = 0$, *then*

$$\operatorname{sch}_{\mathfrak{h}} M(\lambda) = \operatorname{sch}_{\mathfrak{h}} C(\lambda).$$

Proof. If $\operatorname{corank} F_\lambda = 0$, or $\operatorname{corank} F_\lambda = 1$ and $\dim \mathfrak{h}_{\bar{1}}$ is even, then $\dim C(\lambda) \geq \dim C(\lambda + \nu)$ for any ν, so the formula follows from (15). ☐

6.3. $\mathrm{sch}_{\mathfrak{h}}\, L(\lambda)$ *when* corank $F_\lambda \le 1$. We make the following assumptions on our triangular decomposition:

(A1) Setting $Q^- = \mathbb{N}\Delta^-$, we have that $\gamma(Q^-) \subseteq \mathbb{R}^-$ is discrete.

(A2) One has $F_\nu \neq 0$ for each $\nu \in Q^- \setminus \{0\}$.

(A3) For $\lambda \in \mathfrak{t}^*$ such that corank $F_\lambda \le 1$, we have that $M(\lambda)$ has finite length as a $\mathfrak{g}$-module. (One can weaken this to assume that it admits a local composition series in the sense of [4].)

Theorem 6.1. *Let $\Lambda \subset \mathfrak{t}^*$ be an equivalence class of $\sim$ such that*

(A4) corank $F_\lambda \le 1$ *for each $\lambda \in \Lambda$ if* $\dim \mathfrak{h}_{\bar{1}}$ *is even, and otherwise* corank $F_\lambda = 0$ *for each $\lambda \in \Lambda$.*

Then for each $\lambda \in \Lambda$ one has

$$\mathrm{sch}_{\mathfrak{h}}\, L(\lambda) = \sum_{\mu\in\Lambda} m_\mu\, \mathrm{sch}_{\mathfrak{h}}\, C(\mu) \tag{16}$$

for some $m_\mu \in \mathbb{Z}$.

Proof. Assume that the assertion does not hold. Take (μ, μ') such that $\gamma(\mu' - \mu)$ is maximal with the properties

$$\mu \in \Lambda, \quad \mu' \notin \Lambda, \quad \mathrm{sch}_{\mathfrak{h}}\, L(\mu)_{\mu'} \neq 0$$

(the existence of such a pair follows from (A1)).

By (A3), in the Grothendieck ring of $\mathfrak{g}$ we may write

$$[L(\mu)] = [M(\mu)] - \sum_{\substack{\nu\in Q^-\setminus\{0\}\\ \mu+\nu\in\Lambda}} a_{\mu\nu}[L(\mu+\nu)]$$

where $a_{\mu\nu} \in \mathbb{N}$. We may map the above formula to the Grothendieck ring of $\mathfrak{h}$-modules via restriction, and thus learn that

$$\mathrm{sch}_{\mathfrak{h}}\, L(\mu)_{\mu'} = \mathrm{sch}_{\mathfrak{h}}\, M(\mu)_{\mu'} - \sum_{\substack{\nu\in Q^-\setminus\{0\}\\ \mu+\nu\in\Lambda}} \mathrm{sch}_{\mathfrak{h}}\, L(\mu+\nu)_{\mu'}.$$

By (15), $\mathrm{sch}_{\mathfrak{h}}\, M(\mu)_{\mu'} = 0$ and the maximality of $\gamma(\mu' - \mu)$ implies $\mathrm{sch}_{\mathfrak{h}}\, L(\mu+\nu)_{\mu'} = 0$. Hence $\mathrm{sch}_{\mathfrak{h}}\, L(\mu)_{\mu'} = 0$, a contradiction. □

We will see in Section 9 that Theorem 6.1 holds for any irreducible highest weight representation of $\mathfrak{q}_n$, without restrictions on corank F_λ.

7. On $\mathscr{K}_-(\mathfrak{g})$ in the case when $\mathfrak{g}$ is quasireductive

In this section we assume $\mathfrak{g}$ is quasireductive, i.e., $\mathfrak{g}$ is finite-dimensional, $\mathfrak{g}_{\bar{0}}$ is reductive, and $\mathfrak{g}_{\bar{1}}$ is a semisimple $\mathfrak{g}_{\bar{0}}$-module (see [22], [18], and [15] for examples and a partial classification of such algebras). The maximal toral subalgebras $\mathfrak{t}$ in $\mathfrak{g}_{\bar{0}}$ are the Cartan subalgebras in $\mathfrak{g}_{\bar{0}}$, and the maximal quasitoral subalgebras $\mathfrak{h}$ in $\mathfrak{g}$ are the Cartan subalgebras in $\mathfrak{g}$; all such subalgebras are conjugate to one another by inner automorphisms because the same is true on the even part (see [14]). We fix such $\mathfrak{t}$ and $\mathfrak{h}$. Let $P(\mathfrak{g}) \subset \mathfrak{t}^*$ be the set of weights appearing in finite-dimensional $\mathfrak{g}$-modules, which is the same as the set of weights appearing in finite-dimensional $\mathfrak{g}_{\bar{0}}$-modules. We fix a positive system of roots on $\mathfrak{g}$, and thus also on $\mathfrak{g}_{\bar{0}}$. We denote by $P^+(\mathfrak{g})$ the set of dominant weights, i.e.,

$$P^+(\mathfrak{g}) := \{\lambda \in \mathfrak{t}^* \mid \dim L(\lambda) < \infty\}.$$

We have

$$P(\mathfrak{g}) = P^+(\mathfrak{g}) + \mathbb{Z}\Delta = P^+(\mathfrak{g}_{\bar{0}}) + \mathbb{Z}\Delta_{\bar{0}}.$$

Write $\mathcal{C}$ for the subcategory of $\mathcal{F}(\mathfrak{h})$ consisting of modules with weights lying in $P(\mathfrak{g})$.

7.1. *On* $\mathbf{sch}_{\mathfrak{h}}(\mathcal{F}(\mathfrak{g}))$. Define $\mathcal{F}(\mathfrak{g})$ to be the full subcategory of $\mathcal{F}in(\mathfrak{g})$ consisting of modules which are semisimple over $\mathfrak{g}_{\bar{0}}$.

The Weyl group W of $\mathfrak{g}_{\bar{0}}$ acts naturally on $\mathfrak{h}$, and thus we are in the setup of Section 5.6; we refer to that section for the definition of $\mathfrak{t}^*_W$. We see that W preserves the subcategory $\mathcal{C}$. We set

$$P(\mathfrak{g})' = \mathfrak{t}^*_W \cap P(\mathfrak{g}), \quad P^+(\mathfrak{g}_{\bar{0}})' = \mathfrak{t}^*_W \cap P^+(\mathfrak{g}_{\bar{0}}).$$

Recall that

$$P(\mathfrak{g})' \cap I_0 := \big\{\lambda \in P(\mathfrak{g}) \cap I_0 \mid \forall w \in \operatorname{Stab}_W \lambda,\ \det(w|_{\mathfrak{h}_{\bar{1}}/\operatorname{Ker} F_\lambda}) = 1\big\}.$$

From the theory of reductive Lie algebras, in this case we have a natural bijection $P^+(\mathfrak{g}_{\bar{0}})' \to P(\mathfrak{g})'/W$. Recall that for $\nu \in \mathfrak{t}^*_W$ we set

$$a_\nu = \sum_{w \in W/\operatorname{Stab}_W \nu} w[C(\nu)].$$

Theorem 7.1. (i) *The algebra $\mathscr{K}_-(\mathcal{C})^W$ is naturally a spoiled superalgebra; its even part has $\mathbb{Z}$-basis given by a_ν for $\nu \in P^+(\mathfrak{g}_{\bar{0}})' \cap I_0$ and its odd part has a $\mathbb{Z}_2$-basis a_ν for $\nu \in \big(P^+(\mathfrak{g}_{\bar{0}}) \cap I_1\big)$.*

(ii) $\mathrm{sch}_{\mathfrak{h}}$ *defines an embedding $\mathscr{K}_-(\mathfrak{g}) \hookrightarrow \mathscr{K}_-(\mathcal{C})^W$.*

(iii) *For $\lambda \in P^+(\mathfrak{g}) \cap I_0$ one has*

$$\mathrm{sch}_{\mathfrak{h}} L(\lambda) = \sum_\nu k_\nu a_\nu$$

where $k_\nu \in \mathbb{Z}$ with $k_\lambda = 1$.

(iv) *If $\lambda \in P^+(\mathfrak{g}) \cap I_1$, then*

$$\mathrm{sch}_{\mathfrak{h}} L(\lambda) = \sum_{\nu \in I_1} k_\nu a_\nu.$$

with $k_\lambda = 1$.

(v) $P^+(\mathfrak{g}) \subset P^+(\mathfrak{g}_{\bar{0}})'$.

Proof. Part (i) follows from Corollary 5.5. For part (ii) let $L := L(\lambda)$ be a simple finite-dimensional $\mathfrak{g}$-module. Clearly all its weights lie in $P(\mathfrak{g})$. Because W can be realized from inner automorphisms of $\mathfrak{g}$, it is clear that the $\mathrm{sch}_{\mathfrak{h}} L(\lambda)$ must be W-invariant, proving (ii).

Part (iii) uses (7), and part (iv) uses that $L(\lambda) \cong \Pi L(\lambda)$. Finally part (v) then follows from (iii). $\square$

7.1.1. *Remark.* The algebra $\mathscr{K}_-(\mathfrak{g})$ is not naturally spoiled with respect to the embedding into $\mathscr{K}_-(\mathfrak{h})$. For example if $\mathfrak{g}=\mathfrak{q}_2\times\mathfrak{q}_1$, then consider the module $S^2V_{std}^{(1)}\boxtimes V_{std}^{(2)}$, where $V_{std}^{(1)}$, $V_{std}^{(2)}$ denote the standard modules of $\mathfrak{q}_2$, $\mathfrak{q}_1$ respectively. Then this module is of the form $L\oplus\Pi L$ for some simple $\mathfrak{q}_2\times\mathfrak{q}_1$-module L. One may check that $\mathrm{sch}_{\mathfrak{h}} L$ is not a homogeneous element of $\mathscr{K}_-(\mathfrak{h})$, with respect to its spoiled grading (see Section 5.5.1).

Corollary 7.1. *Let $\lambda\in P^+(\mathfrak{g})\cap I_0$ and $w_0\in W$ be such that $w_0\lambda=-\lambda$ and $L(\lambda)^*$ has highest weight λ. Then we have*

$$L(\lambda)^*\cong\Pi^iL(\lambda),\qquad \textit{where}\ (-1)^i=\det(w_0|_{\mathfrak{h}_1/\operatorname{Ker}F_\lambda}).$$

Proof. Recall that $\phi_\lambda:\mathcal{U}(\mathfrak{h})\to\mathcal{C}\ell(\lambda)$, $p_\lambda:\mathcal{C}\ell(\lambda)\to\mathcal{C}\ell(\lambda)/K_\lambda$ stand for the canonical epimorphisms. By Section 3.4.4 we have $\overline{T}_\lambda=H'_{2k}\dots H'_1$, where $H'_1,\dots,H'_{2k}$ is a basis of $p_\lambda\phi_\lambda(\mathfrak{h}_1)\subset\mathcal{C}\ell(\lambda)/K_\lambda$ satisfying $[H'_i,H'_j]=0$ for $i\neq j$. Set

$$\overline{T}_{-\lambda}=\overline{T}_{w_0\lambda}=w_0(H'_{2k})\dots w_0(H'_1).$$

From this definition we have $C(-\lambda)=C(\lambda)^{w_0}$. We have $\overline{\sigma_{-\lambda}}(H'_i)=-H'_i$, and thus

$$\overline{\sigma_{-\lambda}}(\overline{T}_{-\lambda})=(-1)^{k(2k-1)}w_0(H'_1)\dots w_0(H'_k)=(-1)^{k(2k-1)}\det(w_0|_{\mathfrak{h}_{\bar{1}}/K_\lambda})H'_1\dots H'_{2k}=(-1)^i\overline{T}_\lambda.$$

It follows that $C(\lambda)^*=\Pi^iC(-\lambda)$. Therefore

$$(L(\lambda)^*)_\lambda=(L(\lambda)_{-\lambda})^*=(\Pi^iC(\lambda)^*)^*=\Pi^iC(\lambda).$$

and we obtain $L(\lambda)^*\cong\Pi^iL(\lambda)$ as required. □

8. The DS-functor and the reduced Grothendieck group

In Sections 8.1 and 8.2, $\mathfrak{g}$ is any Lie superalgebra. We fix $x\in\mathfrak{g}_{\bar{1}}$ with $[x,x]=c\in\mathfrak{g}_{\bar{0}}$ such that $\operatorname{ad}c$ is semisimple (such elements x are called *homological*).

8.1. DS-*functors: construction and basic properties.* The DS-functors were introduced in [5]; we use a slight generalization (see [6] for a more in-depth treatment). For a $\mathfrak{g}$-module M and $u\in\mathfrak{g}$ we set

$$M^u:=\operatorname{Ker}_M u.$$

Let M be a $\mathfrak{g}$-module on which c acts semisimply. Write $DS_xM=M_x:=M^x/(\operatorname{Im}x\cap M^x)$. Then $\mathfrak{g}^x$ and $\mathfrak{g}_x$ are Lie superalgebras, where x acts via the adjoint action.

Observe that M^x, xM^c are $\mathfrak{g}^x$-invariant and $[x,\mathfrak{g}^c]M^x\subset xM^c$, so $\mathrm{DS}_x(M)$ is a $\mathfrak{g}^x$-module and $\mathfrak{g}_x$-module. This gives the functor $\mathrm{DS}_x:M\mapsto\mathrm{DS}_x(M)$ from the category of $\mathfrak{g}$-modules with semisimple action of c to the category of $\mathfrak{g}_x$-modules.

There are canonical isomorphisms $\mathrm{DS}_x(\Pi(N))\cong\Pi(\mathrm{DS}_x(N))$ and

$$\mathrm{DS}_x(M)\otimes\mathrm{DS}_x(N)\cong\mathrm{DS}_x(M\otimes N).$$

8.2. ***The map* ds_x.** Let $\mathfrak{a}$ be any subalgebra of $\mathfrak{g}$. Write $\mathfrak{a}^x$ for the kernel of $\operatorname{ad} x$ on $\mathfrak{a}$, and similarly for $\mathfrak{a}^c$. We view

$$\mathfrak{a}_x := \mathfrak{a}^x/([\mathfrak{g}^c, x] \cap \mathfrak{a}^c)$$

as a subalgebra of $\mathfrak{g}_x$.

Let $\mathcal{C}(\mathfrak{g})$ be a full subcategory of the category of $\mathfrak{g}$-modules with semisimple action of c and $\mathcal{C}(\mathfrak{a})$ (resp., $\mathcal{C}(\mathfrak{g}^x), \mathcal{C}(\mathfrak{a}^x)$) be a full subcategory of the category of $\mathfrak{a}$-modules (resp., $\mathfrak{g}^x$, $\mathfrak{a}^x$) such that the restriction functors

$$\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{a}} : \mathcal{C}(\mathfrak{g}) \to \mathcal{C}(\mathfrak{a}), \quad \mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x} : \mathcal{C}(\mathfrak{g}) \to \mathcal{C}(\mathfrak{g}^x), \quad \mathrm{Res}^{\mathfrak{a}}_{\mathfrak{a}^x} : \mathcal{C}(\mathfrak{a}) \to \mathcal{C}(\mathfrak{a}^x).$$

are well-defined and that for each $N \in \mathcal{C}(\mathfrak{g})$ the $\mathfrak{g}^x$-modules N^x and xN^c lie in $\mathcal{C}(\mathfrak{g}^x)$ (note that N^x and xN^c are submodules of $\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x}(N)$).

We denote by $\mathcal{C}(\mathfrak{g}_x)$ (resp., by $\mathcal{C}(\mathfrak{a}_x)$) the full category of $\mathfrak{g}_x$-modules N satisfying $\mathrm{Res}^{\mathfrak{g}_x}_{\mathfrak{g}^x}(N) \in \mathcal{C}(\mathfrak{g}^x)$ (resp., $\mathrm{Res}^{\mathfrak{a}_x}_{\mathfrak{a}^x}(N) \in \mathcal{C}(\mathfrak{a}^x)$).

For $\mathfrak{m} = \mathfrak{g}, \mathfrak{a}, \mathfrak{g}^x, \mathfrak{a}^x, \mathfrak{g}_x, \mathfrak{a}_x$ we denote the reduced Grothendieck group $\mathscr{K}_-(\mathcal{C}(\mathfrak{m}))$ by $R(\mathfrak{m})$, for ease of notation.

8.2.1. Take $M \in \mathcal{C}(\mathfrak{g})$ and set $N := \mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x} M$. The action of x gives a $\mathfrak{g}^x$-homomorphism $\theta_N : N^c \to \Pi(N^c)$. One has

$$\theta_N \theta_{\Pi(N)} = 0, \quad \mathrm{Im}\ \theta_{\Pi(N)} = \Pi(\mathrm{Im}\, \theta_N)$$

and $\mathrm{DS}_x(M) = \mathrm{Ker}\, \theta_N / \mathrm{Im}\ \theta_{\Pi(N)}$ as $\mathfrak{g}^x$-modules. Using the exact sequences

$$0 \to \mathrm{Im}\ \theta_{\Pi(N)} \to \mathrm{Ker}\, \theta_N \to \mathrm{DS}_x(M) \to 0, \qquad 0 \to \mathrm{Ker}\, \theta_N \to N^c \to \mathrm{Im}\ \theta_N \to 0$$

we obtain $[\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x} M^c] = [\mathrm{DS}_x(M)]$ in $R(\mathfrak{g}^x)$. However if M_r is the $r \neq 0$ eigenspace of c on M, $x : M_r \to M_r$ will define an $\mathfrak{g}^x$-equivariant isomorphism of M_r, and thus we have $[\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x} M_r] = 0$. It follows that $[\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x} M] = [\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x} M^c] = [\mathrm{DS}_x(M)]$. Since $\mathrm{DS}_x(M)$ is a $\mathfrak{g}_x$-module this gives the commutative diagram

$$\begin{tikzcd} R(\mathfrak{g}) \arrow[r] \arrow[dr, "\mathrm{ds}_x"'] & R(\mathfrak{g}^x) \\ & R(\mathfrak{g}_x) \arrow[u] \end{tikzcd} \tag{17}$$

where $\mathrm{ds}_x : R(\mathfrak{g}) \to R(\mathfrak{g}_x)$ is given by $\mathrm{ds}_x([M]) := [\mathrm{DS}_x(M)]$, and the two other arrows are induced by the restriction functors $\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x}$, $\mathrm{Res}^{\mathfrak{g}_x}_{\mathfrak{g}^x}$ respectively.

8.2.2. *Remark.* If $\mathcal{C}(\mathfrak{g}), \mathcal{C}(\mathfrak{g}_x)$ are closed under $\otimes$, then ds_x is a ring homomorphism.

8.2.3. *Example.* Suppose that $\mathfrak{g}$ is quasireductive. Recall that $\mathcal{F}(\mathfrak{g})$ a rigid tensor category with the duality $N \mapsto N^*$ given by the antiautomorphism $-\mathrm{Id}$; since DS_x is a tensor functor, it preserves $*$-duality, so

$$\mathrm{ds}_x : \mathscr{K}_-(\mathfrak{g}) \ \to \ \mathscr{K}_-(\mathfrak{g}_x)$$

is a ring homomorphism compatible with $*$.

8.3. ds_x *and restriction.* We present results which explain the relationship between ds_x and the restriction functor.

Lemma 8.1. *Suppose that we have a splitting $\mathfrak{g}_x \subseteq \mathfrak{g}^x$ so that $\mathfrak{g}^x = \mathfrak{g}_x \ltimes [x, \mathfrak{g}^c]$. Then for M in $\mathcal{C}(\mathfrak{g})$ we have*

$$\mathrm{ds}_x[M] = [\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}_x} M].$$

Proof. This follows immediately by applying the restriction $R(\mathfrak{g}^x) \to R(\mathfrak{g}_x)$ to our equality $[DS_x M] = [\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{g}^x} M]$. □

Lemma 8.2. *Let $y \in \mathfrak{g}_{\bar{1}}$ with $[y, y] = d$ where $\mathrm{ad}\, d$ acts semisimply on $\mathfrak{g}$ and d, $c + d$ act semisimply on all modules in $\mathcal{C}(\mathfrak{g})$. Suppose further that $[x, y] = 0$, and that we have splittings*

$$\mathfrak{g}^y \subseteq \mathfrak{g}_y \ltimes [y, \mathfrak{g}^d], \quad \mathfrak{g}^{x+y} = \mathfrak{g}_{x+y} \ltimes [x+y, \mathfrak{g}^{c+d}],$$

Furthermore suppose that under these splittings, $x \in \mathfrak{g}_y$ and

$$(\mathfrak{g}_y)^x = \mathfrak{g}_{x+y} \ltimes [x, \mathfrak{g}_y^c].$$

Then we have

$$ds_{x+y} = ds_x \circ ds_y : R(\mathfrak{g}) \to R(\mathfrak{g}_{x+y})$$

Proof. This follows immediately from Lemma 8.1 and the corresponding statement for restriction. □

Proposition 8.1. *We have the following commutative diagram*

$$\begin{tikzcd} R(\mathfrak{g}) \arrow[r] \arrow[d, "\mathrm{ds}_x"'] & R(\mathfrak{a}^x) \\ R(\mathfrak{g}_x) \arrow[r] & R(\mathfrak{a}_x) \arrow[u, "\mathrm{res}^{\mathfrak{a}_x}_{\mathfrak{a}^x}"'] \end{tikzcd}$$

where the horizontal arrows are induced by the corresponding restriction functors and $\mathrm{res}^{\mathfrak{a}_x}_{\mathfrak{a}^x}$ is induced by the morphism $\mathfrak{a}^x \to \mathfrak{a}_x$.

Proof. The restriction functors give the commutative diagram

$$\begin{tikzcd} \mathcal{C}(\mathfrak{g}^x) \arrow[r] & \mathcal{C}(\mathfrak{a}^x) \\ \mathcal{C}(\mathfrak{g}_x) \arrow[r] \arrow[u] & \mathcal{C}(\mathfrak{a}_x) \arrow[u] \end{tikzcd}$$

which, in combination with (17) gives the diagram

$$\begin{tikzcd} R(\mathfrak{g}) \arrow[r] \arrow[dr, "\mathrm{ds}_x"'] & R(\mathfrak{g}^x) \arrow[r] & R(\mathfrak{a}^x) \\ & R(\mathfrak{g}_x) \arrow[r] \arrow[u] & R(\mathfrak{a}_x) \arrow[u] \end{tikzcd}$$

where all arrows except ds_x are induced by the restriction functors. By (17), the above diagram is commutative, and we obtain our result. □

8.3.1. *Example: $\mathcal{F}(\mathfrak{g})$ for $\mathfrak{g}$ quasireductive.* Let E be a simple $\mathfrak{a}^x$-module. By Proposition 8.1 a finite-dimensional $\mathfrak{g}$-module N we have

$$\mathrm{smult}(\mathrm{DS}_x(N);\, E) = \mathrm{smult}(\mathrm{Res}^{\mathfrak{g}}_{\mathfrak{a}^x} N;\, E).$$

For example, let $\mathfrak{g}$ be a classical Lie superalgebra in the sense of [16] and $\mathfrak{a} := \mathfrak{t}$ be a Cartan subalgebra of $\mathfrak{g}_{\bar{0}}$. The map $R(\mathfrak{g}) \to R(\mathfrak{t})$ is given $[N] \mapsto \mathrm{sch}\, N$. If $\mathfrak{t}_x$ is a Cartan subalgebra of $(\mathfrak{g}_x)_{\bar{0}}$, then the composed map $R(\mathfrak{g}) \to R(\mathfrak{t}_x)$ is given by $[N] \mapsto \mathrm{sch}\, \mathrm{DS}_x(N)$. If we fix an embedding $\mathfrak{t}_x \to \mathfrak{t}^x$, we obtain the Hoyt–Reif formula [13]

$$\mathrm{sch}\, \mathrm{DS}_x(N) = (\mathrm{sch}\, N)|_{\mathfrak{t}_x}.$$

8.4. ***A special case.*** Consider the case when $\mathfrak{g}$, $\mathfrak{a} = \mathfrak{h}$ are as in Section 2.1. Denote by F^x the restriction of the form F to $\mathfrak{h}^x$ and set

$$I_x := \{\lambda \in \mathfrak{t}^* \mid \lambda([x, \mathfrak{g}^c] \cap \mathfrak{t}) = 0\}.$$

By assumption, given $\lambda \in I_x$ we have that $\lambda|_{\mathfrak{t}^x}$ lies in the subspace $(\mathfrak{t}_x)^*$. For $\lambda \in I_x$ we denote this element by $\lambda_x \in (\mathfrak{t}_x)^*$.

Note that for $\nu \in I_x$ satisfying $\mathrm{rank}\, F^x_{\nu|_{\mathfrak{t}^x}} = \mathrm{rank}\, F_\nu$ the module $\mathrm{Res}^{\mathfrak{h}}_{\mathfrak{h}_x} C(\nu)$ is simple, so $\mathrm{Res}^{\mathfrak{h}}_{\mathfrak{h}_x} C(\nu) \cong \Pi^i C(\nu_x)$ for some i.

Corollary 8.1. *Take $N \in \mathcal{F}(\mathfrak{g})$. If $\mathrm{sch}_{\mathfrak{h}}(N) = \sum_\nu m_\mu [C(\mu)]$, then*

$$\mathrm{sch}_{\mathfrak{h}_x}\big(\mathrm{DS}_x(N)\big) = \sum_{\mu \in I_x : \mathrm{rank}\, F^x_\mu = \mathrm{rank}\, F_\mu} m_\mu (-1)^{i_\mu}\, [C(\mu_x)].$$

where $\mathrm{Res}^{\mathfrak{h}}_{\mathfrak{h}_x} C(\nu) \cong \Pi^{i_\mu} C(\mu_x)$.

Proof. Recall that $\mathrm{sch}_{\mathfrak{h}}$ gives an embedding of $\mathcal{K}_-(\mathfrak{g})$ to $\mathcal{K}_-(\mathfrak{h})$. Applying Proposition 8.1 to $\mathcal{F}(\mathfrak{g})$ we obtain for $R(\mathfrak{m}) := \mathcal{K}_-(\mathfrak{m})$ the commutative diagram

$$\begin{tikzcd} R(\mathfrak{g}) \arrow[r, "\mathrm{sch}_{\mathfrak{h}^x}"] \arrow[d, "\mathrm{ds}_x"'] & R(\mathfrak{h}^x) \\ R(\mathfrak{g}_x) \arrow[r, "\mathrm{sch}_{\mathfrak{h}_x}"] & R(\mathfrak{h}_x) \arrow[u, "\mathrm{res}^{\mathfrak{h}_x}_{\mathfrak{h}^x}"'] \end{tikzcd}$$

where $\mathrm{res}^{\mathfrak{h}^x}_{\mathfrak{h}_x} : R_{\mathfrak{h}_x} \to R^{\mathfrak{h}_x}$ is induced by the map $\mathfrak{h}^x \to \mathfrak{h}_x$. In light of Proposition 3.1(v) we have $\mathrm{sch}_{\mathfrak{h}^x}(C(\mu)) = 0$ except for the case $\mathrm{rank}\, F^x_\mu = \mathrm{rank}\, F_\mu$. □

8.4.1. *Example.* If, in addition, $(\mathfrak{g}_x)^{\mathfrak{a}_x}_{\bar{0}} = \mathfrak{a}_x$, then, by Corollary 5.2, $\mathrm{sch}_{\mathfrak{a}_x}$ gives an embedding of the reduced Grothendieck ring of $\mathfrak{g}_x$ to $\mathcal{K}_-(\mathfrak{h}_x)$.

For $\mathfrak{g} = \mathfrak{gl}(m|n)$, $\mathfrak{osp}(m|n)$, $\mathfrak{p}_n$, $\mathfrak{q}_n$, $\mathfrak{sq}_n$ and the exceptional Lie superalgebras, for each x we can choose a suitable $\mathfrak{h}$ such that $(\mathfrak{g}_x)^{\mathfrak{t}_x}_{\bar{0}} = \mathfrak{t}_x$; see [5; 23; 10].

9. The reduced Grothendieck ring for $\mathcal{F}(\mathfrak{q}_n)$

In this section we describe the reduced Grothendieck ring for $\mathcal{F}(\mathfrak{q}_n)$, the category of finite-dimensional $\mathfrak{q}_n$-modules with semisimple action of $(\mathfrak{q}_n)_{\bar{0}}$. In addition we will explicitly describe the homomorphisms $\mathrm{ds}_s : \mathscr{K}_-(\mathfrak{q}_n) \to \mathscr{K}_-(\mathfrak{q}_{n-2s})$ induced by the DS functor, to be defined. Wherever it is not stated, we set $\mathfrak{g} := \mathfrak{q}_n$. We will mainly concentrate on the category $\mathcal{F}(\mathfrak{q}_n)_{\mathrm{int}}$ which is the full subcategory of finite-dimensional $\mathfrak{q}_n$-modules with integral weights, and then reduce the corresponding results for $\mathcal{F}(\mathfrak{q}_n)$ to $\mathcal{F}(\mathfrak{q}_n)_{\mathrm{int}}$.

9.1. ***Structure of*** $\mathfrak{q}_n$**.** Recall that $\mathfrak{q}_n$ is the subalgebra of $\mathfrak{gl}(n|n)$ consisting of the matrices with the block form

$$T_{A,B} := \begin{pmatrix} A & B \\ B & A \end{pmatrix}$$

9.1.1. One has $\mathfrak{g}_{\bar{0}} = \mathfrak{gl}_n$. The group GL_n acts on $\mathfrak{g}$ by the inner automorphisms; all triangular decompositions of $\mathfrak{q}_n$ are GL_n-conjugated. We denote by $\mathfrak{t}$ the Cartan subalgebra of $\mathfrak{gl}_n$ spanned by the elements $h_i = T_{E_{ii},0}$ for $i = 1, \dots, n$, where E_{ij} denotes the (i, j) elementary matrix. Let $\{\varepsilon_i\}_{i=1}^n \subset \mathfrak{t}^*$ be the basis dual to $\{h_i\}_{i=1}^n$. The algebra $\mathfrak{h} := \mathfrak{q}_n^{\mathfrak{t}}$ is a Cartan subalgebra of $\mathfrak{q}_n$; one has $\mathfrak{h}_{\bar{0}} = \mathfrak{t}$. The elements $H_i := T_{0,E_{ii}}$ form a basis of $\mathfrak{h}_{\bar{1}}$; one has $[H_i, H_j] = 2\delta_{ij} h_i$.

9.1.2. We write $\lambda \in \mathfrak{t}^*$ as $\lambda = \sum_{i=1}^n \lambda_i \varepsilon_i$ and denote by $\mathrm{Nonzero}(\lambda)$ the set of nonzero elements in the multiset $\{\lambda_i\}_{i=1}^n$ and by $\mathrm{zero}\,\lambda$ the number of zeros in the multiset $\{\lambda_i\}_{i=1}^n$. Recall that rank F_λ is equal to the cardinality of $\mathrm{Nonzero}(\lambda)$ $(= n - \mathrm{zero}\,\lambda)$.

We call a weight $\lambda \in \mathfrak{t}^*$ *integral* (resp., *half-integral*) if $\lambda_i \in \mathbb{Z}$ (resp., $\lambda_i - \frac{1}{2} \in \mathbb{Z}$) for all i. We call a weight λ *typical* if $\lambda_i + \lambda_j \neq 0$ for all i, j and *atypical* otherwise; in particular if λ is typical then $\mathrm{zero}(\lambda) = 0$.

We fix the usual triangular decomposition: $\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}$, where $\Delta^+ = \{\varepsilon_i - \varepsilon_j\}_{1 \le i < j \le n}$.

9.2. ***The monoid*** Ξ ***and cores.*** We denote by Ξ the set of finite multisets $\{a_i\}_{i=1}^s$ with $a_i \in \mathbb{C} \setminus \{0\}$ and $a_i + a_j \neq 0$ for all $1 \le i, j \le s$.

We assign to each finite multiset $A := \{a_i\}_{i=1}^s$ with $a_i \in \mathbb{C}$ the multiset $\mathrm{Core}(A) \in \Xi$, obtained by throwing out all zeros and the maximal number of elements a_i, a_j with $i \neq j$ and $a_i + a_j = 0$; for example,

$$\mathrm{Core}(\{1, 1, -1, -1\} = \varnothing \quad \text{and} \quad \mathrm{Core}(\{1, 1, 0, 0, 0, -1\}) = \{1\}.$$

We view Ξ as a commutative monoid with respect to the operation

$$A \diamond B := \mathrm{Core}(A \cup B)$$

($\varnothing$ is the identity element in Ξ).

For $\lambda \in \mathfrak{t}^*$ we set

$$\mathrm{Core}(\lambda) := \mathrm{Core}(\{\lambda_i\}_{i=1}^n) = \mathrm{Core}(\mathrm{Nonzero}(\lambda)),$$

and denote by χ_λ the central character of $L(\lambda)$. From [24] it follows that $\chi_\lambda = \chi_\nu$ if and only if $\mathrm{Core}(\lambda) = \mathrm{Core}(\nu)$. Observe that $\lambda \in I_0$ if and only if the cardinality of $\mathrm{Core}(\lambda)$ is even.

9.2.1. *The relation* $\sim$. Recall that we write $\lambda \sim \nu$ if $L(\lambda)$, $L(\nu)$ lie in the same block in the BGG category $\mathcal{O}$. By the above,

$$\lambda \sim \nu \implies \mathrm{Core}(\lambda) = \mathrm{Core}(\nu).$$

It is known that the above implication becomes an equivalence if both λ, ν are integral or half-integral; in fact it follows from the following fact:

$$\lambda_i + \lambda_j = 0 \implies \lambda - \varepsilon_i + \varepsilon_j \sim \lambda \tag{18}$$

(this easily follows from the formula for Shapovalov determinants established in [9, Theorem 11.1]: from this formula it follows that for $i < j$ the module $M(\lambda)$ has a primitive vector of weight $\lambda - \varepsilon_i + \varepsilon_j$ if λ is a "generic weight" satisfying $\lambda_i + \lambda_j = 0$; the usual density arguments (see [2],[17]) imply that $M(\lambda)$ has a primitive vector of weight $\lambda - \varepsilon_i + \varepsilon_j$ if λ is any weight satisfying $\lambda_i + \lambda_j = 0$).

9.2.2. *Dominant weights.* Recall that $P^+(\mathfrak{g})$ denotes the set of dominant weights, i.e.,

$$P^+(\mathfrak{g}) := \{\lambda \in \mathfrak{t}^* \mid \dim L(\lambda) < \infty\}.$$

By [19], $\lambda \in P^+(\mathfrak{g})$ if and only if $\lambda_i - \lambda_{i+1} \in \mathbb{N}$ and $\lambda_i = \lambda_{i+1}$ implies $\lambda_i = 0$. This implies the following properties:

(1) $P^+(\mathfrak{g}) \cap I_0 = P^+(\mathfrak{g}_{\bar{0}})' \cap I_0$; in particular

$$P^+(\mathfrak{g}) = \{\lambda \in P^+(\mathfrak{g}_{\bar{0}}) : \det(w|_{\mathfrak{h}_{\bar{1}}/K_\lambda}) = 1 \text{ for all } w \in \mathrm{Stab}_W \lambda\}.$$

(2) if $\lambda \in P^+(\mathfrak{g})$ is atypical, then λ is either integral or half-integral;

(3) the set $\mathrm{Nonzero}(\lambda)$ uniquely determines a dominant weight for a fixed n.

9.2.3. *Grading on* $C(\nu)$. For $\nu \in \mathfrak{t}^*$, we set

$$T_\nu := H_{i_1} \dots H_{i_k}$$

where $\mathrm{Nonzero}(\nu) = \{\nu_{i_1} \geq \nu_{i_2} \geq \dots \geq \nu_{i_k}\}$, and if $\nu_{i_j} = \nu_{i_{j+1}}$ then we require that $i_j \geq i_{j+1}$. This formula determines T_ν uniquely; for $\nu \in P(\mathfrak{g})'$ note that we have $T_{w\nu} = wT_\nu$ for each $w \in W$.

Since $T_\nu^2 = (-1)^{k(k-1)/2} h_{i_1} \dots h_{i_k}$ the function $t :\ \mathfrak{t}^* \to \{c \in \mathbb{C} \mid c > 0\}$ is given by

$$t(\lambda)^2 = (-1)^{k(k-1)/2} \prod_{i:\lambda_i \neq 0} \lambda_i, \quad \text{where } k := \mathrm{rank}\, F_\lambda;$$

for $\lambda \in I_0$ we obtain $t(\lambda)^2 = (-1)^{(\mathrm{rank}\, F_\lambda)/2} \prod_{i:\lambda_i \neq 0} \lambda_i$.

As explained in Section 3.4.3, this function t together with a highest weight λ determines uniquely irreducible modules $L(\lambda)$ for each $\lambda \in P^+(\mathfrak{q}_n)$.

9.2.4. *Remark.* If $\nu\sim 0$, then $t(\nu)\in\mathbb{R}^+$. It follows that if $\lambda\in\mathfrak{t}^*$, $\nu\sim 0$, and $[C(\lambda)][C(\nu)]=\pm[C(\lambda+\nu)]$, then $t(\lambda)t(\nu)=t(\lambda+\nu)$.

9.2.5. Note that $\lambda=-w_0\lambda$ if and only if $\operatorname{Core}(\lambda)=\varnothing$; in this case Corollary 7.1 gives $L(\lambda)^*\cong\Pi^{(\operatorname{rank}F_\lambda)/2}L(\lambda)$.

9.2.6. *Example.* For $\alpha=\varepsilon_i-\varepsilon_j$ one has $T_\alpha=H_jH_i$. Since for $e\in\mathfrak{g}_\alpha\cap\mathfrak{g}_{\bar 0}$ one has $(\operatorname{ad}H_j)(\operatorname{ad}H_i)e=e$, we obtain $\mathfrak{g}_\alpha\cong C_\alpha$. Moreover, $S^kC(\alpha)=C(k\alpha)$ for $\alpha\in\Delta$.

9.3. ***Embedding into exterior algebra.*** Let $\mathfrak{h}^{\mathbb{Z}}$ be the Lie subalgebra of $\mathfrak{h}$ over $\mathbb{Z}$ generated by $H_1,\dots,H_n$, and let $\mathcal{U}(\mathfrak{h}^{\mathbb{Z}})$ be the integral enveloping algebra of $\mathfrak{h}^{\mathbb{Z}}$. Consider the canonical epimorphism $\phi_0:\mathcal{U}(\mathfrak{h}^{\mathbb{Z}})\to\mathcal{S}(\mathfrak{h}^{\mathbb{Z}}_{\bar 1})$, where $\mathcal{S}(\mathfrak{h}^{\mathbb{Z}}_{\bar 1})$ is the exterior ring generated by $\xi_i:=\phi_0(H_i)$. Note that $\mathcal{S}(\mathfrak{h}^{\mathbb{Z}}_{\bar 1})$ is a $\mathbb{Z}$-graded supercommutative ring, free over $\mathbb{Z}$ with basis $\xi_{i_1},\dots,\xi_{i_j}$ with $1\le i_1<i_2<\dots<i_j\le n$.

Let $\mathcal{C}_{\mathbb{R}}$ be the subcategory of $\mathcal{F}(\mathfrak{h})$ consisting of weights λ such that $\lambda_i\in\mathbb{R}$ for all i. In particular, $t_\lambda/|t_\lambda|\in\{1,\sqrt{-1}\}$.

9.3.1. We view the ring $B:=\mathbb{Z}[e^\nu:\nu\in\mathfrak{t}^*]\otimes_{\mathbb{Z}}\mathcal{S}(\mathfrak{h}^{\mathbb{Z}}_{\bar 1})$ as a $\mathbb{Z}$-graded supercommutative ring by defining the degree of $\mathbb{Z}[e^\nu:\nu\in\mathfrak{t}^*]$ to be zero. We construct the ring B^{spoil} as in Section 5.5.

Proposition 9.1. *The map* $[C(\lambda)]\mapsto\varepsilon^i\dfrac{t(\lambda)}{|t(\lambda)|}\cdot e^\lambda\phi_0(T_\lambda)$ *for* $\lambda\in I_i$ *gives a ring monomorphism*

$$\mathscr{K}_-(\mathcal{C}_{\mathbb{R}})\hookrightarrow B^{\text{spoil}}\otimes_{\mathbb{Z}}\mathbb{Z}[\sqrt{-1}]$$

which is compatible with the action of W. *One has*

$$\phi_0(T_\lambda)=\xi_{i_1}\dots\xi_{i_k},$$

where $\operatorname{Nonzero}(\lambda)=\{\lambda_{i_1}\le\dots\le\lambda_{i_k}\}$.

Proof. That the map is injective and W-equivariant is straightforward. We show that it is an algebra homomorphism. First observe that

$$[C(\nu)][C(\lambda)]\neq 0\iff[C(\nu)][C(\lambda)]=\pm[C(\lambda+\nu)]\iff\phi_0(T_\lambda T_\nu)\neq 0\iff T_\lambda T_\nu=\pm T_{\lambda+\nu}.$$

Suppose that $[C(\nu)][C(\lambda)]\neq 0$. Then we write $T_\lambda T_\nu=(-1)^jT_{\lambda+\nu}$ and $\phi_0(T_\lambda)\phi_0(T_\nu)=(-1)^j\phi_0(T_{\lambda+\nu})$.

Fix even vectors $v_\lambda\in C(\lambda)_{\bar 0}$ and $v_\nu\in C(\nu)_{\bar 0}$. Since $H_iC(\lambda)=0$ if $\lambda_i=0$ we have

$$T_\lambda T_\nu(v_\lambda\otimes v_\nu)=T_\lambda v_\lambda\otimes T_\nu v_\nu=\frac{t(\lambda)t(\nu)}{|t(\lambda)t(\nu)|}v_\lambda\otimes v_\nu.$$

Note that $v_\lambda\otimes v_\nu$ is an even vector of $\Pi^iC(\lambda+\nu)\cong C(\lambda)\otimes C(\nu)$. On the other hand we have

$$T_\lambda T_\nu(v_\lambda\otimes v_\nu)=(-1)^jT_{\lambda+\nu}(v_\lambda\otimes v_\nu)=(-1)^{j+i}t(\lambda+\nu)v_\lambda\otimes v_\nu.$$

Thus $\frac{t(\lambda+\nu)}{|t(\lambda+\nu)|}=(-1)^{i+j}\frac{t(\lambda)t(\nu)}{|t(\lambda+\nu)|}$. From these equalities it is easy to check our map is a homomorphism. Injectivity is straightforward. ∎

9.3.2. *Remark.* If we extend scalars to $\mathbb{C}$, the above map defines an embedding

$$\mathscr{K}(\mathfrak{h})_{\mathbb{C}} \hookrightarrow B_{\bar{0}} \otimes_{\mathbb{Z}} \mathbb{C}.$$

However this kills the two-torsion part of $\mathscr{K}(\mathfrak{h})$.

9.3.3. One has

$$\phi_0(T_\lambda T_\nu) \neq 0 \iff T_\lambda T_\nu = \pm T_{\lambda+\nu} \implies \mathrm{Core}(\lambda+\nu) = \mathrm{Core}(\lambda) \diamond \mathrm{Core}(\nu).$$

For example, for $[C(2, 0, 0, 1)][C(0, -2, -1, 0)] = \pm[C(2, -2, -1, 1)]$, we have

$$\mathrm{Core}(\lambda) = \{2, 1\}, \quad \mathrm{Core}(\nu) = \{-2, -1\}, \quad \mathrm{Core}(\lambda+\nu) = \varnothing.$$

9.3.4. Recall that $\mathscr{K}_-(\mathfrak{h})$ has a finite $\mathbb{Z}$-grading (see (11)). By Section 9.3.3, $\mathscr{K}_-(\mathfrak{h})$ is also Ξ-graded: set $\mathscr{K}_-(\mathfrak{h})_A$ to be spanned by $[C(\lambda)]$ with $\mathrm{Core}(\lambda) = A$. Then

$$\mathscr{K}_-(\mathfrak{h}) = \bigoplus_{A\in\Xi} \mathscr{K}_-(\mathfrak{h})_A \quad \text{with} \quad \mathscr{K}_-(\mathfrak{h})_A \mathscr{K}_-(\mathfrak{h})_B \subset \mathscr{K}_-(\mathfrak{h})_{A\diamond B},$$

Note that $\mathscr{K}_-(\mathfrak{h})_A = 0$ if the cardinality of A is odd or greater than n. Further $\mathscr{K}_-(\mathfrak{h})_A$ is W-stable.

Corollary 9.1. (1) *The subring $\mathscr{K}_-(\mathfrak{h})_\varnothing \cap \mathscr{K}_-(\mathfrak{h})_{\mathrm{int}}$ is spanned by $[C(\nu)]$ with $\nu \sim 0$. This subring is generated by $[C(k\alpha)]$ for $\alpha \in \Delta$ and $k \in \mathbb{Z}$.*

(2) *If $\nu \sim 0$ and $[C(\nu)][C(\lambda)] \neq 0$, then $[C(\nu)][C(\lambda)] = \pm[C(\lambda+\nu)]$ and $\lambda+\nu \sim \lambda$.*

Proof. The assertion (i) follows from Section 9.2.1. By (i) it is enough to verify (ii) for $\nu := k(\varepsilon_i - \varepsilon_j)$. In this case the inequality $[C(\nu)][C(\lambda)] \neq 0$ implies $\lambda_i = \lambda_j = 0$. By (18) for such λ one has $\lambda + k(\varepsilon_i - \varepsilon_j) \sim \lambda$. □

9.3.5. *Remark.* All integral, nontypical blocks $\mathcal{B}$ of $\mathfrak{q}_n$ have that $\Pi\mathcal{B} = \mathcal{B}$ (indeed, if $\mathcal{B}$ is integral and not typical then it admits a simple module $L(\lambda)$ such that $\mathrm{zero}(\lambda) > 0$; now we conclude with [12, Theorem 4.1]). Thus we may consider for such blocks the corresponding reduced Grothendieck group $\mathscr{K}_-(\mathcal{B})$. Then the embedding in Proposition 9.1 exists over the coefficient ring $\mathbb{Z}[\sqrt{-1}]$. If we let $\mathcal{B}_0$ denote the principal block of $\mathfrak{q}_n$, then $\mathscr{K}_-(\mathcal{B}_0) = \mathscr{K}_-(\mathfrak{h})_\varnothing \cap \mathscr{K}_-(\mathfrak{h})_{\mathrm{int}}$ is a ring. In this case, because $t(\lambda) \in \mathbb{R}^+$ for all $\lambda \sim 0$, the map in Proposition 9.1 descends to an embedding with *integral* coefficients for $\mathscr{K}_-(\mathcal{B}_0)$.

9.4. *Supercharacters of some highest weight modules.*

Lemma 9.1. $\mathrm{sch}_{\mathfrak{h}} M(\lambda) = \sum_{\nu\sim\lambda} m_\nu [C(\nu)].$

Proof. We start from $\lambda = 0$; in this case $\mathrm{sch}_{\mathfrak{h}} M(0) = \mathrm{sch}_{\mathfrak{h}} \mathcal{U}(\mathfrak{n}^-)$. Let $\alpha_1, \dots, \alpha_N$ denote the negative roots of $\mathfrak{g}$. We have

$$\mathcal{U}(\mathfrak{n}^-)_\nu = \bigoplus_{\sum k_j\alpha_j = \nu} \bigotimes_j S^{k_j} \mathfrak{g}_{\alpha_j}. \tag{19}$$

By Section 9.2.6, $S^k\mathfrak{g}_\alpha \cong C(k\alpha)$. Since $k\alpha \sim 0$, the required formula follows from Corollary 9.1(i).

For arbitrary λ one has $\mathrm{sch}_\mathfrak{h}\, M(\lambda) = \mathrm{sch}_\mathfrak{h}\, \mathcal{U}(\mathfrak{n}^-) \otimes [C(\lambda)]$; the result follows from Corollary 9.1(ii). $\square$

We need some terminology for the following proposition: for roots $\alpha_1 = \epsilon_{i_1} - \epsilon_{j_1}, \alpha_2 = \epsilon_{i_2} - \epsilon_{j_2}$, we write $\alpha_1 \prec \alpha_2$ if $\max(i_1, j_1) < \min(i_2, j_2)$. If $\alpha = \epsilon_i - \epsilon_j$ is any root, we call the set $\{i, j\}$ its support. We note that if we have roots $\alpha_1, \dots, \alpha_j$ with nonoverlapping supports, and positive integers $k_1, \dots, k_j$, then $S^{j_1}\mathfrak{g}_{\alpha_1} \otimes \cdots \otimes S^{k_j}\mathfrak{g}_{\alpha_j}$ is an irreducible $\mathfrak{h}$-module.

Proposition 9.2. *We have*

$$\mathrm{sch}_\mathfrak{h}\, \mathcal{U}(\mathfrak{n}^-) = \sum_{\nu \sim 0} m_\nu [C(\nu)]$$

where each coefficient m_ν is either 0 or ± 1.

Further if $m_\nu \neq 0$, then ν may uniquely be written as

$$\nu = k_1(\alpha_{11} + \cdots + \alpha_{1j_1}) + \cdots + k_r(\alpha_{r1} + \cdots + \alpha_{rj_r}),$$

where all of $k_1, \dots, k_r$ are distinct, the supports of all α_{ij} are distinct, and $\alpha_{i1} \prec \cdots \prec \alpha_{ij_i}$. In this case

$$m_\nu [C(\nu)] = \prod_{i=1}^{r} [S^{k_i}\mathfrak{g}_{\alpha_{i1}}] \cdots [S^{k_i}\mathfrak{g}_{\alpha_{ij_i}}].$$

Proof. We use the embedding of Proposition 9.1. Let y_ν be the image of $[(\mathcal{U}(\mathfrak{n}^-))_\nu]$ in $R(\mathfrak{t}) \otimes \mathcal{S}(\mathfrak{h}_{\bar 1})_{\bar 0}$ (see Proposition 9.1). Since $t_{k\alpha} = k$ for $\alpha \in \Delta$, by (19) one has

$$y_\nu = e^\nu \sum_{\substack{(k_1,\dots,k_N) \\ \sum k_i\alpha_i = \nu}} \phi_0(T_{k\alpha_i}).$$

Recall that $\phi_0(T_{\varepsilon_p - \varepsilon_q}) = \xi_q \xi_p$. In particular, $\phi_0(T_\alpha T_\beta) = 0$ if $(\alpha|\beta) \neq 0$ (where $(-|-)$ stands for the usual form on $\mathfrak{t}^*$). Hence

$$y_\nu = e^\nu \sum_{(k_1,\dots,k_N) \in U} \phi_0(T_{k\alpha_i}), \qquad \text{where } U := \{(k_1, \dots, k_N) \mid \textstyle\sum k_i\alpha_i = \nu,\ k_i k_j = 0 \text{ for } (\alpha_i|\alpha_j) \neq 0\}.$$

If for any $(k_1, \dots, k_N) \in U$ we have $k_i \neq k_j$ for all nonzero k_i, k_j with $i \neq j$ then it is clear U is a singleton set and we are done.

Thus suppose that $(k_1, \dots, k_n)$ has $k_i = k_j \neq 0$ for some $i \neq j$, and without loss of generality suppose $i = 1,\ j = 2$. Write $\alpha_i = \varepsilon_{p_i} - \varepsilon_{q_i}$ for $i = 1, 2$; then $p_i > q_i$ and, since $(\alpha_1|\alpha_2) = 0$, the numbers p_1, q_1, p_2, q_2 are pairwise distinct. If $p_1 > q_2$ and $p_2 > q_1$ then we have negative roots

$$\alpha'_1 := \varepsilon_{p_1} - \varepsilon_{q_2}, \quad \alpha'_2 := \varepsilon_{p_2} - \varepsilon_{q_1}$$

with $\alpha_1 + \alpha_2 = \alpha'_1 + \alpha'_2$. We may assume that $\alpha'_1 = \alpha_3$ and $\alpha'_2 = \alpha_4$. Since $(\alpha_j|\alpha_1), (\alpha_s|\alpha_1) \neq 0$, one has $k_3 = k_4 = 0$, so

$$(k_1, \dots, k_N) = (k_1, k_1, 0, 0, k_5, \dots, k_N).$$

Observe that $(0,0,k_1,k_1,k_5,\dots,k_N)\in U$. One has

$$\phi_0(T_{\alpha_1}T_{\alpha_2}+T_{\alpha_1'}T_{\alpha_2'})=\xi_{p_1}\xi_{q_1}\xi_{p_2}\xi_{q_2}+\xi_{p_1}\xi_{q_2}\xi_{p_2}\xi_{q_1}=0.$$

Therefore we can substitute U by a smaller set, where $k_i=k_j$ implies that $\alpha_i=\varepsilon_{p_i}-\varepsilon_{q_i}$, $\alpha_j=\varepsilon_{p_j}-\varepsilon_{q_j}$ are such that $p_i>q_i>p_j>q_j$ or $p_j>q_j>p_i>q_i$. From this the result follows. □

Corollary 9.2. *One has* $\mathrm{sch}_{\mathfrak{h}}\, M(\lambda)=\sum_{\nu\sim\lambda} m_\nu[C(\nu)]$ *with* $m_\nu\in\{0,\pm1\}$.

Theorem 9.1. *For* $\lambda\in\mathfrak{t}^*$ *we have* $\mathrm{sch}_{\mathfrak{h}}\, L(\lambda)=\sum_{\nu\sim\lambda} k_\nu[C(\nu)]$.

Proof. With the help of Lemma 9.1, the proof works in the exact same fashion as Theorem 6.1. □

Corollary 9.3. *Let* $L(\nu)$ *be a finite-dimensional module. Then*

$$\mathrm{smult}(L(\lambda)\otimes L(\nu):L(\mu))\neq 0 \implies \mathrm{Core}(\mu)=\mathrm{Core}(\lambda)\diamond\mathrm{Core}(\nu).$$

When $\mathrm{smult}(L(\lambda)\otimes L(\nu):L(\mu))\neq 0$, *then either* $\lambda\in I_0$ *or* $\nu\in I_0$.

9.4.1. *Remark.* The Kac–Kazhdan modification of $\mathcal{O}$ introduced in [17] is closed under tensor product; the modules in this category are not always of finite length, but the multiplicity is well-defined (see [4]). In this category the above formula holds for arbitrary $L(\lambda)$ and $L(\nu)$.

9.4.2. Recall that $P(\mathfrak{g}):=P^+(\mathfrak{g})+\mathbb{Z}\Delta$ and for $\nu\in P^+(\mathfrak{g})$ we set

$$a_\nu:=\sum_{w\in W/\mathrm{Stab}_W\,\nu} w[C(\mu)].$$

Lemma 9.2. *Suppose that* $|\mathrm{Nonzero}(\nu)|$ *is odd with* $\nu_i=\nu_j\neq 0$ *for some* $i\neq j$; *then for* $\lambda\in P^+(\mathfrak{q}_n)$, $[L(\lambda)_\nu:C(\nu)]$ *is even.*

Proof. Assume that this does not hold. Without loss of generality we may assume $i=1$, $j=2$, and set $\alpha=\epsilon_1-\epsilon_2$. Let $\mathfrak{h}'$ to be the subalgebra of $\mathfrak{h}$ generated by $H_3,\dots,H_n$, and set $\mathfrak{q}_2(\alpha)$ to be the natural subalgebra of $\mathfrak{q}_n$ isomorphic to $\mathfrak{q}_2$ with weight ϵ_1,ϵ_2. Clearly $\mathfrak{q}_2(\alpha)\times\mathfrak{h}'$ is a subalgebra of $\mathfrak{q}_n$.

Then by Theorem 9.1, $\mathrm{Core}(\nu)=\mathrm{Core}(\lambda)$. We view

$$N:=\sum_{i\in\mathbb{Z}} L(\lambda)_{\nu+i\alpha}$$

as a $\mathfrak{q}_2(\alpha)\times\mathfrak{h}'$-module. We will write $\nu=k(\varepsilon_1+\varepsilon_2)+\nu'$ where ν' is the corresponding $\mathfrak{t}'=\mathfrak{h}'_{\bar 0}$-weight. We assume that $k>0$, with the case of $k<0$ being similar. By the representation theory of $\mathfrak{q}_2$, the only irreducible $\mathfrak{q}_2$-modules with $k(\epsilon_1+\epsilon_2)$ are those of the form $L_{\mathfrak{q}_2(\alpha)}(k+i;k-i)$ for some $i>0$. These are always typical, and are isomorphic to their parity shifts if and only if $i=k$.

Therefore, since rank F_ν is odd, one has

$$L_{\mathfrak{q}_2(\alpha)}(k+i;k-i)\boxtimes L_{\mathfrak{h}'}(\nu')=\begin{cases} L_{\mathfrak{q}_2(\alpha)\times\mathfrak{h}'}(\nu+i\alpha) & \text{if } i\neq k,\\ L_{\mathfrak{q}_2(\alpha)\times\mathfrak{h}'}(\nu+k\alpha)\oplus\Pi L_{\mathfrak{q}_2(\alpha)\times\mathfrak{h}'}(\nu+k\alpha) & i=k.\end{cases}$$

Because $L_{\mathfrak{q}_2(\alpha)}(k+i; k-i)$ is typical,

$$\dim L_{\mathfrak{q}_2(\alpha)\times\mathfrak{h}'}(\nu+i\alpha)_\nu = \begin{cases} 2\dim C(\nu) & \text{if } i\neq k,\\ \dim C(\nu) & \text{if } i=k.\end{cases}$$

Hence, to prove that $\dim N_\nu$ is divisible by $2\dim C(\nu)$ it is enough to show that $\mathrm{mult}(N; L_{\mathfrak{q}_2(\alpha)\times\mathfrak{h}'}(\nu+k\alpha))$ is even, where mult denotes that nongraded multiplicity. By the above, for $i\neq k$ one has

$$\dim L_{\mathfrak{q}_2(\alpha)\times\mathfrak{h}'}(\nu+i\alpha)_{\nu+k\alpha} = 2\dim C(\nu+k\alpha).$$

Hence it is enough to verify that $\dim N_{\nu+k\alpha}$ is divisible by $2\dim C(\nu+k\alpha)$ that is

$$\mathrm{mult}(L(\lambda); C(\nu+k\alpha)) \equiv 0 \mod 2.$$

Let $\{\nu_i\}_{i=1}^n$ contain i_+ copies of k and i_- copies of $-k$. Then $\{\nu+k\alpha\}_{i=1}^n$ contains $j_+=i_+-2$ copies of k and j_- copies of $-k$. Therefore $i_+-i_-\neq j_+-j_-$, so

$$\mathrm{Core}(\nu+k\alpha)\neq\mathrm{Core}(\nu)=\mathrm{Core}(\lambda)$$

and thus by Theorem 9.1, $\mathrm{mult}(L(\lambda); C(\nu+k\alpha)\equiv 0 \bmod 2$ as required. □

Combining Theorem 7.1, Theorem 9.1, and the results of [11], we obtain our main result:

Theorem 9.2. $\mathrm{sch}_{\mathfrak{h}}$ *defines a morphism of spoiled superalgebras* $\mathcal{K}_-(\mathfrak{g})\to\mathcal{K}_-(\mathfrak{h})^W$. *Further*:

(1) *For* $\lambda\in P^+(\mathfrak{q}_n)\cap I_0$ *one has*

$$\mathrm{sch}_{\mathfrak{h}} L(\lambda) = \sum_{\substack{\nu\in P^+(\mathfrak{g})\cap I_0:\ \nu\sim\lambda\\ \mathrm{zero}\,\lambda\geq\mathrm{zero}\,\nu\\ \mathrm{zero}\,\lambda-\mathrm{zero}\,\nu\equiv 0 \bmod 4}} k_\nu a_\nu,\quad k_\nu\in\mathbb{N}.$$

(2) *For* $\lambda\in P^+(\mathfrak{g})\cap I_1$ *one has*

$$\mathrm{sch}_{\mathfrak{h}} L(\lambda) = \sum_{\substack{\nu\in P^+(\mathfrak{g})\cap I_1:\ \nu\sim\lambda\\ \mathrm{zero}\,\lambda\geq\mathrm{zero}\,\nu}} k_\nu a_\nu,\quad k_\nu\in\{0,1\}.$$

In both cases $k_\lambda=1$, *and* $\mathrm{sch}_{\mathfrak{h}} L(\lambda)=a_\lambda$ *if* $\lambda\in P^+(\mathfrak{g})$ *is typical.*

Proof. The only part that remains to be justified is the inequality $\mathrm{zero}\,\lambda\geq\mathrm{zero}\,\nu$. For this we invoke the results of [11], where it is shown that if $x=H_1\in(\mathfrak{q}_n)_{\bar 1}$, and $\lambda\in P^+(\mathfrak{q}_n)$, then $ds_x^{\mathrm{zero}\,\lambda+1}L(\lambda)=0$. □

Corollary 9.4. *For* $\lambda\in P^+(\mathfrak{g})$ *and* $\nu\notin W P^+(\mathfrak{g})$ *we have* $2\dim E_\nu$ *divides* $\dim L(\lambda)_\nu$.

Corollary 9.5. (1) *The map* $N\mapsto\mathrm{sch}_{\mathfrak{h}} N$ *induces an isomorphism*

$$\mathcal{K}_-(\mathfrak{q}_n)_{\mathbb{Q}} \xrightarrow{\sim} \mathcal{K}_-(\mathcal{C})^W_{\mathbb{Q}}.$$

(2) *For* $A\in\Xi$, *let* $\mathcal{F}(\mathfrak{q}_n)_A$ *be the full subcategory of* $\mathcal{F}(\mathfrak{q}_n)$ *consisting of modules of central character corresponding to* A. *Then* $\mathrm{sch}_{\mathfrak{h}}$ *restricts to an isomorphism* $\mathcal{K}_-(\mathcal{F}(\mathfrak{q}_n)_A)_{\mathbb{Q}}\to(\mathcal{K}_-(\mathfrak{h})_A)^W_{\mathbb{Q}}$.

(3) *The duality is given by* $a_\nu^*=a_{-w_0(\nu)}$

Proof. It is clear that $\mathrm{sch}_{\mathfrak{h}}$ has image lying in $\mathscr{K}_-(\mathcal{C})_{\mathbb{Q}}^W$, and by Corollary 5.2, $\mathrm{sch}_{\mathfrak{h}}$ is an embedding. From Theorem 9.2 and the fact that for each $\lambda \in P^+(\mathfrak{g})$ the number of elements $\nu \in P^+(\mathfrak{g})$ satisfying $\nu < \lambda$ and $\nu \sim \lambda$ is finite, it follows that the image of $\mathscr{K}_-(\mathfrak{g})$ contains a_ν for each $\nu \in P^+(\mathfrak{g})$. Hence the image is equal to $\mathscr{K}_-(\mathfrak{h})_{\mathbb{Q}}^W$, proving surjectivity giving (i). Part (ii) is an easy consequence of (i).

For (iii) we simply apply Theorem 9.1, and for (iii) we observe that $T_{-\lambda} = \sigma(T_\lambda)$ and $t(\lambda) = t_{-\lambda}$. Thus it is easy check that $C(\lambda)^* \cong C(-\lambda)$. □

9.5. *The map* $\mathbf{ds}_s$. For $s = \frac{1}{2}, \dots, \frac{n}{2}$, set

$$x_s = H_{n+1-2s} + \cdots + H_n.$$

Then $c = x_s^2$ is a semisimple element of $\mathfrak{g}_{\bar{0}}$, and we have $DS_x\mathfrak{q}_n = \mathfrak{q}_{n-2s}$. Further we have a splitting $\mathfrak{q}_n^c = \mathfrak{q}_{n-2s} \ltimes [x, \mathfrak{q}_n^c]$, where $\mathfrak{q}_{n-2s} \subseteq \mathfrak{q}_n$ is the natural embedding such that $\mathfrak{t}_x \subseteq \mathfrak{t}$ is spanned by $h_1, \dots, h_{n-2s}$ and $\Delta(\mathfrak{g}_x) = \{\varepsilon_i - \varepsilon_j\}_{1 \le i \neq j \le n-2s}$.

We write $DS_s := DS_{x_s}$ and $ds_s : R(\mathfrak{q}_n) \to R(\mathfrak{q}_{n-2s})$ for the induced homomorphism on reduced Grothendieck rings. These splittings of $\mathfrak{g}_{x_s}$ in $\mathfrak{g}^{x_s}$ satisfy the hypotheses of Lemma 8.2, thus we have $ds_i \circ ds_j = ds_{i+j}$.

9.5.1. *Remark.* For $i = 1, \dots, \lfloor \frac{n}{2} \rfloor$ set

$$\alpha_i := \varepsilon_{n-2s+1} - \varepsilon_{n-2s+2}, \quad \alpha_i^\vee := h_{n-2s+1} - h_{n-2s+2}.$$

Let y_α be a nonzero odd element in $\mathfrak{g}_\alpha$. For each $1 \le s \le \frac{n}{2}$, consider $y_s := \sum_{i=1}^{s} x_{\alpha_i}$. Then clearly $[y_s, y_s] = 0$, and one can show that $ds_{y_s} = ds_{2s}$ as defined above.

We view $\mathfrak{t}_{x_s}^*$ as a subspace in $\mathfrak{t}^*$ via the natural embedding $\iota_{n,s} : \mathfrak{t}_{x_s}^* \hookrightarrow \mathfrak{t}^*$ (given by $\varepsilon_i \mapsto \varepsilon_i \in \mathfrak{t}^*$ for $i = 1, \dots, n-2s$).

Corollary 9.6. *Take $N \in \mathcal{F}(\mathfrak{g})$. For each $\nu \in \mathfrak{t}_{x_s}^*$ we have*

$$\mathrm{smult}(\mathrm{DS}_s(N) : C(\nu)) = \mathrm{smult}(N : C(\iota_{n,s}(\nu))).$$

Proof. Retain the notation from Section 8.4 and set $\mu := \iota_{n,s}(\nu) \in \mathfrak{t}^*$. Note that $\mathfrak{t}_{x_s}$ is spanned by $h_1, h_2, \dots, h_{n-2s}$ and $\mathfrak{t}^{x_s} = \mathfrak{t}$.

Since $\mathfrak{t} \cap [x_s, \mathfrak{g}^c]$ is spanned by $h_{n-s+1}, \dots, h_n$, for $\mu \in I_x$ the restriction of F_μ to $\mathfrak{t}^{x_s}$ written with respect to the above basis has the diagonal entries $\mu_1, \dots, \mu_{n-s}$ and zeros on the last s places. In particular, for $\mu \in I_x$ one has $\mathrm{rank}\, F_\mu^{x_s} = \mathrm{rank}\, F_\mu$ if and only if $\mu_i = 0$ for $i > n-s$, i.e., $\mu \in \mathfrak{t}_x^*$. By Corollary 8.1 we obtain

$$\mathrm{smult}(\mathrm{DS}_s(N); C(\nu)) = (-1)^{i_\mu}\, \mathrm{smult}(N; C(\mu))$$

where $\mathrm{Res}^{\mathfrak{h}}_{\mathfrak{h}_{x_s}} C(\mu) \cong \Pi^{i_\mu} C(\nu))$. The formulae

$$T_\mu = \prod_{i:\ \mu_i \neq 0} H_i = \prod_{i:\ \nu_i \neq 0} H_i = T_\nu, \quad t(\mu) = t(\nu)$$

give $\mathrm{Res}^{\mathfrak{h}}_{\mathfrak{h}_{x_s}} C(\mu) \cong C(\nu)$ as required. ☐

Corollary 9.7. *The map* $\mathrm{ds}_s : \mathscr{K}_-(\mathfrak{g}) \to \mathscr{K}_-(\mathfrak{g}_{x_s})$ *is given by*

$$\mathrm{ds}_s(a_\mu) = \begin{cases} 0 & \text{if } \mathrm{zero}\,\mu < s, \\ a_{\mu'} & \text{if } \mathrm{zero}\,\mu \geq s, \end{cases}$$

where $\mu' \in P^+(\mathfrak{g}_x)$ *is such that* $\mathrm{Nonzero}(\mu) = \mathrm{Nonzero}(\mu')$.

Proof. Write $\mathrm{sch}_{\mathfrak{h}} N = \sum_{\mu \in P^+(\mathfrak{g})} k_\mu a_\mu$ and $\mathrm{sch}_{\mathfrak{h}_x} \mathrm{DS}_s(N) = \sum_{\nu \in P^+(\mathfrak{g}_x)} m_\nu a_\nu$. By Corollary 9.6 $k_\mu = m_\nu$ if $\mu = \iota_{n,s}(\nu)$. ☐

9.5.2. We denote by $\mathcal{F}(\mathfrak{g})_{\mathrm{int}}$ the full subcategory of $\mathcal{F}(\mathfrak{g})$ with the modules whose weights with nonzero weight spaces lie in the lattice generated by $\epsilon_1, \dots, \epsilon_n$.

Corollary 9.8. *The kernel of* $\mathrm{ds}_s : \mathscr{K}_-(\mathfrak{g}) \to \mathscr{K}_-(\mathfrak{g}_x)$ *is spanned by* a_μ *with* $\mathrm{zero}\,\mu < s$ *and the image of* ds_s *is equal to* $\mathscr{K}_-(\mathcal{F}(\mathfrak{g}_{x_s})_{\mathrm{int}})$.

Corollary 9.9. *For* $\lambda \in P^+(\mathfrak{g})$ *and* $\nu \in P^+(\mathfrak{g}_{x_s})$ *one has*

$$\mathrm{zero}\,\lambda - \mathrm{zero}\,\nu - s \not\equiv 0 \bmod 4 \implies \mathrm{smult}(\mathrm{DS}_s(L(\lambda)) : L_{\mathfrak{g}_x}(\nu)) = 0.$$

Proof. Combining Theorem 9.2 and Corollary 9.7 we conclude that $\mathrm{ds}_s(\mathrm{sch}\, L(\lambda))$ lies in the span of a_ν with $\nu \in P^+(\mathfrak{g}_x)$ such that $n - zero(\lambda) - (n - s - \mathrm{zero}(\nu)) \equiv 0 \bmod 4$ that is $\mathrm{zero}\,\lambda - \mathrm{zero}\,\nu - s \equiv 0 \bmod 4$.

Let ν_0 be maximal (with respect to the standard partial order in $\mathfrak{t}_x^*$) such that

$$\mathrm{smult}(\mathrm{DS}_s(L(\lambda)); L_{\mathfrak{g}_x}(\nu)) \neq 0 \quad \text{and} \quad \mathrm{zero}\,\lambda - \mathrm{zero}\,\nu - s \not\equiv 0 \bmod 4.$$

The maximality of ν_0 forces $\mathrm{zero}\,\nu \not\equiv \mathrm{zero}\,\nu_0$ if $\nu > \nu_0$ and $L_{\mathfrak{g}_x}(\nu)$ is a subquotient of $\mathrm{DS}_s(L_{\mathfrak{g}}(\lambda))$. By Theorem 9.2 we obtain $[L_{\mathfrak{g}_x}(\nu) : C(\nu_0)] = 0$ if $\nu \neq \nu_0$ and $L_{\mathfrak{g}_x}(\nu)$ is a subquotient of $\mathrm{DS}_s(L_{\mathfrak{g}}(\lambda))$. Therefore $\mathrm{smult}(\mathrm{DS}_s(L(\lambda)); L_{\mathfrak{g}_x}(\nu_0))$ is equal to the coefficient of a_{ν_0} in $\mathrm{sch}_{\mathfrak{h}} \mathrm{DS}_s(L(\lambda)) = \mathrm{ds}_s(\mathrm{sch}_{\mathfrak{h}} L(\lambda))$, which is zero by above. ☐

9.6. *Example:* $\mathfrak{q}_2$. Recall that the atypical dominant weights for $\mathfrak{q}_2$ are of the form $s(\varepsilon_1 - \varepsilon_2)$ for $s \in \frac{1}{2}\mathbb{N}$.

Proposition 9.3. *Take* $\mathfrak{g} = \mathfrak{q}_2$ *with* $\Delta^+ = \{\alpha\}$. *For* $\lambda \in P^+(\mathfrak{q}_2)$ *one has*

$$\mathrm{sch}_{\mathfrak{h}} L(\lambda) = \begin{cases} a_\lambda & \text{if } \lambda \text{ is typical}, \\ \sum_{i=1}^{s} a_{i\alpha} & \text{if } \lambda = s\alpha,\ s \in \mathbb{N}, \\ \sum_{i=0}^{s} a_{i+\frac{1}{2}\alpha} & \text{if } \lambda = \left(s + \frac{1}{2}\right)\alpha,\ s \in \mathbb{N}. \end{cases}$$

Proof. If λ is typical, the assertion follows from Theorem 6.1. Consider the case when $\mu \in P^+(\mathfrak{q}_2)$ is atypical and $\mu \neq 0$. Write $K(\mu)$ for the maximal finite-dimensional quotient of $\mathrm{Ind}^{\mathfrak{q}_2}_{\mathfrak{b}} C(\mu)$. It is known that $\mathrm{sch}_{\mathfrak{h}} K(\mu) = a_\mu$. Further, if $\mu = s\alpha$ for $s \in \mathbb{N}$, then it is known (for example see Section 7 of [9]) that we have short exact sequences

$$0 \to V_0 \to K(\alpha) \to L(\alpha) \to 0, \quad 0 \to \Pi L((s-1)\alpha) \to K(s\alpha) \to L(s\alpha) \to 0$$

where V_0 is a nontrivial extension of $\mathbb{C}$ and $\Pi\mathbb{C}$. On the other hand if $\mu = (s+\frac{1}{2})\alpha$ then for $s=0$ we have $K(\alpha/2) = L(\alpha/2)$ and for $s>0$ we have short exact sequences

$$0 \to \Pi L\big((s-\tfrac{1}{2})\alpha\big) \to K\big((s+\tfrac{1}{2})\alpha\big) \to L\big((s+\tfrac{1}{2})\alpha\big) \to 0.$$

From these results we can obtain the desired formulas by induction. □

9.7. ***Realization of $\mathscr{K}_-(\mathfrak{q}_n)$.*** Recall that $\mathcal{F}(\mathfrak{g})_{\text{int}}$ denotes the full subcategory of $\mathcal{F}(\mathfrak{g})$ consisting of the modules with integral weights and that $\mathcal{C}$ denotes the full subcategory of $\mathcal{F}(\mathfrak{h})$ consists of modules with weights lying in $P(\mathfrak{g})$. We write $\mathcal{C}_{\text{int}}$ for the subcategory of $\mathcal{C}$ consisting of those modules with integral weights.

9.7.1. By Theorem 9.2, the map sch gives embeddings

$$\mathscr{K}_-(\mathfrak{g}) \hookrightarrow \mathscr{K}_-(\mathcal{C})^W, \quad \mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\text{int}}) \hookrightarrow \mathscr{K}_-(\mathcal{C}_{\text{int}})^W.$$

Further, we have identified the image of $\mathscr{K}_-(\mathfrak{g})$ (resp., $\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\text{int}})$) with the subalgebra spanned by a_ν with $\nu \in P^+(\mathfrak{g})$ (resp., $\nu \in P^+(\mathfrak{g})_{\text{int}}$).

One has $a_0 = 1$; we denote by $\text{ev}_0 : \mathscr{K}_-(\mathcal{C}) \to \mathbb{Z}$ the counit map given by $\text{ev}_0(a_\nu) = \delta_{0,\nu}$ for $\nu \in P^+(\mathfrak{g})$. One has

$$\mathscr{K}_-(\mathcal{C}) = \mathscr{K}_-(\mathcal{C}_{\text{int}})^W \oplus \mathscr{K}_-(\mathcal{C}_{nint})^W,$$

where $\mathcal{C}_{nint}$ consists of modules with weights that lie in $P(\mathfrak{g})$ and are nonintegral. For $b \in \mathscr{K}_-(\mathcal{C})^W$ and $b' \in \mathscr{K}_-(\mathcal{C}_{nint})^W$ one has $bb' = \text{ev}_0(b)b'$, so $\mathscr{K}_-(\mathcal{C}_{nint})$ is an ideal of $\mathscr{K}_-(\mathcal{C})$. By Corollary 9.8, for $x \neq 0$, $\text{ds}_x(\mathscr{K}_-(\mathcal{C}_{nint})^W) = 0$ and the image of ds_x lies in $\mathscr{K}_-(\mathcal{C}_{\text{int}})$. This reduces a study of $\mathscr{K}_-(\mathcal{C})$ to a study of $\mathscr{K}_-(\mathcal{C}_{\text{int}})$.

9.7.2. The ring $\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\text{int}}) \otimes_{\mathbb{Z}} \mathbb{Z}[\sqrt{-1}]$ can be realized in the following way.

Let V be a free $\mathbb{Z}[\sqrt{-1}]$-module with a basis $\{v_i \mid i \in \mathbb{Z}\setminus\{0\}\}$. Denote by $\bigwedge V$ the external ring of V. This is a $\mathbb{N}$-graded supercommutative ring; we consider $(\bigwedge V)^{\text{spoil}}$, which has

$$\big(\textstyle\bigwedge V\big)^{\text{spoil}} = \bigoplus_{i=0}^{\infty} \textstyle\bigwedge^{2i} V \oplus \bigoplus_{i=0}^{\infty} \textstyle\bigwedge^{2i+1} V\varepsilon$$

where, we recall, ε is a formal variable satisfying $\varepsilon^2 = 2\varepsilon = 0$. This is an $\mathbb{N}$-graded commutative and supercommutative ring (which means that $(\bigwedge V)_i^{\text{spoil}}(\bigwedge V)_j^{\text{spoil}} = 0$ if i, j are odd).

We denote by Ξ_{int} the set of finite multisets $\{a_i\}_{i=1}^s$ with $a_i \in \mathbb{Z}\setminus\{0\}$ and $a_i + a_j \neq 0$ for all $1 \le i, j \le s$. For $A \in \Xi_{\text{int}}$ we denote by $(\bigwedge V)_A^{\text{spoil}}$ the span of the elements $v_{i_1} \wedge v_{i_2} \wedge \ldots \wedge v_{i_p}$ (resp., the elements $\varepsilon v_{i_1} \wedge v_{i_2} \wedge \ldots \wedge v_{i_p}$) with $\text{Core}(\{i_1, \ldots, i_p\}) = A$ if A has an even (resp., an odd) cardinality. Clearly, this gives a grading

$$\big(\textstyle\bigwedge V\big)_A^{\text{spoil}} \big(\textstyle\bigwedge V\big)_B^{\text{spoil}} \subseteq \big(\textstyle\bigwedge V\big)_{A\diamond B}^{\text{spoil}}$$

and $(\bigwedge V)^{\text{spoil}}_{\varnothing}$ is the subring of $(\bigwedge V)^{\text{spoil}}$ generated by $v_i \wedge v_{-i}$. In other words $\left(\bigwedge V\right)^{\text{spoil}}$ is a Ξ_{int}-graded algebra. For each k we consider the ideal

$$J_k := \sum_{i=k+1}^{\infty} (\bigwedge V)^{\text{spoil}}_i.$$

Recall that the map $\lambda \mapsto \text{Nonzero}(\lambda)$ gives a one-to-one correspondence between $P^+(\mathfrak{g})_{\text{int}}$ and the subsets $S \subset \mathbb{Z}\setminus\{0\}$ of cardinality at most n. The weights in I_0 correspond to the subsets of even cardinality; the zero weight corresponds to $\varnothing$. We define

$$\psi : \mathscr{K}_{-}(\mathcal{F}(\mathfrak{g})_{\text{int}}) \otimes_{\mathbb{Z}} \mathbb{Z}[\sqrt{-1}] \xrightarrow{\sim} (\bigwedge V)^{\text{spoil}}/J_n$$

by setting, for $\lambda \in I_i \cap P^+(\mathfrak{g})$,

$$\psi(a_\lambda) := \varepsilon^i \frac{t(\lambda)}{|t(\lambda)|} \cdot v_\lambda,$$

where $v_\lambda := v_{i_1} \wedge v_{i_2} \wedge \ldots \wedge v_{i_k}$, and $\text{Nonzero}(\lambda) = \{i_1 > i_2 > \ldots > i_k\}$.

Combining Corollary 9.5, Proposition 9.1 and Corollary 9.7 we obtain:

Theorem 9.3. (1) *The map $\psi : \mathscr{K}_{-}(\mathcal{F}(\mathfrak{g})_{\text{int}}) \otimes_{\mathbb{Z}} \mathbb{Z}[\sqrt{-1}] \xrightarrow{\sim} \left(\bigwedge V\right)^{\text{spoil}}/J_n$ is an isomorphism of Ξ_{int}-graded spoiled super rings.*

(2) *In particular, for $A \in \Xi_{\text{int}}$, ψ restricts to an isomorphism of vector spaces $\psi_A : \mathscr{K}_{-}(\mathcal{F}(\mathfrak{g})_A) \to \left(\bigwedge V\right)^{\text{spoil}}_A/J_n$.*

(3) *The map $\text{ds}_s : R(\mathfrak{q}_n)_{\text{int}} \to R(\mathfrak{q}_{n-s})_{\text{int}}$ corresponds to the natural quotient map*

$$\left(\bigwedge V\right)^{\text{spoil}}/J_n \to \left(\bigwedge V\right)^{\text{spoil}}/J_{n-s}$$

(one has $J_{n-s} \supset J_n$ for $s > 0$).

Proof. It is clear that ψ is an isomorphism of $\mathbb{Z}$-modules which preserves the Ξ-grading. The result on ds_s is also clear from the map. Thus it remains to check that ψ respects multiplication.

Let $\lambda, \mu \in P^+(\mathfrak{g})$ with $\lambda, \mu \in I_0$; the case when either is in I_1 is much easier. If $a_\lambda a_\mu = 0$ then the result is clear, thus we assume that $a_\lambda a_\mu \neq 0$. In this case, there exists unique elements $\lambda' \in W\lambda$ and $\mu' \in W\mu$ such that $\lambda' + \mu'$ is dominant, and thus we have $a_\lambda a_\mu = (-1)^i a_{\lambda'+\mu'}$, where $E_{\lambda'} \otimes E_{\mu'} = \Pi^i E_{\lambda'+\mu'}$. We may write $T_{\lambda'+\mu'} = (-1)^j T_{\lambda'} T_{\mu'}$ for some j, and it is clear that this j also satisfies $v_{\lambda'+\mu'} = (-1)^j v_\lambda v_\mu$. Then as in the proof of Proposition 9.1, we have $E_{\lambda'} \otimes E_{\mu'} = \Pi^{j+\ell} E_{\lambda'+\mu'}$, where

$$(-1)^\ell = \frac{t(\lambda)t(\mu)}{t(\lambda'+\mu')} \in \{\pm 1\}.$$

Therefore $a_\lambda a_\mu = (-1)^j \frac{t(\lambda)t(\mu)}{t(\lambda'+\mu')} a_{\lambda'+\mu'}$. Therefore

$$\psi(a_\lambda a_\mu) = (-1)^j \frac{t(\lambda)t(\mu)}{t(\lambda'+\mu'} \left(\frac{t(\lambda'+\mu')}{|t(\lambda'+\mu')|} v_{\lambda'+\mu'} \right) = \left(\frac{t(\lambda)}{|t(\lambda)|} v_\lambda \right) \left(\frac{t(\mu)}{|t(\mu)|} v_\mu \right)$$

where we have used that $|t(\lambda'+\mu')| = |t(\lambda)||t(\mu)|$. $\square$

For $A \in \Xi_{\mathrm{int}}$, we define

$$\bar{A} := \begin{cases} \dfrac{\#\{i : a_i > 0\} - \#\{i | a_i < 0\}}{2} \mod 2 & \text{for } |A| \text{ even,} \\ \dfrac{\#\{i : a_i > 0\} - \#\{i | a_i < 0\} - 1}{2} \mod 2 & \text{for } |A| \text{ odd.} \end{cases}$$

Note that we have $\overline{A \diamond B} = \bar{A} + \bar{B}$ if $|A|$ or $|B|$ is even.

Lemma 9.3. *The grading*

$$\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\mathrm{int}})_{\bar{0}} = \bigoplus_{\bar{A}=\bar{0}} \mathscr{K}_-(\mathcal{F}(\mathfrak{g})_A), \qquad \mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\mathrm{int}})_{\bar{1}} = \bigoplus_{\bar{A}=\bar{1}} \mathscr{K}_-(\mathcal{F}(\mathfrak{g})_A)$$

defines another super ring structure on $\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\mathrm{int}})$; *note that it does not have the structure of a spoiled super ring under this grading.*

Proof. This follows from the fact that $\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_{\mathrm{int}})$ is Ξ_{int}-graded, and $\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_A)\mathscr{K}_-(\mathcal{F}(\mathfrak{g})_B) = 0$ if $|A|$ and $|B|$ are odd. □

Corollary 9.10. *If $|A|$ is even, then $t(\lambda) \in \mathbb{R}^+$, and thus the isomorphism ψ_A admits an integral structure. In particular we have an isomorphism of graded rings*:

$$\bigoplus_{\bar{A}=\bar{0}} \mathscr{K}_-(\mathcal{F}(\mathfrak{g})_A) \to \bigoplus_{\bar{A}=\bar{0}} \left(\bigwedge V^{\mathbb{Z}}\right)_A^{\mathrm{spoil}}.$$

Here $V^{\mathbb{Z}}$ denotes the free $\mathbb{Z}$-module with basis $\{v_i : i \in \mathbb{Z} \setminus \{0\}\}$.

Acknowledgements

The authors are grateful to I. Entova-Aizenbud, D. Grantcharov, and V. Hinich for numerous helpful discussions. Maria Gorelik was supported in part by ISF Grant 1957/21 and the Minerva foundation with funding from the Federal German Ministry for Education and Research. Vera Serganova was supported in part by NSF grant 2001191. Alexander Sherman was supported in part by ISF Grant 711/18, NSF-BSF Grant 2019694, and ARC grant DP210100251.

References

[1] A. Aramova, L. L. Avramov, and J. Herzog, "Resolutions of monomial ideals and cohomology over exterior algebras", *Trans. Amer. Math. Soc.* **352**:2 (2000), 579–594. MR Zbl

[2] J. N. Bernstein, I. M. Gelfand, and S. I. Gelfand, "Category of $\mathfrak{g}$-modules", *Funkcional. Anal. i Priložen.* **10**:2 (1976), 1–8. In Russian; translated in *Funct. Anal. Appl.* **10**:2 (1976), 87–92. MR

[3] S.-J. Cheng, "Supercharacters of queer Lie superalgebras", *J. Math. Phys.* **58**:6 (2017), art. id. 061701. MR Zbl

[4] V. V. Deodhar, O. Gabber, and V. Kac, "Structure of some categories of representations of infinite-dimensional Lie algebras", *Adv. Math.* **45**:1 (1982), 92–116. MR Zbl

[5] M. Duflo and V. Serganova, "On associated variety for Lie superalgebras", preprint, 2005. arXiv math/0507198

[6] I. Entova-Aizenbud and V. Serganova, "Jacobson–Morozov lemma for algebraic supergroups", *Adv. Math.* **398** (2022), art. id. 108240. MR Zbl

[7] A. Frisk, "Typical blocks of the category $\mathscr{O}$ for the queer Lie superalgebra", *J. Algebra Appl.* **6**:5 (2007), 731–778. MR Zbl

[8] M. Gorelik, "On the ghost centre of Lie superalgebras", *Ann. Inst. Fourier* (*Grenoble*) **50**:6 (2000), 1745–1764. MR Zbl

[9] M. Gorelik, "Shapovalov determinants of Q-type Lie superalgebras", *Int. Math. Res. Pap.* **2006** (2006), art. id. 96895. MR Zbl

[10] M. Gorelik, "Depths and cores in the light of DS-functors", preprint, 2020. arXiv 2010.05721

[11] M. Gorelik and A. Sherman, "On the Duflo–Serganova functor for the queer Lie superalgebra", *Adv. Math.* **441** (2024), art. id. 109552. MR Zbl

[12] N. Grantcharov and V. Serganova, "Extension quiver for Lie superalgebra $\mathfrak{q}(3)$", *SIGMA Symmetry Integrability Geom. Methods Appl.* **16** (2020), art. id. 141. MR Zbl

[13] C. Hoyt and S. Reif, "Grothendieck rings for Lie superalgebras and the Duflo–Serganova functor", *Algebra Number Theory* **12**:9 (2018), 2167–2184. MR Zbl

[14] J. E. Humphreys, *Introduction to Lie algebras and representation theory*, Grad. Texts in Math. **9**, Springer, 1972. MR Zbl

[15] N. I. Ivanova and A. L. Onishchik, "Parabolic subalgebras and gradings of reductive Lie superalgebras", *Sovrem. Mat. Fundam. Napravl.* **20** (2006), 5–68. In Russian; translated in *J. Math. Sci.* (*NY*) **152**:1 (2008), 1–60. MR Zbl

[16] V. G. Kac, "Lie superalgebras", *Adv. Math.* **26**:1 (1977), 8–96. MR Zbl

[17] V. G. Kac and D. A. Kazhdan, "Structure of representations with highest weight of infinite-dimensional Lie algebras", *Adv. Math.* **34**:1 (1979), 97–108. MR Zbl

[18] V. Mazorchuk, "Parabolic category $\mathscr{O}$ for classical Lie superalgebras", pp. 149–166 in *Advances in Lie superalgebras* (Rome, 2012), edited by M. Gorelik and P. Papi, Springer INdAM Ser. **7**, Springer, 2014. MR Zbl

[19] I. B. Penkov, "Characters of typical irreducible finite-dimensional $\mathfrak{q}(n)$-modules", *Funktsional. Anal. i Prilozhen.* **20**:1 (1986), 37–45. In Russian; translated in *Funct. Anal. Appl.* **20**:1 (1986), 30–37. MR Zbl

[20] I. Penkov and V. Serganova, "Generic irreducible representations of finite-dimensional Lie superalgebras", *Internat. J. Math.* **5**:3 (1994), 389–419. MR Zbl

[21] D. H. Peterson and V. G. Kac, "Infinite flag varieties and conjugacy theorems", *Proc. Nat. Acad. Sci. U.S.A.* **80**:6 (1983), 1778–1782. MR Zbl

[22] V. Serganova, "Quasireductive supergroups", pp. 141–159 in *New developments in Lie theory and its applications* (Córdoba, Argentina, 2009), edited by C. Boyallian et al., Contemp. Math. **544**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[23] V. Serganova, "Finite dimensional representations of algebraic supergroups", pp. 603–632 in *Proceedings of the International Congress of Mathematicians*, *I* (Seoul, 2014), edited by S. Y. Jang et al., Kyung Moon Sa, Seoul, 2014. MR Zbl

[24] A. N. Sergeev, "The centre of enveloping algebra for Lie superalgebra $Q(n, \mathbb{C})$", *Lett. Math. Phys.* **7**:3 (1983), 177–179. MR Zbl

[25] A. N. Sergeev and A. P. Veselov, "Grothendieck rings of basic classical Lie superalgebras", *Ann. of Math.* (2) **173**:2 (2011), 663–703. MR Zbl

Communicated by Sergey Fomin
Received 2023-09-07 Revised 2024-11-08 Accepted 2024-12-23

maria.gorelik@gmail.com — *Department of Mathematics, Weizmann Institute of science, 76100 Rehovot, Israel*

serganov@math.berkeley.edu — *Department of Mathematics, University of California, Berkeley, CA 94720-3840, United States*

xandersherm@gmail.com — *School of Mathematics and Statistics, University of Sydney, Camperdown NSW 2050, Australia*

ALGEBRA AND NUMBER THEORY 19:12 (2025)
https://doi.org/10.2140/ant.2025.19.2409

Combing a hedgehog over a field

Alexey Ananyevskiy and Marc Levine

We investigate the question of the existence of a nonvanishing section of the tangent bundle on a smooth affine quadric hypersurface Q^o over a given perfect field k. If Q^o admits a k-rational point, we give a number of necessary and sufficient conditions for such existence. We apply these conditions in a number of examples, including the case of the *algebraic n-sphere* over k, $S^n_k \subset \mathbb{A}^{n+1}_k$, defined by the equation $\sum_{i=1}^{n+1} x_i^2 = 1$.

1. Introduction

It is an elementary but nonetheless beautiful result found in nearly all introductory courses in differential topology that, for all $n \geq 1$, the tangent bundle $T_{S^{2n}}$ does not admit a nonvanishing section. One proof uses the Gauss–Bonnet theorem to show that Euler class of $T_{S^{2n}}$ is nonzero by computing its degree as the Euler characteristic of S^{2n}, namely 2, while the existence of a nonvanishing section would force the Euler class to vanish. For the odd-dimensional case, the Euler characteristic vanishes, and hence the Euler class vanishes as well; one can also easily write down explicitly a nonvanishing section of $T_{S^{2n+1}}$.

Writing the n-sphere S^n as the hypersurface in $\mathbb{R}^{n+1}$ defined by the equation $\sum_{i=1}^{n+1} x_i^2 = 1$, one can ask the corresponding question in the algebro-geometric setting: let k be a field of characteristic $\neq 2$ and let $S^n_k \subseteq \mathbb{A}^{n+1}_k$ be the hypersurface defined by the equation $\sum_{i=1}^{n+1} x_i^2 = 1$. Does the tangent bundle $T_{S^n_k}$ admit a nonvanishing section? (To avoid any possible misunderstanding, for $E \to X$ a vector bundle on a k-variety X, a section $s : X \to E$ is said to be nonvanishing if the scheme-theoretic intersection of $s(X)$ with the zero-section of E is the empty scheme. Equivalently, letting $\bar{k}$ denote the algebraic closure of k, the set of $\bar{k}$ points x of X with $s(x) = 0$ is empty.)

This question for $S^2_{\mathbb{Q}_p}$ was originally raised by Umberto Zannier (see Remark 1.7 below for his original formulation). He showed that $S^2_{\mathbb{Q}_p}$ admits a nonvanishing vector field for odd p and he asked if there is a nonvanishing vector field on the 2-sphere over $\mathbb{Q}_2$, motivating our interest in the question of the existence of nonvanishing vector fields on S^n_k for arbitrary n and k.

We give an essentially complete answer to this question; if k is perfect, this is in fact a special case of the more general Theorem 1.4 about smooth affine quadric hypersurfaces.

This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 832833).
MSC2020: 11E81, 14J60.
Keywords: splitting vector bundles, Euler class, Chow–Witt ring, Grothendieck–Witt group.

Theorem 1.1 (see Theorem 4.1 (1), (3) and Remark 1.10). *Let k be a field of characteristic $\neq 2$.*

(1) *If n is odd, then $T_{S^n_k}$ admits a nonvanishing section.*

(2) *If $n > 0$ is even, then $T_{S^n_k}$ admits a nonvanishing section if and only if -1 is in the subgroup of $k^\times$ generated by the nonzero values of the function $\sum_{i=1}^{n+1} x_i^2$ on k^{n+1}.*

As the condition in (2) is not very explicit, we reformulate this as follows.

Corollary 1.2 (Corollary 4.3). *Let k be a field of characteristic $\neq 2$. For $n > 0$ even, $T_{S^n_k}$ admits a nonvanishing section if and only if the equation*

$$\sum_{i=1}^{2n+1} x_i^2 = -1$$

has a solution with the $x_i \in k$.

Examples 1.3. Let S^n_k be as above.

(1) Suppose that $\mathrm{char}\, k = p > 2$. Then $T_{S^n_k}$ has a nonvanishing section for all $n > 0$.

(2) Suppose k contains a p-adic field $\mathbb{Q}_p$. Then $T_{S^n_k}$ has a nonvanishing section for all $n > 0$.

(3) Suppose that k is a number field, and take $n > 0$ to be even. Then $T_{S^n_k}$ has a nonvanishing section if and only if k has no real embeddings.

To see this we apply Corollary 1.2. For (1), since $\mathbb{F}_p \subseteq k$, it suffices to take $k = \mathbb{F}_p$. Since every element $x \in \mathbb{F}_p^\times$ is a sum of two squares [Lam 2005, Proposition II.3.4], the condition of Corollary 1.2 is satisfied for all $n \geq 1$. See also Remark 1.10 for an explicit nonvanishing section.

For (2), we reduce as above to the case $k = \mathbb{Q}_p$. If p is odd, then by Hensel's lemma, each solution to $\sum_{i=1}^{2n+1} x_i^2 = -1$ in $\mathbb{F}_p$ lifts to a solution in $\mathbb{Z}_p$, so the criterion is satisfied. For $p = 2$, the class of a unit u in $\mathbb{Z}_2$ modulo squares is given by the image of u in $(\mathbb{Z}/8)^\times$, so it suffices to write 7 as the sum of ≤ 5 squares in $\mathbb{Z}_2$ and it turns out that four squares are enough: $7 = 1 + 1 + 1 + 4$. For those more intrinsically minded, one has the general result that *every* nondegenerate quadratic form ϕ in at least five variables over a local field has a nontrivial zero [Lam 2005, Theorem VI.2.12], which we apply to $\phi = \sum_{i=1}^5 x_i^2$.

For (3), it is clear that the equation $\sum_{i=1}^{2n+1} x_i^2 = -1$ has no solution in k^{2n+1} if k admits a real embedding. Conversely, we may use the Hasse–Minkowski principle for quadratic forms (see, e.g., [Lam 2005, Hasse–Minkowski principle VI.3.1]) to see that $\sum_{i=1}^4 x_i^2 = -1$ has a solution in k if k is a purely imaginary number field. Indeed, it suffices to show that $\sum_{i=1}^4 x_i^2 = -1$ has a solution in k_v for every place v of k. This is clear if v is an infinite place, as $k_v = \mathbb{C}$ by assumption. If v is a finite place, then $k_v \supset \mathbb{Q}_p$ for some prime p, and we have just seen that $\sum_{i=1}^4 x_i^2 = -1$ has a solution in $\mathbb{Q}_p$ for every prime p.

One can also ask about a general smooth affine quadric $Q^o \subseteq \mathbb{A}_k^{n+1}$, with k a field of characteristic $\neq 2$. Since every quadratic form over k can be diagonalized, we may assume that Q^o is defined by an equation of the form $q = 1$, where $q = \sum_{i=1}^{n+1} a_i x_i^2 \in k[x_1, \ldots, x_{n+1}]$, with $\prod_i a_i \neq 0$. Here one has a result of essentially the same form as for S^n_k, with the extra condition that, for even n, $Q^o(k)$ should be nonempty, that is, $q = 1$ has a solution in k.

Let $D(q)$ be the set of nonzero values of q on k^{n+1}, let $D(q)^2 = \{a \cdot b \mid a, b \in D(q)\} \subseteq k^\times$, and let $[D(q)]$, $[D(q)^2]$ be the subgroups of $k^\times$ generated by $D(q)$, $D(q)^2$, respectively.

Theorem 1.4 (Theorem 4.1 (1), (3)). *Let k be a perfect field of characteristic $\neq 2$, let $q = \sum_{i=1}^{n+1} a_i x_i^2$ with $a_1, \dots, a_{n+1} \in k^\times$ and let $Q^o \subseteq \mathbb{A}_k^{n+1}$ be the affine quadric hypersurface $q = 1$.*

(1) *If n is odd, then T_{Q^o} has a nonvanishing section.*

(2) *Suppose $Q^o(k) \neq \varnothing$. If $n > 0$ is even, then T_{Q^o} has a nonvanishing section if and only if $-1 \in [D(q)]$.*

If $Q^o(k) = \varnothing$ and n is even, we only have a necessary condition for the existence of a nonvanishing section of T_{Q^o}.

Theorem 1.5 (Theorem 4.1 (2)). *Let k, q and Q^o be as above. If n is even and T_{Q^o} has a nonvanishing section, then $-\prod_{i=1}^{n+1} a_i \in [D(q)^2]$.*

Since $a_i \in D(q)$ for each i, the above condition is the same as asking for $-a_i$ to be in $[D(q)^2]$ for *some* i. Note that $Q^o(k) \neq \varnothing$ if and only if $1 \in D(q)$, so if $Q^o(k) \neq \varnothing$, we have $[D(q)^2] = [D(q)]$, and $-1 \in [D(q)]$ if and only if $-\prod_{i=1}^{n+1} a_i \in [D(q)]$.

Here is a version of Examples 1.3 for general q.

Corollary 1.6 (Corollaries 4.5, 4.7, and 4.9). *Let k, q and Q^o be as in Theorem 1.4.*

(1) *Let $k = \mathbb{F}_{p^m}$ with $p > 2$. Then T_{Q^o} has a nonvanishing section for all $n > 0$.*

(2) *Suppose k is a non-Archimedean local field of characteristic zero, the perfection of a local field of characteristic $p > 2$, or the perfection of a function field of a curve over a finite field of characteristic $p > 2$. Then for n odd, or $n \geq 4$ even, T_{Q^o} has a nonvanishing section. If $n = 2$, then T_{Q^o} has a nonvanishing section if $Q^o(k) \neq \varnothing$.*

(3) *Suppose k is a number field, $Q^o(k) \neq \varnothing$ and $n > 0$ is even. Then T_{Q^o} has a nonvanishing section if and only if the equation $q = 0$ has a nontrivial solution in k_v for every real place v of k. Equivalently, for each real embedding $\sigma : k \hookrightarrow \mathbb{R}$, $\sigma(a_i) < 0$ for some i. T_{Q^o} also has a nonvanishing section if n is odd.*

(4) *Let k be a perfect field of cohomological dimension ≤ 2. Suppose that n is odd, or that $n > 0$ is even and $Q^o(k) \neq \varnothing$. Then T_{Q^o} has a nonvanishing section.*

Remark 1.7 (unimodular rows and unimodular matrices). Let $q = \sum_{i=1}^{n+1} a_i x_i^2 \in k[x_1, \dots, x_{n+1}]$ be a quadratic form, defining $Q^o \subseteq \mathbb{A}_k^{n+1}$ as $V(q-1)$, and let R be the coordinate ring

$$R := k[x_1, \dots, x_{n+1}]/(q-1).$$

We are assuming that q is nondegenerate, that is, $\prod_i a_i \neq 0$, and that $n \geq 1$.

Let $\nabla(q)$ denote the gradient

$$\nabla(q) := (\partial q/\partial x_1, \dots, \partial q/\partial x_{n+1})$$

and let

$$\tilde{\nabla}(q) := (a_1 x_1, \dots, a_{n+1} x_{n+1}),$$

so $2\tilde{\nabla}(q) = \nabla(q)$.

We first assume that 2 is invertible in k, so we can rewrite the tangent-normal sequence for $Q^o \subseteq \mathbb{A}_k^{n+1}$ using $\tilde{\nabla}(q)$, as

$$0 \to T_{Q^o} \to \mathcal{O}_{Q^o}^{n+1} \xrightarrow{\tilde{\nabla}(q)^t} \mathcal{O}_{Q^o} \to 0, \tag{1.8}$$

where $\tilde{\nabla}(q)^t \in M_{n+1\times 1}(R)$ is the transpose of $\tilde{\nabla}(q)$. Since Q^o is affine, we can rephrase everything in terms of R-modules, giving the exact sequence

$$0 \to \mathfrak{T}_{Q^o} \to R^{n+1} \xrightarrow{\tilde{\nabla}(q)^t} R \to 0,$$

with $\mathfrak{T}_{Q^o}$ the R-module of global sections of T_{Q^o}. Since

$$(x_1, \dots, x_{n+1}) \cdot \tilde{\nabla}(q) = 1,$$

we may split the surjection by $(x_1, \dots, x_{n+1}) : R \to R^{n+1}$, exhibiting $\mathfrak{T}_{Q^o}$ as a stably free R-module, and showing that $(x_1, \dots, x_{n+1})$ is a *unimodular row*, i.e., $(x_1, \dots, x_{n+1})R$ is the unit ideal.

It is straightforward to see that the stably free R-module $\mathfrak{T}_{Q^o}$ is free if and only if there is a matrix $M \in \mathrm{GL}_{n+1}(R)$ with the first row $(x_1, \dots, x_{n+1})$; by dividing the last row of M by $\det M$, we may in fact take M to have $\det M = 1$, so $\mathfrak{T}_{Q^o}$ is a free R-module if and only if there is a unimodular matrix M over R with the first row $(x_1, \dots, x_{n+1})$.

More generally, we may take k to be an arbitrary commutative ring (even with 2 not a unit), and let $(a_{ij})_{1\le i,j\le n+1} \in \mathrm{GL}_{n+1}(k)$ be an invertible symmetric matrix. Let

$$q := \sum_{i,j=1}^{n+1} a_{ij}x_ix_j, \quad R := k[x_1, \dots, x_{n+1}]/(q-1), \quad Q^o := \operatorname{Spec} R,$$
$$\tilde{\nabla}(q) := \Big(\sum_j a_{1j}x_j, \sum_j a_{2j}x_j, \dots, \sum_j a_{n+1j}x_j\Big).$$

Then the map $R^{n+1} \xrightarrow{\tilde{\nabla}(q)^t} R$ is surjective, and we may define a stably free R-module $\mathfrak{T}_{Q^o/k}$ by the exact sequence

$$0 \to \mathfrak{T}_{Q^o/k} \to R^{n+1} \xrightarrow{\tilde{\nabla}(q)^t} R \to 0. \tag{1.9}$$

Since $(x_1, \dots, x_{n+1}) \cdot \tilde{\nabla}(q) = 1 \in R$, $(x_1, \dots, x_{n+1})$ is a unimodular row over R, and the R-module $\mathfrak{T}_{Q^o/k}$ is free if and only if there is a matrix $M \in \mathrm{SL}_{n+1}(R)$ with first row $(x_1, \dots, x_{n+1})$. Furthermore, if $2 \in k^\times$, then Q^o is smooth over k, and $\mathfrak{T}_{Q^o/k}$ is the R-module of global sections of the relative tangent bundle $T_{Q^o/k} \to Q^o$ of Q^o over $\operatorname{Spec} k$.

Note that, for $n = 2$ and $2 \in k^\times$, the existence of a unimodular matrix over R with the first row (x_1, x_2, x_3) is equivalent to the existence of a nonvanishing section of $T_{Q^o/k}$. Indeed, $T_{Q^o/k}$ admits a nonvanishing section if and only if we can write $\mathfrak{T}_{Q^o/k} \cong R \oplus P$, with P a rank one projective R-module, which yields the isomorphism $\bigwedge_R^2 \mathfrak{T}_{Q^o/k} \cong P$. The exact sequence (1.9) gives an isomorphism $\bigwedge_R^2 \mathfrak{T}_{Q^o/k} \cong R$, so $P \cong R$ and thus $\mathfrak{T}_{Q^o/k} \cong R^2$. This gives the existence of $M \in \mathrm{SL}_3(R)$ with first row (x_1, x_2, x_3).

The original form of Zannier's question was in the following terms: taking $q := x^2 + y^2 + z^2 \in k[x, y, z]$, for which fields k is the unimodular row (x, y, z) over R the first row in a unimodular matrix over R? He showed this was the case for $k=\mathbb{Q}_p$, p odd, and asked about $k=\mathbb{Q}_2$. Note that one can just as well ask the question for k an arbitrary commutative ring and general q as above; our results only give a criterion for a positive answer to this question if k is a perfect field with 2 invertible and $n=2$. Noting our positive answer for $k=\mathbb{Q}_2$, Zannier asked in a recent private communication about the case $k=\mathbb{Z}_2$.

Remark 1.10 (explicit sections). (1) If n is odd or if the quadratic form $q:=\sum_{i=1}^{n+1} a_i x_i^2$ is isotropic over k (i.e., $q=0$ has a nontrivial solution in k), one can write down explicit nonvanishing sections of T_{Q^o}.

For n odd, the tangent-normal sequence for $Q^o\subseteq \mathbb{A}_k^{n+1}$,

$$0\to T_{Q^o}\to \mathcal{O}_{Q^o}^{n+1}\xrightarrow{(2a_1x_1,\dots,2a_{n+1}x_{n+1})^t}\mathcal{O}_{Q_o}\to 0,$$

says a section s of T_{Q^o} is given by an $(n+1)$-tuple of regular functions $(s_1,s_2,\dots,s_{n+1})$ with

$$\sum_{i=1}^{n+1} a_i x_i s_i = 0.$$

One can take

$$s=(a_2x_2,-a_1x_1,\dots,a_{n+1}x_{n+1},-a_nx_n),$$

which is clearly nonvanishing.

This is a special case of the following general result. Let A be a commutative ring, let $(a_1,\dots,a_{2m})$ be a unimodular row in A^{2m} (i.e., $a_1,\dots,a_{2m}$ generate the unit ideal in A), and let M be the stably free A-module defined by the exactness of the sequence

$$0\to M\to A^{2m}\xrightarrow{(a_1,\dots,a_{2m})^t} A\to 0.$$

Then M admits the free rank one summand defined by

$$0\to A\xrightarrow{(-a_2,a_1,\dots,-a_{2m},a_{2m-1})} M\subset A^{2m}.$$

Now take n even and suppose q is isotropic. Then after a k-linear change of coordinates and change of notation, we may assume that

$$q=2x_1x_2+\sum_{i=3}^{n+1} a_i x_i^2$$

(see, e.g., [Lam 2005, Theorem I.3.4]). In this case, the tangent-normal sequence for $Q^o\subseteq\mathbb{A}_k^{n+1}$ is

$$0\to T_{Q^o}\to \mathcal{O}_{Q^o}^{n+1}\xrightarrow{(2x_2,2x_1,2a_3x_3,\dots,2a_{n+1}x_{n+1})^t}\mathcal{O}_{Q_o}\to 0.$$

Letting

$$s=(0,a_3x_3,-x_1,a_5x_5,-a_4x_4,\dots,a_{n+1}x_{n+1},-a_nx_n)$$

gives a section of T_{Q^o} with $s=0$ given by $Q^o\cap(x_1=x_3=\dots=x_{n+1}=0)$, which is clearly empty.

In particular, let k be a field of characteristic $p>2$. In a finite field, -1 is a sum of two squares [Lam 2005, Proposition II.3.4], whence the quadratic form $x_1^2+x_2^2+\dots+x_{n+1}^2$ is isotropic over k provided that $n\geq 2$. Hence the tangent bundle $T_{S_k^n}$ to the algebraic n-sphere over k admits a nonvanishing section for every $n\geq 1$.

(2) Our main results for even n and q anisotropic only give criteria for the existence of a nonvanishing section, without giving an explicit formula. In the case that Zannier had asked about originally, $S^2_{\mathbb{Q}_2}$, Peter Müller (private communication, Universität Würzburg, June 27, 2024), noting our existence result and following a suggestion of Zannier, found an explicit trivialization of $T_{S^2_{\mathbb{Q}_2}}$. We quote from his private communication:

> Indeed, some sophisticated computations eventually gave an explicit example over $\mathbb{Q}_2$ (in fact even over $\mathbb{Q}(\sqrt{-7})$), ...

Here is Müller's example. Let $R = \mathbb{Q}_2[x, y, z]/(x^2 + y^2 + z^2 - 1)$, the coordinate ring of $S^2_{\mathbb{Q}_2}$. The polynomial $T^2 - T + 2$ has two roots in $\mathbb{Z}_2$, exactly one of which, which we denote by ω, is a unit in $\mathbb{Z}_2$. In particular, $2 - \omega$ is also a unit in $\mathbb{Z}_2$. Müller gives his example in the form of a 3×3 matrix over R with determinant $2 - \omega$ and first row (x, y, z). The explicit matrix is

$$\begin{pmatrix} x & y & z \\ -y+z+1 & (1-\omega)x+y+(1+\omega)z+\omega & -x-y+2z+(1-\omega) \\ \omega y+(2-\omega)z & (1-2\omega)x+(1+\omega)y+3z+1 & -2x+(2-\omega)z-\omega \end{pmatrix}.$$

Let λ_i be the dot product of (x, y, z) with the i-th row. Noting that $(x, y, z) \cdot (x, y, z) = 1$ in R, this gives the following two independent nonvanishing sections of $T_{S^2_{\mathbb{Q}_2}}$:

$$s_1(x, y, z) = (-y+z+1,\, (1-\omega)x+y+(1+\omega)z+\omega,\, -x-y+2z+(1-\omega)) - \lambda_2 \cdot (x, y, z),$$
$$s_2(x, y, z) = (\omega y+(2-\omega)z,\, (1-2\omega)x+(1+\omega)y+3z+1,\, -2x+(2-\omega)z-\omega) - \lambda_3 \cdot (x, y, z).$$

Müller notes that this also works over $\mathbb{Q}(\sqrt{-7})$, where we take ω to be either of the two roots of $T^2 - T + 2$ in $\mathbb{Q}(\sqrt{-7})$.

In addition, Müller's example gives a positive answer to Zannier's question over $\mathbb{Z}_2$ instead of $\mathbb{Q}_2$, just divide the last row by the determinant $2 - \omega \in \mathbb{Z}_2^\times$.

Remark 1.11 (some nonexamples). Suppose n is even. We have already seen in Remark 1.10 that T_{Q^o} has a nonvanishing section if q is isotropic over k. On the other hand, q being isotropic over k implies that Q^o has a k-rational point [Lam 2005, Theorem I.3.4(3)], so if $Q^o(k) = \varnothing$, then q is anisotropic over k and we do not have any explicit method for finding a (possible) nonvanishing section of T_{Q^o}. Moreover, Theorem 1.5 is our only result that considers the case n even and $Q^o(k) = \varnothing$, and it only gives us a necessary condition for T_{Q^o} to have a nonvanishing section. Here is a series of examples that are not covered by any of our results.

Take $k = \mathbb{Q}_p$ with $p > 2$. Let $u \in \mathbb{Z}_p^\times$ be a nonsquare modulo p, and let $q = ux_1^2 + px_2^2 - upx_3^2$. It follows from [Lam 2005, Theorem VI.2.2] that $q - x_0^2$ is anisotropic over $\mathbb{Q}_p$; hence $Q^o(\mathbb{Q}_p) = \varnothing$ and also q is anisotropic over $\mathbb{Q}_p$. Moreover $-(u \cdot p \cdot (-up)) = 1$ in $\mathbb{Q}_p^\times/\mathbb{Q}_p^{\times 2}$, so $-(u \cdot p \cdot (-up))$ is in $[D(q)^2]$; hence the necessary condition in Theorem 1.5 is satisfied. We do not know whether T_{Q^o} has a nonvanishing section in any of these cases.

One final nonexample. Take $k = \mathbb{R}$, $q = \sum_{i=1}^{n+1} -x_i^2$, with n even, and let R be the coordinate ring of Q^o. Then we have $Q^o(\mathbb{R}) = \varnothing$. By a theorem of Ojanguren, Parimala and Sridharan [Ojanguren et al.

1986, Theorem 3.2], there is an $M \in \mathrm{SL}_{n+1}(R)$ with

$$(x_1, \dots, x_{n+1}) = (1, 0, \dots, 0)M$$

in other words, $(x_1, \dots, x_{n+1})$ is the first row of the unimodular matrix M. Thus the R-module $\mathfrak{T}_{Q^o}$ is free, so T_{Q^o} is a trivial vector bundle over Q^o, and hence admits a nonvanishing section. Our results only yield the necessary condition

$$-(-1)^{n+1} = 1 \in |D(q)^2|,$$

which (fortunately!) is true in this case.

The main idea behind the proofs of our results is quite close to the case of the real spheres. We have already disposed of the case of odd n in Remark 1.10. For $n > 0$ even, we replace the Euler class $e_{\mathrm{top}}(T_{S^n}) \in H^n(S^n, \mathbb{Z})$ with the Euler class $e(T_{Q^o})$ in the *Chow–Witt group* $\widetilde{\mathrm{CH}}^n(Q^o)$. For a smooth k-variety X and a rank-r vector bundle E on X, one has an Euler class $e(E)$ in the (twisted) Chow–Witt group $\widetilde{\mathrm{CH}}^n(X, \det^{-1}(E))$. In our case, the tangent-normal sequence for T_{Q^o} gives a canonical isomorphism $\det T_{Q^o} \cong \mathcal{O}_{Q^o}$, which induces an isomorphism $\widetilde{\mathrm{CH}}^n(Q^o, \det^{-1} T_{Q^o}) \cong \widetilde{\mathrm{CH}}^n(Q^o)$ with the untwisted version of the Chow–Witt group, giving us our Euler class $e(T_{Q^o}) \in \widetilde{\mathrm{CH}}^n(Q^o)$. A fundamental result of Morel [2012, Section 8.2] says that for a smooth affine k-scheme X of dimension n over a perfect field k and a rank n vector bundle E over X, E admits a nonvanishing section if and only if the Euler class $e(E)$ vanishes (in this form, the result also relies on work of Asok and Fasel [2016] and Asok, Hoyois and Wendt [Asok et al. 2017]; see Theorem 2.2 for the discussion).

Since Q^o is not proper over k, we do not have a nice analog of the Gauss–Bonnet theorem for Q^o, so we pass to its projective closure $Q \subseteq \mathbb{P}^{n+1}$, defined by the equation $\sum_{i=1}^{n+1} a_i x_i^2 = x_0^2$, and let $Q^\infty \subseteq Q$ be the hyperplane section defined by $x_0 = 0$.

Let $\mathrm{GW}(k)$ denote the Grothendieck–Witt ring of (virtual) regular quadratic forms over k. For $p : X \to \operatorname{Spec} k$ a smooth projective variety, we have the pushforward map

$$p_* : \widetilde{\mathrm{CH}}_0(X) \to \widetilde{\mathrm{CH}}_0(\operatorname{Spec} k) = \mathrm{GW}(k),$$

which we denote by $\deg_{GW} : \widetilde{\mathrm{CH}}_0(X) \to \mathrm{GW}(k)$; we call this the *quadratic degree map*. X has a *quadratic Euler characteristic* $\chi(X/k) \in \mathrm{GW}(k)$ and we have a quadratic Gauss–Bonnet theorem [Déglise et al. 2021, Theorem 4.6.1; Levine and Raksit 2020, Theorem 5.3]: letting T_X be the tangent bundle of X, we have

$$\deg_{GW}(e(T_X)) = \chi(X/k),$$

so we are all set up to argue as in differential topology.

Getting back to our quadrics, let us first assume that Q^o has a k-rational point. We show in Section 2 that $e(T_{Q^o}) = 0$ if and only if $\chi(Q/k)$ is in the subgroup $\deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty)) \subseteq \mathrm{GW}(k)$, and we have an explicit expression for $\chi(Q/k)$:

$$\chi(Q/k) = \Big\langle 2, 2\prod_i a_i \Big\rangle + \frac{n}{2}\langle 1, -1\rangle,$$

where $\langle a, b\rangle$ is the quadratic form $ax^2 + by^2$, and $m \cdot \langle a, b\rangle$ is the quadratic form $\sum_{i=1}^m a x_i^2 + b y_i^2$.

Putting this together, we see that T_{Q^o} admits a nonvanishing section if and only if $\langle 2, 2\prod_i a_i\rangle + \frac{n}{2}\langle 1, -1\rangle$ is in $\deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty)) \subseteq \mathrm{GW}(k)$. We conclude Section 2 by using this criterion to handle the cases discussed in Examples 1.3 (1), (2) above.

The next step is to use the theory of quadratic forms to rephrase the condition

$$\Big\langle 2, 2\prod_i a_i\Big\rangle + \frac{n}{2}\langle 1, -1\rangle \in \deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty)) \subseteq \mathrm{GW}(k)$$

in terms of the subgroups $[D(q)]$ and $[D(q)^2]$. This is done in Section 3, relying on properties of Scharlau's transfer, Knebusch's norm principle and basic facts about Pfister forms and Pfister neighbors. We apply these tools in Section 4 to give our main results, which yield criteria that are much easier to apply than the one derived in Section 2. We conclude by using this to compute the remaining examples described above; the case in which Q^o does not have a k-rational point is trickier to handle, and we are only able to arrive at the necessary condition stated in Theorem 1.5.

Throughout the paper we employ the following notation:

k	a perfect field with $\operatorname{char} k \neq 2$
Sm_k	the category of smooth, separated, finite-type k-schemes
T_X	the tangent bundle of $X \in \mathrm{Sm}_k$
$X(F)$	the set of rational points of X_F for $X \in \mathrm{Sm}_k$ and a field extension F/k
$\mathrm{GW}(F)$	the Grothendieck–Witt ring of (virtual) regular quadratic forms over a field F
$\mathrm{W}(F)$	the Witt ring, the quotient of $\mathrm{GW}(F)$ by the hyperbolic forms
$\mathrm{GI}(F)$	the ideal in $\mathrm{GW}(F)$ generated by the even-dimensional forms
$I(F)$	the image in $\mathrm{W}(F)$ of $\mathrm{GI}(F)$
$F^\times$	the multiplicative group of nonzero elements of the field F
$\langle a_1, a_2, \dots, a_n\rangle$	the quadratic form $a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2$

2. Recollections on Chow–Witt groups and a computational criterion

Definition 2.1. Assume k to be a perfect field. We will use the Chow–Witt groups, also known as Chow groups of oriented cycles, that were introduced in [Barge and Morel 2000]. These groups provide refined cohomological obstructions to the existence of nonvanishing sections of algebraic vector bundles (see, e.g., [Asok and Fasel 2015; 2023]). We refer the reader to the expositions in [Fasel 2020; Déglise 2023; Asok and Fasel 2016, Sections 2, 3] for the properties of these groups that we list below.

We recall from [Morel 2012, Chapter 2] the *Milnor-Witt K-theory sheaves* $\mathcal{K}_n^{\mathrm{MW}}$, $n \in \mathbb{Z}$. These are Nisnevich sheaves of abelian groups on Sm_k, with products $\mathcal{K}_n^{\mathrm{MW}} \times \mathcal{K}_m^{\mathrm{MW}} \to \mathcal{K}_{n+m}^{\mathrm{MW}}$ making the graded object $\mathcal{K}_*^{\mathrm{MW}} := \bigoplus_{n\in\mathbb{Z}} \mathcal{K}_n^{\mathrm{MW}}$ into a sheaf of associative, unital, graded rings on Sm_k. Given $X \in \mathrm{Sm}_k$ and a line bundle L on X, we have the L-twisted version $\mathcal{K}_n^{\mathrm{MW}}(L)$, giving a Nisnevich sheaf on Sm_k/X. Letting $\mathcal{GW}$ denote the Nisnevich sheaf of Grothendieck–Witt rings, there is a canonical isomorphism $\mathcal{K}_0^{\mathrm{MW}} \cong \mathcal{GW}$. For a field F, and L a dimension one F-vector space, we write $K_n^{\mathrm{MW}}(L)(F)$ for $\mathcal{K}_n^{\mathrm{MW}}(L)(\operatorname{Spec} F)$.

For a smooth variety X over k, a line bundle L over X and an integer $n \geq 0$, the *Chow–Witt group* $\widetilde{\mathrm{CH}}^n(X, L)$ is defined as

$$\widetilde{\mathrm{CH}}^n(X, L) = H^n_{\mathrm{Zar}}(X; \mathcal{K}^{\mathrm{MW}}_n(L)).$$

We will also use the homological notation with

$$\widetilde{\mathrm{CH}}_n(X, L) = \widetilde{\mathrm{CH}}^{d-n}(X, L \otimes \omega_X),$$

where $d = \dim X$ and ω_X is the canonical bundle of X. We put

$$\widetilde{\mathrm{CH}}^n(X) := \widetilde{\mathrm{CH}}^n(X, \mathcal{O}_X), \quad \widetilde{\mathrm{CH}}_n(X) := \widetilde{\mathrm{CH}}_n(X, \mathcal{O}_X).$$

Chow–Witt groups have the following properties that we will use below.

(1) $\widetilde{\mathrm{CH}}^n(X, L)$ is canonically identified by [Morel 2012, Theorem 5.47] with the n-th cohomology group of the Rost–Schmid complex

$$\bigoplus_{x \in X^{(0)}} K^{\mathrm{MW}}_n(L_x \otimes \omega_{x/X})(k(x)) \to \bigoplus_{x \in X^{(1)}} K^{\mathrm{MW}}_{n-1}(L_x \otimes \omega_{x/X})(k(x)) \to \cdots \to \bigoplus_{x \in X^{(d)}} K^{\mathrm{MW}}_{n-d}(L_x \otimes \omega_{x/X})(k(x)),$$

where the sums are taken over all the points of X of the respective codimension, L_x is the restriction of L to x, $\omega_{x/X}$ is the determinant of the normal bundle for the embedding $x \to X$ and $d = \dim X$.

(2) For a line bundle L' over X there is a canonical isomorphism [Morel 2012, Remark 5.13]

$$\widetilde{\mathrm{CH}}^n(X, L \otimes (L')^{\otimes 2}) \cong \widetilde{\mathrm{CH}}^n(X, L).$$

(3) For a morphism $f : Y \to X$ of smooth varieties over k one has a functorial pullback homomorphism

$$f^* : \widetilde{\mathrm{CH}}^n(X, L) \to \widetilde{\mathrm{CH}}^n(Y, f^* L)$$

given by the pullback in the cohomology of sheaves. Further, if f is proper then one has a functorial pushforward homomorphism [Fasel 2020, Section 2.3]

$$f_* : \widetilde{\mathrm{CH}}_n(Y, f^* L) \to \widetilde{\mathrm{CH}}_n(X, L)$$

induced by the transfers on the Rost–Schmid complexes. For a closed embedding $i : Z \to X$ of smooth varieties, with $j : X \setminus Z \to X$ the open embedding of the complement, the localization sequence

$$\widetilde{\mathrm{CH}}_n(Z, i^* L) \xrightarrow{i_*} \widetilde{\mathrm{CH}}_n(X, L) \xrightarrow{j^*} \widetilde{\mathrm{CH}}_n(X \setminus Z, j^* L)$$

is exact [Fasel 2020, Section 2.2].

(4) Let F/k be a field extension of finite degree. Since k is perfect, F is separable over k, so the field trace $\mathrm{tr}^F_k : F \to k$ is a nonzero k-linear functional. This gives rise to the Scharlau transfer

$$(\mathrm{tr}^F_k)_* : \mathrm{GW}(F) \to \mathrm{GW}(k),$$

an additive homomorphism, which is given on generators $\langle a \rangle \in \mathrm{GW}(F)$ by defining $(\mathrm{tr}^F_k)_*(\langle a \rangle)$ to be the quadratic form $x \mapsto \mathrm{tr}^F_k(a x^2)$ on the k-vector space F. The pushforward homomorphism in Chow–Witt groups

$$\pi_* : \widetilde{\mathrm{CH}}_0(\operatorname{Spec} F) \to \widetilde{\mathrm{CH}}_0(\operatorname{Spec} k)$$

for the morphism $\pi : \operatorname{Spec} F \to \operatorname{Spec} k$ coincides by [Fasel 2020, Example 1.23] with the Scharlau transfer $(\mathrm{tr}_k^F)_*$ under the identifications

$$\widetilde{\mathrm{CH}}_0(\operatorname{Spec} F) = K_0^{\mathrm{MW}}(F) \cong \mathrm{GW}(F), \quad \widetilde{\mathrm{CH}}_0(\operatorname{Spec} k) = K_0^{\mathrm{MW}}(k) \cong \mathrm{GW}(k).$$

(5) For a rank-n vector bundle E over a smooth variety X over k one has an *Euler class*

$$e(E) = s^* s_*(1_X) \in \widetilde{\mathrm{CH}}^n(X, (\det E)^\vee),$$

where $s : X \to E$ is the zero section. This class is natural with respect to pullbacks [Asok and Fasel 2016, Proposition 3.1.1].

(6) Let E be a rank n vector bundle over a smooth affine variety X of dimension n over k. Suppose that $\det E \cong \mathcal{O}_X$. Then $e(E) = 0$ if and only if E has a nonvanishing section. For $n = 1$ there is nothing to prove, for $n = 2$ this was shown in [Barge and Morel 2000, Theorem 2.2] and for general n this follows from the results of [Morel 2012, Chapter 8; Asok et al. 2017, Theorem 1; Asok and Fasel 2016, Theorem 5.6]; see Theorem 2.2 below for the details.

Theorem 2.2 (Barge–Morel, Morel, Asok–Fasel, Asok–Hoyois–Wendt). *Let k be a perfect field and E be a rank n vector bundle over a smooth affine variety X of dimension n over k. Suppose that* $\det E \cong \mathcal{O}_X$. *Then E admits a nonvanishing section if and only if $e(E) = 0$.*

Proof. If $n = 1$, there is nothing to prove, so assume $n \geq 2$. It follows from [Morel 2012, Theorem 8.14] that E admits a nonvanishing section if and only if a certain obstruction-theoretic Euler class $e_{\mathrm{ob}}(E) \in \widetilde{\mathrm{CH}}^n(X)$ vanishes. Note that in that work it was assumed $n \geq 4$ because of the assumption $r \neq 2$ in [Morel 2012, Theorem 8.1 (3)], which can be removed using [Asok et al. 2017, Theorem 1]. It follows from [Asok and Fasel 2016, Theorem 5.6] that $e_{\mathrm{ob}}(E) = 0$ if and only if $e(E) = 0$, whence the claim. □

Remark 2.3. We expect that Theorem 2.2 generalizes to vector bundles with possibly nontrivial determinant and to general fields, removing the assumptions $\det E \cong \mathcal{O}_X$ and k being perfect.

Remark 2.4. Let X be a smooth hypersurface in $\mathbb{A}_k^{n+1}$. Since $k[x_1, \ldots, x_{n+1}]$ is a UFD, the ideal of X is principal, $I_X = (F)$. We have the tangent-normal sequence describing the tangent bundle T_X as

$$0 \to T_X \to \mathcal{O}_X^{n+1} \xrightarrow{\nabla F} \mathcal{O}_X \to 0,$$

where $\nabla F := (\partial F/\partial x_1, \ldots, \partial F/\partial x_{n+1})$ is the usual gradient of F. Since X is affine, this sequence splits, in particular, T_X is stably trivial and $\det T_X \cong \mathcal{O}_X$, so Theorem 2.2 is applicable to T_X.

Definition 2.5. Let X be a smooth proper variety over a perfect field k with the structure morphism $\pi : X \to \operatorname{Spec} k$. Then the *quadratic degree map*

$$\deg_{GW} := \pi_* : \widetilde{\mathrm{CH}}_0(X) \to \widetilde{\mathrm{CH}}_0(\operatorname{Spec} k) \cong \mathrm{GW}(k)$$

is the pushforward homomorphism for the structure morphism.

Lemma 2.6. *Let Q be a smooth projective quadric over a perfect field k. Suppose that $Q(k) \neq \varnothing$. Then the quadratic degree map*

$$\deg_{GW} : \widetilde{\mathrm{CH}}_0(Q) \to \mathrm{GW}(k)$$

is an isomorphism.

Proof. Let $n = \dim Q$ and consider the commutative square

$$\begin{tikzcd} Y \arrow[r, "g"] \arrow[d, "f"'] & Q \arrow[d, "\pi"] \\ \mathbb{P}^n \arrow[r, "p"] & \operatorname{Spec} k \end{tikzcd}$$

where $Q \xleftarrow{g} Y \xrightarrow{f} \mathbb{P}^n$ is the resolution of the birational morphism $Q \dashrightarrow \mathbb{P}^n$ given by the projection from a rational point on Q. Note that all maps in this square are proper. This gives a commutative diagram of pushforward homomorphisms:

$$\begin{tikzcd} \widetilde{\mathrm{CH}}_0(Y) \arrow[r, "g_*"] \arrow[d, "f_*"'] & \widetilde{\mathrm{CH}}_0(Q) \arrow[d, "\pi_*"] \\ \widetilde{\mathrm{CH}}_0(\mathbb{P}^n) \arrow[r, "p_*"] & \mathrm{GW}(k) \end{tikzcd}$$

It follows from the birational invariance of $\widetilde{\mathrm{CH}}_0$ [Feld 2022, Corollary 2.2.11] that f_* and g_* are isomorphisms. Recall that $\omega_{\mathbb{P}^n} = \mathcal{O}_{\mathbb{P}^n}(-n-1)$, whence

$$\widetilde{\mathrm{CH}}_0(\mathbb{P}^n) = \widetilde{\mathrm{CH}}^n(\mathbb{P}^n, \mathcal{O}(-n-1)) \cong \begin{cases} \widetilde{\mathrm{CH}}^n(\mathbb{P}^n), & n \text{ is odd}, \\ \widetilde{\mathrm{CH}}^n(\mathbb{P}^n, \mathcal{O}_{\mathbb{P}^n}(-1)), & n \text{ is even}. \end{cases}$$

The homomorphism p_* is an isomorphism by [Fasel 2013, Corollary 11.8]. Thus $\pi_* = \deg_{GW}$ is an isomorphism as well. □

Definition 2.7. A smooth projective scheme X over k has a *quadratic Euler characteristic* $\chi(X/k) \in \mathrm{GW}(k)$, arising from the categorical Euler characteristic of the dualizable object $\Sigma^\infty_{\mathbb{P}^1} X_+$ in the motivic stable homotopy category $\mathrm{SH}(k)$, together with Morel's theorem [2004, Theorem 6.4.1, Remark 6.4.2] identifying the endomorphisms of the unit in $\mathrm{SH}(k)$ with $\mathrm{GW}(k)$ (see [Hoyois 2014, Section 1; Levine 2020, Section 2.1] for details). The motivic Gauss–Bonnet theorem [Déglise et al. 2021, Theorem 4.6.1; Levine and Raksit 2020, Theorem 5.3] gives the identity

$$\chi(X/k) = \deg_{GW}(e(T_X)) \in \mathrm{GW}(k). \tag{2.8}$$

Theorem 2.9. *Let Q^o be the affine quadric over a perfect field k given by the equation*

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2 = 1,$$

with $a_1, \ldots, a_{n+1} \in k^\times$ and let Q^∞ be the projective quadric given by the equation

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2 = 0.$$

Then the following hold:

(1) *If n is odd, then the tangent bundle T_{Q^o} has a nonvanishing section.*

(2) *If $n>0$ is even and the tangent bundle T_{Q^o} has a nonvanishing section, then*

$$\frac{n}{2}\cdot\langle 1,-1\rangle+\left\langle 2,2\cdot\prod_{i=1}^{n+1}a_i\right\rangle\in \deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty))\subseteq \mathrm{GW}(k).$$

(3) *If $n>0$ is even and Q^o has a rational point, then the tangent bundle T_{Q^o} has a nonvanishing section if and only if*

$$\frac{n}{2}\cdot\langle 1,-1\rangle+\left\langle 2,2\cdot\prod_{i=1}^{n+1}a_i\right\rangle\in \deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty))\subseteq \mathrm{GW}(k).$$

Proof. (1) We have settled the case of odd n in Remark 1.10.

(2), (3) Let Q be the compactification of Q^o given by the equation

$$a_1x_1^2+a_2x_2^2+\cdots+a_{n+1}x_{n+1}^2=x_0^2$$

in the projective space $\mathbb{P}^{n+1}$ and let $j:Q^o\to Q$ be the open embedding. Then $Q^\infty=Q\setminus j(Q^o)$; let $i:Q^\infty\to Q$ be the closed embedding. Consider the localization sequence and the quadratic degree homomorphisms

$$\begin{tikzcd} \widetilde{\mathrm{CH}}_0(Q^\infty) \arrow[r, "i_*"] \arrow[dr, "\deg_{GW}"'] & \widetilde{\mathrm{CH}}_0(Q) \arrow[r, "j^*"] \arrow[d, "\deg_{GW}"] & \widetilde{\mathrm{CH}}_0(Q^o) \\ & \mathrm{GW}(k) & \end{tikzcd}$$

We have the identifications

$$\widetilde{\mathrm{CH}}_0(Q)=\widetilde{\mathrm{CH}}^n(Q,\omega_Q)=\widetilde{\mathrm{CH}}^n(Q,(\det T_Q)^\vee),$$

so we consider the Euler class $e(T_Q)\in\widetilde{\mathrm{CH}}^n(Q,(\det T_Q)^\vee)$ as being in $\widetilde{\mathrm{CH}}_0(Q)$.

Exactness of the localization sequence yields that the Euler class $e(T_{Q^0})=e(j^*T_Q)=j^*e(T_Q)$ vanishes if and only if $e(T_Q)\in i_*\widetilde{\mathrm{CH}}_0(Q^\infty)$. By (2.8) and [Levine 2020, Corollary 12.2] we have

$$\deg_{GW}(e(T_Q))=\chi(Q/k)=\frac{n}{2}\cdot\langle 1,-1\rangle+\left\langle 2,2\cdot\prod_{i=1}^{n+1}a_i\right\rangle. \tag{2.10}$$

Suppose that T_{Q^o} has a nonvanishing section. Then $e(T_{Q^0})=0$ and hence $e(T_Q)$ is in $i_*\widetilde{\mathrm{CH}}_0(Q^\infty)$. Taking quadratic degrees and using formula (2.10) we obtain

$$\deg_{GW}(e(T_Q))=\frac{n}{2}\cdot\langle 1,-1\rangle+\left\langle 2,2\cdot\prod_{i=1}^{n+1}a_i\right\rangle\in\deg_{GW}(i_*\widetilde{\mathrm{CH}}_0(Q^\infty))=\deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty)),$$

proving (2) and one implication in (3).

Now suppose that Q^o has a rational point and $\frac{n}{2}\cdot\langle 1,-1\rangle+\langle 2,2\cdot\prod_{i=1}^{n+1}a_i\rangle\in\deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty))$. Note that $\deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty))=\deg_{GW}(i_*\widetilde{\mathrm{CH}}_0(Q^\infty))$, whence Lemma 2.6 combined with (2.10) show that

$e(T_Q) \in i_*\widetilde{\mathrm{CH}}_0(Q^\infty)$, yielding $e(T_{Q^o}) = 0$. Remark 2.4 provides an isomorphism $\det T_{Q^o} \cong \mathcal{O}_{Q^o}$, whence Theorem 2.2 implies that T_{Q^o} has a nonvanishing section, completing the proof of (3). $\square$

Example 2.11. Let S_k^2 be the quadric over a field k given by the equation $x^2 + y^2 + z^2 = 1$.

(1) If the equation $x^2 + y^2 = -1$ has a solution over k then the conic $C_k \subseteq \mathbb{P}_k^2$ given by the equation $x^2 + y^2 + z^2 = 0$ has a rational point whence $\deg_{GW}(\widetilde{\mathrm{CH}}_0(C_k)) = \mathrm{GW}(k)$ and Theorem 2.9 yields that $T_{S_k^2}$ has a nonvanishing section. In particular [Lam 2005, Example XI.2.4(2) and (6)] yield that this holds for $k = \mathbb{Q}_p$ the field of p-adic numbers and for $k = \mathbb{F}_{p^n}$ a finite field, with $p \neq 2$ in both cases. By base-change, the same follows if $k \supset \mathbb{Q}_p$ or k has characteristic p, with $p \neq 2$. An explicit nonvanishing section of $T_{S_k^2}$ in these cases may be found as in Remark 1.10.

(2) Let $k = \mathbb{Q}_2$ be the field of dyadic numbers. Then the equation $x^2 + y^2 = -1$ has no solution over k by, e.g., [Lam 2005, Example XI.2.4(7)] and the conic $C_k \subseteq \mathbb{P}_k^2$ given by the equation $x^2 + y^2 + z^2 = 0$ has no rational points. However, it is clear that C_k has a rational point over $\mathbb{Q}_2(\sqrt{-2})$. Moreover, since 2 is equivalent to -14 modulo squares in $\mathbb{Q}_2$ (see, e.g., [Lam 2005, Corollary VI.2.24]), C_k has the point $(6 + \sqrt{-14},\, 10,\, -2 + 3\sqrt{-14})$ over $\mathbb{Q}_2(\sqrt{2})$. A straightforward computation shows that

$$(\mathrm{tr}_{\mathbb{Q}_2}^{\mathbb{Q}_2(\sqrt{\pm 2})})_*(\langle 1\rangle) = \langle 2, \pm 1\rangle,$$

whence

$$\langle 1, -1\rangle + \langle 2, 2\rangle = (\mathrm{tr}_{\mathbb{Q}_2}^{\mathbb{Q}_2(\sqrt{2})})_*(\langle 1\rangle) + (\mathrm{tr}_{\mathbb{Q}_2}^{\mathbb{Q}_2(\sqrt{-2})})_*(\langle 1\rangle) \in \deg_{GW}(\widetilde{\mathrm{CH}}_0(C_k))$$

and Theorem 2.9 (3) yields that $T_{S_k^2}$ has a nonvanishing section. Alternatively, we have

$$(\mathrm{tr}_{\mathbb{Q}_2}^{\mathbb{Q}_2(\sqrt{2})})_*(\langle 1 + \sqrt{2}\rangle) = \langle -2, 1\rangle,$$

whence

$$\langle 1, -1\rangle + \langle 2, 2\rangle = \langle 2, -2\rangle + \langle 1, 1\rangle = (\mathrm{tr}_{\mathbb{Q}_2}^{\mathbb{Q}_2(\sqrt{2})})_*(\langle 1, 1 + \sqrt{2}\rangle) \in \deg_{GW}(\widetilde{\mathrm{CH}}_0(C_k)).$$

Note that C_k does not have a $\mathbb{Q}_2$-point, so we cannot apply Remark 1.10. We were not able to produce an explicit nonvanishing section in this case.

Example 2.12. We can expand on the last example as follows. Let S_k^{2n} be the quadric over a field k given by the equation $\sum_{i=1}^{2n+1} x_i^2 = 1$, $n > 0$, and suppose k contains a p-adic field $\mathbb{Q}_p$. Then $T_{S_k^{2n}}$ has a nonvanishing section. Indeed, letting $C_k^{2n} \subseteq \mathbb{P}_k^{2n}$ be the projective quadric defined by $\sum_{i=1}^{2n+1} x_i^2 = 0$, we have just seen that $\langle 1, -1\rangle + \langle 2, 2\rangle$ is in $\deg_{GW}(\widetilde{\mathrm{CH}}_0(C_k^2)) \subseteq \mathrm{GW}(k)$. But for an arbitrary quadratic extension $k(\sqrt{a})$ of k, we have

$$(\mathrm{tr}_k^{k(\sqrt{a})})_*(\langle \sqrt{a}\rangle) = \langle 1, -1\rangle,$$

so $\langle 1, -1\rangle$ is in $\deg_{GW}(\widetilde{\mathrm{CH}}_0(Q))$ for *every* smooth projective quadric Q over k, and thus we have

$$n \cdot \langle 1, -1\rangle + \langle 2, 2\rangle \in \deg_{GW}(\widetilde{\mathrm{CH}}_0(C_k^2)) \subseteq \deg_{GW}(\widetilde{\mathrm{CH}}_0(C_k^{2n})) \subseteq \mathrm{GW}(k).$$

We then apply Theorem 2.9 (3) to conclude that $T_{S_k^{2n}}$ has a nonvanishing section. Just as in Example 2.11 (1), we can produce an explicit nonvanishing section if C_k^{2n} has a k-rational point. This is the case if $\mathbb{Q}_p \subseteq k$ with p an odd prime or if $n \geq 2$ and $\mathbb{Q}_2 \subseteq k$ (for this last case, see [Lam 2005, Theorem VI.2.12]).

Remark 2.13. Returning to the example of $S^2_{\mathbb{Q}_2}$, Nanjun Yang asked if one could completely compute $\widetilde{\mathrm{CH}}_0(S^2_{\mathbb{Q}_2})$. From the localization exact sequence in the proof of Theorem 2.9, we have the isomorphism

$$\widetilde{\mathrm{CH}}_0(S^2_{\mathbb{Q}_2}) \cong \mathrm{GW}(\mathbb{Q}_2)/\deg_{GW}(\widetilde{\mathrm{CH}}_0(C_{\mathbb{Q}_2})).$$

We may identify $C_{\mathbb{Q}_2}$ with the Severi–Brauer variety associated to the standard quaternions $\mathbb{H}_{\mathbb{Q}_2}$ over $\mathbb{Q}_2$. By [Serre 1962, XIII, Proposition 6], the Brauer group of a local field K is isomorphic to $\mathbb{Q}/\mathbb{Z}$ by the map

$$\mathrm{inv}_K : \mathrm{Br}(K) \to \mathbb{Q}/\mathbb{Z},$$

so, for a degree-2 extension $k \supset \mathbb{Q}_2$, $C_{\mathbb{Q}_2}(k) \neq \varnothing$ if and only if k splits $\mathbb{H}_{\mathbb{Q}_2}$, i.e., if and only if the invariant $\mathrm{inv}_k(\mathbb{H}_k)$ in $\mathbb{Q}/\mathbb{Z}$ is zero. But by [Serre 1962, XIII, Proposition 7],

$$\mathrm{inv}_k(\mathbb{H}_k) = 2 \cdot \mathrm{inv}_{\mathbb{Q}_2}(\mathbb{H}_{\mathbb{Q}_2}) = 2 \cdot \tfrac{1}{2} = 0,$$

so $C_{\mathbb{Q}_2}(\mathbb{Q}_2(\sqrt{a})) \neq \varnothing$ for every nonsquare $a \in \mathbb{Q}_2^\times$.

Thus, $\langle 2, 2a\rangle = \mathrm{tr}_{\mathbb{Q}_2(\sqrt{a})/\mathbb{Q}_2}(\langle 1\rangle)$ is in $\deg_{GW}(\widetilde{\mathrm{CH}}_0(C_{\mathbb{Q}_2}))$ for all nonsquares a, so the ideal I in $\mathrm{GW}(\mathbb{Q}_2)$ generated by the forms

$$\{\langle 2, u\rangle \mid u = 3, 5, 7\} \cup \{\langle 2, 2u\rangle \mid u = 1, 3, 5, 7\}$$

is contained in $\deg_{GW}(\widetilde{\mathrm{CH}}_0(C_{\mathbb{Q}_2}))$. It is easy to see that I is exactly the ideal in $\mathrm{GI}(\mathbb{Q}_2)$ of even-rank forms in $\mathrm{GW}(\mathbb{Q}_2)$; since $\deg_{GW}(\widetilde{\mathrm{CH}}_0(C_{\mathbb{Q}_2}))$ is clearly contained $\mathrm{GI}(\mathbb{Q}_2)$, we have $\deg_{GW}(\widetilde{\mathrm{CH}}_0(C_{\mathbb{Q}_2})) = \mathrm{GI}(\mathbb{Q}_2)$ and

$$\widetilde{\mathrm{CH}}_0(S^2_{\mathbb{Q}_2}) \cong \mathbb{Z}/2$$

via the mod 2 rank map $\mathrm{GW}(\mathbb{Q}_2) \to \mathbb{Z}/2$.

3. Scharlau's transfer for closed points on a quadric

Definition 3.1. Let F be a field, char $F \neq 2$. We denote by $\mathrm{GI}(F) \subseteq \mathrm{GW}(F)$ the ideal consisting of the even-dimensional (virtual) regular quadratic forms in the Grothendieck–Witt ring of F.

Definition 3.2. Let E/F be a field extension of finite degree, char $F \neq 2$, and $s : E \to F$ be a nonzero F-linear functional. Then the *Scharlau transfer* [1969]

$$s_* : \mathrm{GW}(E) \to \mathrm{GW}(F)$$

is the additive homomorphism such that $s_*(\langle a\rangle)$ is the quadratic form $x \mapsto s(ax^2)$ on the F-vector space E. See [Lam 2005, Chapter VII] for some of the properties of the Scharlau transfer.

Let X be a variety over k. The *transfer ideal of* X is given by

$$\mathrm{GI}_X^{\mathrm{tr}} = \sum_{x \in X_{(0)}} (s_x)_*(\mathrm{GW}(k(x))) \subseteq \mathrm{GW}(k),$$

with the sum taken over all the closed points of X and $\{s_x : k(x) \to k\}_{x \in X_{(0)}}$ being a chosen set of nonzero k-linear functionals. Note that $\mathrm{GI}_X^{\mathrm{tr}}$ does not depend on the choices of s_x [Lam 2005, Remark VII.1.6(C)]. It is easy to see that the transfer ideal admits an alternative description as

$$\mathrm{GI}_X^{\mathrm{tr}} = \sum_{\substack{F/k \text{ finite} \\ X(F) \neq \varnothing}} (s_F)_*(\mathrm{GW}(F)) \subseteq \mathrm{GW}(k),$$

with the sum taken over all the (isomorphism classes) of field extensions F/k of finite degree such that X_F has a rational point and $\{s_F : F \to k\}_F$ being some chosen set of nonzero k-linear functionals.

Remark 3.3. The transfer ideal $s_*(\mathrm{GW}(F)) \subseteq \mathrm{GW}(k)$ for an extension of fields F/k of finite degree is a classical object of study; see, e.g., [Lam 2005, Chapter VII]. This agrees with the notion introduced above if one considers F as a zero-dimensional variety Spec F over k.

Lemma 3.4. *Let X be a smooth proper variety over a perfect field k. Then*

$$\mathrm{GI}_X^{\mathrm{tr}} = \deg_{GW}(\widetilde{\mathrm{CH}}_0(X)).$$

Proof. This follows from the description of $\widetilde{\mathrm{CH}}_0(X)$ via the cohomology of the Rost–Schmid complex and the fact that the pushforward in $\widetilde{\mathrm{CH}}_0$ for a separable field extension of finite degree coincides with the Scharlau transfer for the field trace. □

Lemma 3.5. *Let E/F be a field extension of even degree*, char $F \neq 2$, *and $s : E \to F$ be an F-linear nonzero functional. Then $s_*(\mathrm{GW}(E)) \subseteq s_*(\mathrm{GI}(E)) + \langle 1, -1\rangle \cdot \mathrm{GW}(F)$.*

Proof. Note that the claim does not depend on the choice of s (see [Lam 2005, Remark VII.1.6(C)]). Without loss of generality we may assume that $E = F(\alpha)/F$ is a simple extension. Then there is a nonzero functional s such that $s_*(\langle\alpha\rangle) = \frac{1}{2}[E : F] \cdot \langle 1, -1\rangle$ [Lam 2005, Theorem VII.2.3]. The claim follows, since $s_*(\phi) = s_*(\phi + \langle\alpha\rangle) - \frac{1}{2}[E : F] \cdot \langle 1, -1\rangle$. □

Definition 3.6. Let q be a regular quadratic form over k. Then we use the following notation:

- $D(q)$ is the set of nonzero values of q.
- $D(q)^2 = \{a \cdot b \mid a, b \in D(q)\}$ is the set of products of pairs of nonzero values of q.
- $[D(q)]$ and $[D(q)^2]$ are the multiplicative subgroups of $k^\times$ generated by the respective sets.

Note that if $1 \in D(q)$ then $[D(q)] = [D(q)^2]$. Since all the sets introduced above are stable under multiplication by squares $(k^\times)^2$, we will sometimes abuse the notation and denote in the same way the corresponding subsets of $k^\times/(k^\times)^2$.

Definition 3.7. For a regular quadratic form q over k of dimension n the *signed discriminant* $\delta_{\pm}(q)$ is given by the formula

$$\delta_{\pm}(q) := (-1)^{n(n-1)/2} \det A_q,$$

where A_q is a symmetric matrix representing q. This gives a well-defined map

$$\delta_{\pm} : \mathrm{GW}(k) \to k^{\times}/(k^{\times})^2.$$

When restricted to the ideal $\mathrm{GI}(k)$, this map becomes a homomorphism.

Lemma 3.8. *Let Q be a smooth projective quadric over k defined by a quadratic form q and take $\phi \in \mathrm{GI}^{\mathrm{tr}}_Q$. Then $\delta_{\pm}(\phi)$ is in $[D(q)^2]$.*

Proof. We may assume $Q(k) = \varnothing$; otherwise q is isotropic and $D(q) = k^{\times}$, whence there is nothing to prove. Springer's theorem [Lam 2005, Theorem VII.2.7] yields that for every closed point $x \in Q_{(0)}$ the degree $[k(x) : k]$ is even. Hence $\mathrm{GI}^{\mathrm{tr}}_Q \subseteq \mathrm{GI}(k)$, whence $\delta_{\pm}$ restricted to $\mathrm{GI}^{\mathrm{tr}}_Q$ is a homomorphism. Thus it is sufficient to check the claim for $\phi = s_*(\psi)$ with $\psi \in \mathrm{GW}(k(x))$ for a closed point $x \in Q_{(0)}$ and s a chosen k-linear functional $s : k(x) \to k$. Furthermore, by Lemma 3.5, we may assume $\psi \in \mathrm{GI}(k(x))$. By [Scharlau 1985, Chapter II, Theorem 5.12] we have

$$\delta_{\pm}(s_*(\psi)) = N_{k(x)/k}(\delta_{\pm}(\psi)) \in k^{\times}/(k^{\times})^2.$$

The quadric $Q_{k(x)}$ has a rational point, whence $q_{k(x)}$ is isotropic and $D(q_{k(x)}) = k(x)^{\times}$; in particular, $\delta_{\pm}(\psi) \in D(q_{k(x)})$. Then Knebusch's norm principle [Lam 2005, Theorem VII.5.1] implies that $N_{k(x)/k}(\delta_{\pm}(\psi)) \in [D(q)^2]$. □

Remark 3.9. If q is a Pfister form then the last result was obtained in [Bhatwadekar et al. 2014, Lemma 3.6].

Lemma 3.10. *Let Q be a smooth projective quadric over k defined by a quadratic form q. Then $\langle a, b\rangle \in \mathrm{GI}^{\mathrm{tr}}_Q$ if and only if $-ab \in [D(q)^2]$.*

Proof. Since $\delta_{\pm}(\langle a, b\rangle) = -ab$, one implication follows from Lemma 3.8. For the other implication, first note that we may assume $Q(k) = \varnothing$, since otherwise $\mathrm{GI}^{\mathrm{tr}}_Q = \mathrm{GW}(k)$ and there is nothing to prove. Then there exists a closed point $x \in Q_{(0)}$ such that $[k(x) : k] = 2$ and we may choose $\alpha \in k$ such that $k(x) \cong k(\sqrt{\alpha})$. Then for the k-linear functional

$$s : k(\sqrt{\alpha}) \to k, \quad s(1) = 0,\ s(\sqrt{\alpha}) = 1,$$

one has $s_*(\langle 1\rangle) = \langle 1, -1\rangle$, whence

$$\langle 1, -1\rangle \in \mathrm{GI}^{\mathrm{tr}}_Q .$$

Taking $c_1, c_2 \in k^{\times}$, we have

$$\langle 1, -c_1c_2\rangle = \langle c_1\rangle\langle 1, -c_2\rangle + \langle 1, -c_1\rangle - \langle 1, -1\rangle.$$

Recalling that $\mathrm{GI}^{\mathrm{tr}}_Q$ is an ideal in $\mathrm{GW}(k)$, it follows that

$$\langle 1, -c_1\rangle, \langle 1, -c_2\rangle \in \mathrm{GI}^{\mathrm{tr}}_Q \quad \Longrightarrow \quad \langle 1, -c_1c_2\rangle \in \mathrm{GI}^{\mathrm{tr}}_Q\,. \tag{3.11}$$

We claim that

$$c, d \in D(q) \quad \Longrightarrow \quad \langle 1, -cd\rangle \in \mathrm{GI}^{\mathrm{tr}}_Q\,. \tag{3.12}$$

Accepting our claim for the moment, write $-ab \in [D(q)^2]$ as a product, $-ab := \prod_i a_i b_i$, with $a_i, b_i \in D(q)$. By (3.12), we have $\langle 1, -a_ib_i\rangle \in \mathrm{GI}^{\mathrm{tr}}_Q$ for each i. By (3.11), it follows that $\left\langle 1, -\prod_i a_ib_i\right\rangle = \langle 1, ab\rangle$ is in $\mathrm{GI}^{\mathrm{tr}}_Q$. We proceed to prove (3.12).

First suppose that $\dim Q = 0$. Then we may assume $q = x_1^2 - \alpha x_2^2$ and $Q \cong \operatorname{Spec} k(\sqrt{\alpha})$. A straightforward computation with the same functional s as above shows that

$$s_*(\langle w_1 + \sqrt{\alpha} w_2\rangle) = \langle 1, -(u_1^2 - \alpha u_2^2)(v_1^2 - \alpha v_2^2)\rangle$$

for $w_1 = (u_1v_1 + \alpha u_2v_2)/(u_1v_2 + u_2v_1)$ and $w_2 = 1$, which proves (3.12) in this case.

Now suppose that $\dim Q \geq 1$. Let $q(u) = a,\ q(v) = b$ and choose some $w \neq 0$ such that $\psi_q(u, w) = \psi_q(v, w) = 0$, where ψ_q is the symmetric bilinear form associated to q. Put $c = q(w)$ and let $\alpha = -a/c$. Then

$$q(u + \sqrt{\alpha} w) = q(u) + \alpha q(w) + 2\sqrt{\alpha}\psi_q(u, w) = q(u) + \alpha q(w) = 0.$$

Thus there is a closed point $x \in Q_{(0)}$ such that $k(x) \cong k(\sqrt{-a/c})$. Then for the same functional s as above one has

$$s_*(\langle\sqrt{-a/c}\rangle) = \langle 1, -a/c\rangle \in \mathrm{GI}^{\mathrm{tr}}_Q\,.$$

The same argument shows that $\langle 1, -b/c\rangle$ is in $\mathrm{GI}^{\mathrm{tr}}_Q$. Using (3.11), we see that $\langle 1, -ab\rangle$ also belongs to $\mathrm{GI}^{\mathrm{tr}}_Q$ and (3.12) follows. □

Remark 3.13. Let Q be a smooth projective quadric over a field k defined by a quadratic form q. Then the group $[D(q)^2]$ coincides with the group of norms $N_Q(k)$ of Q, i.e., with the multiplicative subgroup of $k^\times$ generated by the norms $N_{F/k}(a)$, with $a \in F^\times$ and F/k being an extension of fields of finite degree such that Q_F has a rational point [Colliot-Thélène and Skorobogatov 1993, Lemma 2.2].

Definition 3.14. For $a_1, a_2, \ldots, a_n \in k^\times$ an *n-fold Pfister form* $\langle\langle a_1, a_2, \ldots, a_n\rangle\rangle$ is the quadratic form $\prod_{i=1}^n \langle 1, -a_i\rangle$ of dimension 2^n. A regular quadratic form q over k is called a *Pfister neighbor* if there exists $a \in k^\times$ such that $\langle a\rangle \cdot q$ is a subform of an n-fold Pfister form with $2^{n-1} < \dim q$. Note that the Pfister form containing $\langle a\rangle \cdot q$ for a Pfister neighbor q is unique [Lam 2005, Proposition X.4.17].

Lemma 3.15. *Let q be a Pfister neighbor over a field k with the associated Pfister form ϕ and let Q and Φ be the projective quadrics given by $q = 0$ and $\phi = 0$ respectively. Then $\mathrm{GI}^{\mathrm{tr}}_Q = \mathrm{GI}^{\mathrm{tr}}_\Phi$ and $[D(q)^2] = [D(\phi)^2] = D(\phi)$.*

Proof. Suppose ϕ is an n-fold Pfister form and q has dimension $m > 2^{n-1}$. Let $a \in k^\times$ be such that $\langle a\rangle \cdot q$ is a subform of ϕ. Since the quadrics associated to q and $\langle a\rangle \cdot q$ are the same and $[D(q)^2] = [D(\langle a\rangle \cdot q)^2]$, we may assume that q is a subform of ϕ. We claim that for a field extension F/k, $Q(F) \neq \varnothing$ if and only if

$\Phi(F) \neq \varnothing$. Indeed, since q is a subform of ϕ, we have $Q \subseteq \Phi \subset \mathbb{P}_k^{2^n-1}$, and, moreover, $Q = \Phi \cap L$, where L is some codimension-$(2^n - m)$ linear subspace of $\mathbb{P}_k^{2^n-1}$. Thus if $Q(F) \neq \varnothing$ then $\Phi(F) \neq \varnothing$. Now let F be such that $\Phi(F) \neq \varnothing$. Then ϕ_F is isotropic, whence hyperbolic [Lam 2005, Theorem X.1.7], so Φ_F contains a linear subspace $L' \subset \mathbb{P}_F^{2^n-1}$ of dimension $2^{n-1} - 1$. Letting $L_F \subset \mathbb{P}_F^{2^n-1}$ be the base-extension of $L \subset \mathbb{P}_k^{2^n-1}$ to F, we see that Q_F contains the linear subspace $L' \cap L_F \subset \mathbb{P}_F^{2^n-1}$ of dimension at least $2^{n-1} - 1 - (2^n - m) = m - 2^{n-1} - 1 \geq 0$, whence $Q(F) \neq \varnothing$.

The transfer ideals are generated by the Scharlau transfers for the field extensions F/k of finite degree such that $Q(F) \neq \varnothing$ (respectively, $\Phi(F) \neq \varnothing$), whence it follows from the above that $\mathrm{GI}_Q^{\mathrm{tr}} = \mathrm{GI}_\Phi^{\mathrm{tr}}$. Then the equality $[D(q)^2] = [D(\phi)^2]$ follows from Lemma 3.10 because $[D(q)^2]$ and $[D(\phi)^2]$ coincide with the sets of signed discriminants of the binary forms from the respective transfer ideals. Since $1 \in D(\phi)$, we have $[D(\phi)^2] = [D(\phi)]$ and the last equality $[D(\phi)] = D(\phi)$ follows from [Lam 2005, Theorem XI.1.1].

Alternatively, for the equality $[D(q)^2] = [D(\phi)^2]$ one could apply the description of these groups as the groups of norms [Colliot-Thélène and Skorobogatov 1993, Lemma 2.2]. □

4. Nonvanishing vector fields on affine quadrics via groups of values

Using the results of the previous section, we can reformulate Theorem 2.9 in a more manageable form.

Theorem 4.1. *Let $q = a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2$ be a quadratic form over a perfect field k with $a_1, \ldots, a_{n+1} \in k^\times$, and let Q^o be the affine quadric given by the equation*

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2 = 1.$$

Then the following hold:

(1) *If n is odd then the tangent bundle T_{Q^o} has a nonvanishing section.*

(2) *If $n > 0$ is even and the tangent bundle T_{Q^o} has a nonvanishing section then*
$$-\prod_{i=1}^{n+1} a_i \in [D(q)^2].$$

(3) *If $n > 0$ is even and $Q^o(k) \neq \varnothing$ then the tangent bundle T_{Q^o} has a nonvanishing section if and only if $-1 \in [D(q)]$.*

Proof. The case of odd n follows from Theorem 2.9 (1), so we assume $n > 0$ is even. Let $Q^\infty \subseteq \mathbb{P}^n$ be the quadric given by $q = 0$. By Lemma 3.4 we have

$$\frac{n}{2} \cdot \langle 1, -1\rangle + \Big\langle 2, 2 \cdot \prod_i a_i \Big\rangle \in \deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty)) \iff \frac{n}{2} \cdot \Big\langle 1, -1\rangle + \langle 2, 2 \cdot \prod_i a_i \Big\rangle \in \mathrm{GI}_{Q^\infty}^{\mathrm{tr}}.$$

Note one trivially has $1 \in [D(q)^2]$. Thus, Lemma 3.10 yields $\langle 1, -1\rangle \in \mathrm{GI}_{Q^\infty}^{\mathrm{tr}}$ and implies in addition that

$$\frac{n}{2} \cdot \langle 1, -1\rangle + \Big\langle 2, 2 \cdot \prod_i a_i \Big\rangle \in \mathrm{GI}_{Q^\infty}^{\mathrm{tr}} \iff -\prod_{i=1}^{n+1} a_i \in [D(q)^2].$$

Applying Theorem 2.9(2) proves (2).

For (3), note that $Q^o(k) \neq \varnothing$ if and only if $1 \in D(q)$, which implies that $[D(q)] = [D(q)^2]$. Since each a_i is in $D(q)$, we see that $-\prod_{i=1}^{n+1} a_i \in [D(q)^2]$ if and only if $-1 \in [D(q)]$. Thus, $-1 \in [D(q)]$ if and only if $\frac{n}{2}\langle 1, -1\rangle + \langle 2, 2 \cdot \prod_i a_i\rangle$ is in $\deg_{GW}(\widetilde{\mathrm{CH}}_0(Q^\infty))$, and then Theorem 2.9 (3) implies the claim. □

Definition 4.2. Let F be a field. The *level* of F, denoted by $s(F)$, is the minimal integer n such that $-1 \in D(x_1^2 + x_2^2 + \cdots + x_n^2)$. If no such n exists then $s(F) = \infty$. The level of a field is either infinite or a power of 2 [Lam 2005, Pfister's Level Theorem XI.2.2].

Corollary 4.3. *Let S_k^n, $n \geq 1$, be the affine quadric over a field k given by the equation*

$$x_1^2 + x_2^2 + \cdots + x_{n+1}^2 = 1.$$

Then the tangent bundle $T_{S_k^n}$ has a nonvanishing section if and only if one of the following holds:

(1) *n is odd.*

(2) *$n > 0$ is even and $s(k) \leq 2n + 1$.*

Proof. First assume that $\operatorname{char} k > 2$. Then Remark 1.10 yields that the tangent bundle $T_{S_k^n}$ has a nonvanishing section for every $n \geq 1$. At the same time -1 is a square or a sum of two squares in a finite field, whence in k, thus $s(k)$ is 1 or 2 and, in particular, $s(k) \leq 2n + 1$ for $n \geq 1$. This yields the claim in the positive characteristic.

Assume $\operatorname{char} k = 0$, in particular, k is perfect. By Theorem 4.1 we need to show that for even n one has

$$-1 \in [D(x_1^2 + x_2^2 + \cdots + x_{n+1}^2)]$$

if and only if $s(k) \leq 2n + 1$. Let m be such that $2^{m-1} < n + 1 \leq 2^m$. Then $x_1^2 + x_2^2 + \cdots + x_{n+1}^2$ is a Pfister neighbor with the associated Pfister form $x_1^2 + x_2^2 + \cdots + x_{2^m}^2$. Lemma 3.15 yields

$$[D(x_1^2 + x_2^2 + \cdots + x_{n+1}^2)] = D(x_1^2 + x_2^2 + \cdots + x_{2^m}^2).$$

Thus $-1 \in [D(x_1^2 + x_2^2 + \cdots + x_{n+1}^2)]$ if and only if $s(k) \leq 2^m$. The claim follows since $s(k)$ is a power of 2. □

Example 4.4. Let S_k^2 be the quadric over a field k given by the equation $x^2 + y^2 + z^2 = 1$. If $k = \mathbb{R}$ then $T_{S_k^2}$ has no nonvanishing sections since by a classical result of Poincaré the real vector bundle $T_{S_k^2}(\mathbb{R})$ has no nonvanishing continuous sections in the Euclidean topology [tom Dieck 2008, Theorem 6.5.5]. More generally, Corollary 4.3 yields that $T_{S_k^2}$ has no nonvanishing sections if and only if $s(k) \geq 8$ (including the case of $s(k) = \infty$). In particular, $T_{S_k^2}$ has a nonvanishing section for the following fields (cf. Example 2.11):

(1) k a quadratically closed field,

(2) k a field of characteristic $p > 2$,

(3) k a non-Archimedean local field,

(4) k a purely imaginary number field.

See [Lam 2005, Example XI.2.4] for the relevant computations of $s(k)$.

Corollary 4.5. *Let k be a perfect field of characteristic $\neq 2$ such that every Pfister form of dimension 8 is hyperbolic, i.e., that $I(k)^3 = 0$,*[1] *and let Q^o be the affine quadric over k given by the equation*

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2 = 1,$$

with $a_i \in k^\times$. Suppose that n is odd, or that $n > 0$ is even and $Q^o(k) \neq \varnothing$. Then the tangent bundle T_{Q^o} has a nonvanishing section.

Proof. By Theorem 4.1 it is sufficient to show that for an even $n > 0$ one has $-1 \in [D(\langle a_1, a_2, \ldots, a_{n+1}\rangle)]$. We claim that already

$$-1 \in [D(\langle a_1, a_2, a_3\rangle)^2] \subseteq [D(\langle a_1, a_2, \ldots, a_{n+1}\rangle)^2] = [D(\langle a_1, a_2, \ldots, a_{n+1}\rangle)].$$

Indeed, note that the quadratic form $\langle a_1, a_2, a_3\rangle$ is a Pfister neighbor since

$$\langle a_1\rangle \cdot \langle a_1, a_2, a_3\rangle = \langle 1, a_1a_2, a_1a_3\rangle$$

is a subform of the 2-fold Pfister form $\langle\langle -a_1a_2, -a_1a_3\rangle\rangle = \langle 1, a_1a_2, a_1a_3, a_2a_3\rangle$. Then Lemma 3.15 yields

$$[D(\langle a_1, a_2, a_3\rangle)^2] = D(\langle 1, a_1a_2, a_1a_3, a_2a_3\rangle). \tag{4.6}$$

Now, the form $\langle 1, a_1a_2, a_1a_3, a_2a_3, 1\rangle$ is again a Pfister neighbor with the associated 3-fold Pfister form $\langle\langle -a_1a_2, -a_1a_3, -1\rangle\rangle$. The latter form is hyperbolic by the assumption, whence its subform $\langle 1, a_1a_2, a_1a_3, a_2a_3, 1\rangle$ is isotropic by the same dimension count argument as in the proof of Lemma 3.15. It follows that the equation

$$x_1^2 + a_1a_2x_2^2 + a_1a_3x_3^2 + a_2a_3x_4^2 + x_5^2 = 0$$

has a solution over k. This means that

$$-1 \in D(\langle 1, a_1a_2, a_1a_3, a_2a_3\rangle),$$

yielding by (4.6) that $-1 \in [D(\langle a_1, a_2, a_3\rangle)^2]$ and thereby the claim. □

Corollary 4.7. *Let k be a number field and let Q^o be the affine quadric over k given by the equation*

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2 = 1,$$

with $a_i \in k^\times$. Suppose $Q^o(k) \neq \varnothing$.

(1) *If n is odd, then T_{Q^o} has a nonvanishing section.*

(2) *If $n > 0$ is even, then T_{Q^o} has a nonvanishing section if and only if, for each real embedding $\sigma : k \to \mathbb{R}$, $\sigma(a_i) < 0$ for some i.*

[1] It follows from the Milnor conjecture [Voevodsky 2003, Corollary 7.5; Orlov et al. 2007, Theorem 4.1; Röndigs and Østvær 2016, Theorem 1.1] that $I(k)^3 = 0$ is equivalent to k being of 2-cohomological dimension at most 2.

Proof. The case of odd n follows from Theorem 4.1 (1).

For even $n > 0$, put $q = a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2$. Let $\sigma : k \to \mathbb{R}$ be an embedding such that $\sigma(a_i) > 0$ for all i and let $q_{\mathbb{R}}$ be q extended to $\mathbb{R}$ using this embedding. Then $[D(q_{\mathbb{R}})] \subseteq \mathbb{R}_{>0}$, whence $-1 \notin [D(q)]$. Thus Theorem 4.1 (3) yields one direction of the desired implication.

For the other direction it suffices to show that if $n > 0$ is even and for every embedding $\sigma : k \to \mathbb{R}$ one has $\sigma(a_i) < 0$ for some i then $-1 \in [D(q)]$. Note that the assumption that $Q^o(k) \neq \varnothing$ implies that, for each real embedding σ of k, there is a j with $\sigma(a_j) > 0$.

First assume $n \geq 4$. Let v be a place of k and consider the quadratic form $q + x_{n+2}^2 = a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2 + x_{n+2}^2$ over k_v. If v is a finite place, then $q + x_{n+2}^2$ is isotropic since every quadratic form of dimension ≥ 5 is isotropic over a local field [Lam 2005, Theorem VI.2.12]. If v is an infinite place then the assumption $\sigma(a_i) < 0$ for some i implies that $q + x_{n+2}^2$ is isotropic. Then [Lam 2005, Hasse–Minkowski Principle VI.3.1] implies that $q + x_{n+2}^2$ is isotropic over k, whence $-1 \in D(q)$.

Now assume $n = 2$. The form $\langle a_1, a_2, a_3\rangle = a_1x_1^2 + a_2x^2 + a_3x^2$ is a Pfister neighbor with the associated Pfister form $\langle 1, a_1a_2, a_1a_3, a_2a_3\rangle$ and Lemma 3.15 yields

$$[D(\langle a_1, a_2, a_3\rangle)^2] = D(\langle 1, a_1a_2, a_1a_3, a_2a_3\rangle). \tag{4.8}$$

Let v be a place of k and consider the form $q' = \langle 1, a_1a_2, a_1a_3, a_2a_3, 1\rangle$ over k_v. As above, if v is a finite place then the form q' is isotropic by [Lam 2005, Theorem VI.2.12]. If v is a complex place then q' is clearly isotropic. If v is a real place with the real embedding $\sigma_v : k \to \mathbb{R}$ then as $Q^o(k) \neq \varnothing$, we have $Q^o(k_v) \neq \varnothing$; hence there is a j such that $\sigma(a_j) > 0$. Combined with our assumption that $\sigma_v(a_i) < 0$ for some i, we see that at least one of $\sigma_v(a_1a_2)$, $\sigma_v(a_1a_3)$ and $\sigma_v(a_2a_3)$ is negative, whence q' is isotropic over k_v. Then [Lam 2005, Hasse–Minkowski Principle VI.3.1] implies that q' is isotropic over k meaning that $-1 \in D(\langle 1, a_1a_2, a_1a_3, a_2a_3\rangle)$. By (4.8), we thus have $-1 \in [D(\langle a_1, a_2, a_3\rangle)^2]$ and the claim follows. □

Corollary 4.9. *Let k be a field of one of the following types*:

(1) *a finite field $\mathbb{F}_{p^n}$, $p > 2$,*

(2) *a non-Archimedean local field of characteristic zero,*

(3) *the perfection of a local field of characteristic $p > 2$,*

(4) *the perfection of the function field of a curve over a finite field.*

Let Q^o be the affine quadric over k given by the equation

$$q := a_1x_1^2 + a_2x_2^2 + \cdots + a_{n+1}x_{n+1}^2 = 1,$$

with $a_i \in k^\times$. Suppose that $n > 0$, and if $n = 2$ and k is of type (2, 3, 4), *suppose in addition that Q^o has a k-rational point. Then T_{Q^o} has a nonvanishing section.*

Proof. In all the above cases, k is a perfect field of cohomological dimension ≤ 2, and the result follows from Corollary 4.5, once we know that $Q^o(k) \neq \varnothing$ if $n \geq 2$ is even. For k of type (1) every regular quadratic form in at least three variables is isotropic [Lam 2005, Proposition I.3.4] and for k of type (2, 3, 4), every regular quadratic form in at least five variables is isotropic [Lam 2005, Theorem VI.2.12, Corollary VI.3.5]; applying this to $q - x_0^2$ shows that $Q^o(k) \neq \varnothing$ in all cases to be considered. □

Acknowledgements

Ananyevskiy is supported by the DFG research grant AN 1545/4-1 and DFG Heisenberg grant AN 1545/1-1. He would like to thank Fabien Morel who drew his attention to the problem of the existence of nonvanishing vector fields on a 2-sphere over the dyadic numbers. Fabien Morel in turn learned about this problem from Umberto Zannier some time ago. This paper grew out of this particular question during Ananyevskiy's research visit to the University of Duisburg-Essen.

Levine is supported by the ERC grant QUADAG. He would like to thank Christopher Deninger, who passed on the problem of the existence of nonvanishing vector fields on a 2-sphere over the dyadic numbers from Umberto Zannier in 2020, as well as Umberto Zannier himself for subsequent discussions. We would like to thank the referee for their very helpful comments and suggestions, and we thank Peter Müller for kindly allowing us to present his example here.

References

[Asok and Fasel 2015] A. Asok and J. Fasel, "Splitting vector bundles outside the stable range and $\mathbb{A}^1$-homotopy sheaves of punctured affine spaces", *J. Amer. Math. Soc.* **28**:4 (2015), 1031–1062. MR

[Asok and Fasel 2016] A. Asok and J. Fasel, "Comparing Euler classes", *Q. J. Math.* **67**:4 (2016), 603–635. MR

[Asok and Fasel 2023] A. Asok and J. Fasel, "Vector bundles on algebraic varieties", pp. 2146–2170 in *International Congress of Mathematicians*, *III* (Helsinki, 2022), edited by D. Beliaev and S. Smirnov, Eur. Math. Soc., Berlin, 2023. MR

[Asok et al. 2017] A. Asok, M. Hoyois, and M. Wendt, "Affine representability results in $\mathbb{A}^1$-homotopy theory, I: Vector bundles", *Duke Math. J.* **166**:10 (2017), 1923–1953. MR

[Barge and Morel 2000] J. Barge and F. Morel, "Groupe de Chow des cycles orientés et classe d'Euler des fibrés vectoriels", *C. R. Acad. Sci. Paris Sér. I Math.* **330**:4 (2000), 287–290. MR

[Bhatwadekar et al. 2014] S. M. Bhatwadekar, J. Fasel, and S. Sane, "Euler class groups and 2-torsion elements", *J. Pure Appl. Algebra* **218**:1 (2014), 112–120. MR

[Colliot-Thélène and Skorobogatov 1993] J.-L. Colliot-Thélène and A. N. Skorobogatov, "Groupe de Chow des zéro-cycles sur les fibrés en quadriques", *K-Theory* **7**:5 (1993), 477–500. MR

[Déglise 2023] F. Déglise, "Notes on Milnor–Witt K-theory", preprint, 2023. arXiv 2305.18609

[Déglise et al. 2021] F. Déglise, F. Jin, and A. A. Khan, "Fundamental classes in motivic homotopy theory", *J. Eur. Math. Soc.* **23**:12 (2021), 3935–3993. MR

[tom Dieck 2008] T. tom Dieck, *Algebraic topology*, Eur. Math. Soc., Zürich, 2008. MR

[Fasel 2013] J. Fasel, "The projective bundle theorem for $\mathbf{I}^j$-cohomology", *J. K-Theory* **11**:2 (2013), 413–464. MR

[Fasel 2020] J. Fasel, "Lectures on Chow–Witt groups", pp. 83–121 in *Motivic homotopy theory and refined enumerative geometry* (Essen, Germany, 2018), edited by F. Binda et al., Contemp. Math. **745**, Amer. Math. Soc., Providence, RI, 2020. MR

[Feld 2022] N. Feld, "Birational invariance of the Chow–Witt group of zero-cycles", preprint, 2022. arXiv 2210.03995

[Hoyois 2014] M. Hoyois, "A quadratic refinement of the Grothendieck–Lefschetz–Verdier trace formula", *Algebr. Geom. Topol.* **14**:6 (2014), 3603–3658. MR

[Lam 2005] T. Y. Lam, *Introduction to quadratic forms over fields*, Grad. Stud. in Math. **67**, Amer. Math. Soc., Providence, RI, 2005. MR

[Levine 2020] M. Levine, "Aspects of enumerative geometry with quadratic forms", *Doc. Math.* **25** (2020), 2179–2239. MR

[Levine and Raksit 2020] M. Levine and A. Raksit, "Motivic Gauss–Bonnet formulas", *Algebra Number Theory* **14**:7 (2020), 1801–1851. MR

[Morel 2004] F. Morel, "An introduction to $\mathbb{A}^1$-homotopy theory", pp. 357–441 in *Contemporary developments in algebraic K-theory* (Trieste, Italy, 2002), edited by M. Karoubi et al., ICTP Lect. Notes **15**, Abdus Salam Int. Cent. Theoret. Phys., Trieste, Italy, 2004. MR

[Morel 2012] F. Morel, *$\mathbb{A}^1$-algebraic topology over a field*, Lecture Notes in Math. **2052**, Springer, 2012. MR

[Ojanguren et al. 1986] M. Ojanguren, R. Parimala, and R. Sridharan, "Symplectic bundles over affine surfaces", *Comment. Math. Helv.* **61**:3 (1986), 491–500. MR

[Orlov et al. 2007] D. Orlov, A. Vishik, and V. Voevodsky, "An exact sequence for $K_*^M/2$ with applications to quadratic forms", *Ann. of Math.* (2) **165**:1 (2007), 1–13. MR

[Röndigs and Østvær 2016] O. Röndigs and P. A. Østvær, "Slices of Hermitian K-theory and Milnor's conjecture on quadratic forms", *Geom. Topol.* **20**:2 (2016), 1157–1212. MR

[Scharlau 1969] W. Scharlau, "Zur Pfisterschen Theorie der quadratischen Formen", *Invent. Math.* **6** (1969), 327–328. MR

[Scharlau 1985] W. Scharlau, *Quadratic and Hermitian forms*, Grundl. Math. Wissen. **270**, Springer, 1985. MR

[Serre 1962] J.-P. Serre, *Corps locaux*, Publ. Inst. Math. Univ. Nancago **8**, Hermann, Paris, 1962. MR

[Voevodsky 2003] V. Voevodsky, "Motivic cohomology with $\mathbb{Z}/2$-coefficients", *Publ. Math. Inst. Hautes Études Sci.* **98** (2003), 59–104. MR

Communicated by Jean-Louis Colliot-Thélène
Received 2024-02-11 Revised 2024-07-05 Accepted 2024-10-18

alseang@gmail.com Mathematisches Institut, LMU München, München, Germany

marc.levine@uni-due.de Fakultät Mathematik, Universität Duisburg-Essen, Essen, Germany

ALGEBRA AND NUMBER THEORY 19:12 (2025)
https://doi.org/10.2140/ant.2025.19.2433

Self-correlations of Hurwitz class numbers

Alexander Walker

The asymptotic study of class numbers of binary quadratic forms is a foundational problem in arithmetic statistics. Here, we investigate finer statistics of class numbers by studying their self-correlations under additive shifts. Specifically, we produce uniform asymptotics for the shifted convolution sum $\sum_{n<X} H(n)H(n+\ell)$ for fixed $\ell \in \mathbb{Z}$, in which $H(n)$ denotes the Hurwitz class number.

1. Introduction

The study of class numbers of binary quadratic forms has a rich history, dating back to Lagrange and Gauss. In *Disquisitiones arithmeticae*, Gauss made several conjectures about the distribution of class numbers, including the famous statement that the class number $h(-D)$ of binary quadratic forms of discriminant $-D$ should diverge to infinity as $D \to \infty$. Gauss' conjecture was established by Heilbronn [12], with effective lower bounds first obtained through the combined work of Goldfeld [7] and Gross and Zagier [11].

Moment estimates for class numbers have been studied by many authors, often using Dirichlet's class number to reduce the problem to estimates for families of quadratic Dirichlet L-functions at the special point 1. For example, Wolke [34] proved that

$$\sum_{n \leq X} \tilde{h}(-n)^{\alpha} = c(\alpha) X^{1+\frac{\alpha}{2}} + O_{\alpha}\big(X^{1+\frac{\alpha}{2}-\frac{1}{4}}\big) \tag{1-1}$$

for fixed $\alpha > 0$, where $\tilde{h}(-n)$ denotes the number of classes of *primitive* binary quadratic forms of discriminant $-n$. Later work of Granville and Soundararajan [10] implies that the main term in (1-1) holds with some uniform error for any $\alpha \ll \log X$.

In comparison, shifted convolution estimates for class numbers are far less understood. Recent work of Kumaraswamy [23] considers

$$D(X,\ell) := \sum_{n \leq X}{}^{\flat}\, h(-n)h(-n-\ell),$$

in which $\sum^{\flat}$ denotes restriction to n such that both $-n$ and $-n-\ell$ are fundamental discriminants, with neither congruent to $1 \bmod 8$. Kumaraswamy applies the circle method to prove that

$$D(X,\ell) = c_{\ell} X^{\frac{3}{2}} (X+\ell)^{\frac{1}{2}} + O_{\epsilon}\big(X^{\frac{3}{2}-\frac{1}{30}} (X+\ell)^{\frac{1}{2}+\frac{1}{60}+\epsilon}\big)$$

MSC2020: primary 11E41; secondary 11F72, 11M99.
Keywords: Hurwitz class number, shifted convolution, harmonic Maass form, mock modular form.

for $\ell \geq 1$ and all $\epsilon > 0$, uniformly in ℓ (cf. [23, Theorem 1.1]). For fixed ℓ, this gives a power-saving of $O(X^{1/60-\epsilon})$ in the error term.

Unfortunately, the peculiar restriction to $n, n+\ell \not\equiv 1 \bmod 8$ in [23] is essential, as this work uses the identity

$$r_3(n) = 12\Big(1 - \Big(\frac{-n}{2}\Big)\Big)h(-n), \tag{1-2}$$

(cf. [3, Proposition 5.3.10]) to relate the class number to the Kronecker symbol and $r_3(n)$, the number of representations of n as a sum of 3 squares, which holds when $-n$ is a fundamental discriminant. Since $\big(\frac{-n}{2}\big) = 1$ for $n \equiv \pm 1 \bmod 8$, the identity (1-2) gives no information about $h(-n)$ on the residue class $-n \equiv 1 \bmod 8$.

This article presents an alternative method for studying correlations of class numbers, via the spectral theory of automorphic forms. In this setting, it is convenient to consider a version of the class number $h(-n)$ called the Hurwitz class number $H(n)$, in which the classes containing a multiple of x^2+y^2 or x^2-xy+y^2 are weighted by $\frac{1}{2}$ and $\frac{1}{3}$, respectively. By convention, we set $H(0) = -\frac{1}{12}$. Hurwitz class numbers feature, for one example, in Eichler–Selberg "class number relation" formulas, such as

$$\sum_{m \in \mathbb{Z}} H(4n-m^2) = 2\sigma_1(n) + \sum_{d \mid n} \min\Big(d, \frac{n}{d}\Big), \tag{1-3}$$

which appear in the work of Kronecker and Hurwitz. Here, $\sigma_\nu(n) = \sum_{d \mid n} d^\nu$.

More recently, Zagier [36] showed that Hurwitz class numbers arise as the coefficients of a mock modular form. Specifically, Zagier proved that

$$\mathcal{H}(z) := \sum_{n \geq 0} H(n)e(nz) + \frac{1}{8\pi\sqrt{y}} + \sum_{n \geq 1} \frac{n\,\Gamma\big(-\frac{1}{2}, 4\pi n^2 y\big)}{4\sqrt{\pi}} e(-n^2 z) \tag{1-4}$$

defines a harmonic Maass form of weight $\frac{3}{2}$ on $\Gamma_0(4)$. Here, $z = x+iy$,

$$e(z) = e^{2\pi i n z},$$

and $\Gamma(\beta, y)$ denotes the incomplete gamma function. In particular, one may study Hurwitz class numbers using automorphic forms.

In this article, we leverage the analytic theory of harmonic Maass forms and mock modular forms to study the shifted convolution Dirichlet series

$$D_\ell(s) := \sum_{n \geq 1} \frac{H(n)H(n+\ell)}{(n+\ell)^{s+\frac{1}{2}}}, \tag{1-5}$$

where $\ell \geq 1$ is a fixed integer. We prove that $D_\ell(s)$ admits meromorphic continuation to $s \in \mathbb{C}$ and use this information to study the self-correlations of Hurwitz class numbers under additive shifts. Our main theorem is the following result.

Theorem 1.1. *Let $\sigma_\nu(m) = \sum_{d \mid m} d^\nu$ denote the sum-of-divisors function, with the convention that $\sigma_\nu(m) = 0$ for $m \notin \mathbb{Z}$. Fix $\ell \geq 1$ and let ℓ_o denote the odd part of ℓ. Then, for all $\epsilon > 0$, we have*

$$\sum_{n \leq X} H(n)H(n+\ell) = \frac{\pi^2 X^2}{252\,\zeta(3)}\big(2\sigma_{-2}\big(\tfrac{\ell}{4}\big) - \sigma_{-2}\big(\tfrac{\ell}{2}\big) + \sigma_{-2}(\ell_o)\big) + O_\epsilon\big(X^{\frac{5}{3}+\epsilon} + X^{1+\epsilon}\ell\big).$$

For $\ell \ll X^{2/3}$, this result achieves a uniform error of size $O_\epsilon(X^{5/3+\epsilon})$. For larger ℓ, the error term depends on ℓ but remains nontrivial when $\ell \ll X^{1-\epsilon}$.

Since Hurwitz class numbers agree with the ordinary class numbers $h(-n)$ for n not of the form $3m^2$ or $4m^2$, the rough upper bound $H(n) \ll n^{1/2+o(1)}$ (cf. Lemma 7.2) implies the following result as an immediate corollary.

Corollary 1.2. *With notation as above, we have*

$$\sum_{n \leq X} h(-n-\ell)h(-n) = \frac{\pi^2 X^2}{252\,\zeta(3)}\big(2\sigma_{-2}\big(\tfrac{\ell}{4}\big) - \sigma_{-2}\big(\tfrac{\ell}{2}\big) + \sigma_{-2}(\ell_o)\big) + O_\epsilon\big(X^{\frac{5}{3}+\epsilon} + X^{1+\epsilon}\ell^{1+\epsilon}\big).$$

The error bounds in Theorem 1.1 are of course not sharp. We conjecture that Theorem 1.1 should hold with a secondary main term and an error of size $O_\epsilon((X\ell)^{1+\epsilon})$; specifically, that

$$\begin{aligned}\sum_{n \leq X} H(n)H(n+\ell) = {} & \frac{\pi^2 X^2}{252\,\zeta(3)}\big(2\sigma_{-2}\big(\tfrac{\ell}{4}\big) - \sigma_{-2}\big(\tfrac{\ell}{2}\big) + \sigma_{-2}(\ell_o)\big) \\ & - \frac{2X^{\frac{3}{2}}}{9\pi}\big(2\sigma_{-1}\big(\tfrac{\ell}{4}\big) - \sigma_{-1}\big(\tfrac{\ell}{2}\big) + \sigma_{-1}(\ell_o)\big) + O_\epsilon\big((X\ell)^{1+\epsilon}\big). \end{aligned} \tag{1-6}$$

To support this conjecture, we show (cf. Remark 10.1) that (1-6) holds when the cutoff $n \leq X$ is replaced by a certain class of truncations with smoothing.

Paper methodology and outline

To produce shifted convolution estimates that treat all congruence classes equally, we abandon (1-2) in favor of the generating function $\mathcal{H}(z)$ from (1-4). In particular, we treat shifted convolutions involving weak harmonic Maass forms instead of ordinary modular forms. We also depart from [23] in that we treat shifted convolutions using the spectral theory of automorphic forms, as opposed to the circle method.

Following some background material on harmonic weak Maass forms and mock modular forms in Section 2, we relate the Dirichlet series $D_\ell(s)$ defined in (1-5) to the Petersson inner product $\langle y^{3/2}|\mathcal{H}|^2, P_\ell(\cdot,\bar{s})\rangle$, in which $P_\ell(z,s)$ is a particular Poincaré series.

We obtain a meromorphic continuation for $D_\ell(s)$ by first producing a meromorphic continuation of $\langle y^{3/2}|\mathcal{H}|^2, P_\ell(\cdot,\bar{s})\rangle$. This task is complicated by the fact that $F(z) := y^{3/2}|\mathcal{H}(z)|^2$ is not square-integrable. To address this, we show in Section 4 that $F(z)$ may be written in the form $\mathcal{V}(z) + \mathcal{E}(z)$, in which $\mathcal{V} \in L^2$ and $\mathcal{E}$ is an explicit function involving Eisenstein series and the Jacobi theta function.

The meromorphic continuations of $\langle \mathcal{E}, P_\ell(\cdot,\bar{s})\rangle$ and $\langle \mathcal{V}, P_\ell(\cdot,\bar{s})\rangle$ are then computed in Sections 5 and 6, respectively. While $\langle \mathcal{E}, P_\ell(\cdot,\bar{s})\rangle$ can be understood directly, the meromorphic continuation of $\langle \mathcal{V}, P_\ell(\cdot,\bar{s})\rangle$ is accomplished through spectral expansion of the Poincaré series.

The methods described up to this point apply more generally. To illustrate this, the major results of Sections 3–6 are presented with $\mathcal{H}$ replaced by a generic weak harmonic Maass form "of polynomial growth" (cf. Section 2B). Our first significant specialization to $\mathcal{H}(z)$ occurs in Section 6, where we leverage the fact that the contribution from the nonholomorphic part of $\mathcal{H}(z)$ is unusually simple (cf. Remark 6.4) to more easily classify the poles and residues of $D_\ell(s)$ in the right half-plane $\operatorname{Re} s > \frac{1}{2}$.

Our main application, Theorem 1.1, also requires uniform bounds for the growth of $D_\ell(s)$ in vertical strips. In Section 7, we address various elementary terms to reduce this problem to growth estimates for $\langle \mathcal{V}, P_\ell(\cdot,\bar{s})\rangle$.

The spectral expansion of $P_\ell(z,s)$ gives a decomposition $\langle \mathcal{V}, P_\ell(\cdot,\bar{s})\rangle = \Sigma_{\mathrm{disc}}(s) + \Sigma_{\mathrm{cont}}(s)$ corresponding to contributions from the discrete and continuous spectra of the hyperbolic Laplacian. While Σ_{cont} is readily handled, the problem of bounding $\Sigma_{\mathrm{disc}}(s)$ with respect to $|\operatorname{Im} s|$ is particularly complicated and represents the central difficulty of this work.

Ultimately, our bounds for Σ_{disc} rely on decay estimates for triple inner products of the form $\langle y^{3/2}|\mathcal{H}|^2, \mu_j\rangle$, in which $\mu_j(z)$ runs through an orthonormal basis for Hecke–Maass cusp forms on $\Gamma_0(4)$. Similar inner products, of the form $\langle y^k \phi_1 \overline{\phi_2}, \mu_j\rangle$ (with ϕ_1, ϕ_2 automorphic forms of weight k) have been studied in numerous works, and we mention a few:

a. ϕ_1, ϕ_2 weight $k \in \mathbb{Z}$ holomorphic cusp forms on $\Gamma_0(N)$, by [8];

b. ϕ_1, ϕ_2 weight 0 Eisenstein series on $\Gamma_0(1)$, by [33];

c. $\phi_1 \overline{\phi_2}$ replaced by any polynomial in Maass cusp forms, by [28];

d. ϕ_1, ϕ_2 weight 0 Maass cusp forms on $\Gamma_0(1)$, by [19; 20];

e. ϕ_1, ϕ_2 weight $k \in \frac{1}{2}\mathbb{Z}$ modular forms on $\Gamma_0(N)$, by [22].

Of these prior works, (a) and (b) use the Rankin–Selberg method directly, (c) and (e) use the automorphic kernel, and (d) uses a modified Rankin–Selberg method that introduces an auxiliary Eisenstein series for the express purpose of unfolding.

Our treatment of $\langle y^{3/2}|\mathcal{H}|^2, \mu_j\rangle$ appears in Section 8. More generally, this section produces bounds for triple inner products of the form $\langle y^k |f|^2, \mu_j\rangle$, where f is a harmonic Maass form of polynomial growth of weight $k \in \frac{1}{2} + \mathbb{Z}$. In particular, we prove the following result:

Theorem 1.3. *Let f be a harmonic Maass form of polynomial growth of weight $k \in \frac{1}{2} + \mathbb{Z}$ and level N. Let μ be an L^2-normalized Hecke–Maass cusp form of weight* 0 *on $\Gamma_0(N)$, with spectral type $t \in \mathbb{R}$. For all $\epsilon > 0$, we have*

$$\langle y^k |f|^2, \mu\rangle \ll_{N,\epsilon} \left(|t|^{2k-1} + |t|^{3-2k}\right)|t|^{\epsilon} e^{-\frac{\pi}{2}|t|}.$$

We remark that the space of harmonic Maass forms of polynomial growth includes $M_k(\Gamma_0(N))$, the space of modular forms. In this setting, Theorem 1.3 can be used to improve the spectral dependence in certain results of [22]. (In particular, see [22, Proposition 14].)

Our proof of Theorem 1.3 draws heavy inspiration from [19; 20], though our work is more complicated in several respects, such as the change from $\Gamma_0(1)$ to $\Gamma_0(N)$, the change in Whittaker functions (from K-Bessel functions to incomplete gamma functions), the generalization to half-integral weight, and the introduction of terms related to the fact that f need not be cuspidal. We also depart from Jutila by considering individual inner products instead of spectral large sieve inequalities. We suspect that a spectral large sieve inequality would not improve Theorem 1.1.

In Section 9, we apply these triple product estimates to complete our quantification of the growth of $D_\ell(s)$. At this point, our main result follows from a version of Perron's formula with truncation, as presented in Section 10.

2. Harmonic weak Maass forms and mock modular forms

The theory of harmonic Maass forms was introduced by Bruinier and Funke in the context of geometric theta lifts [2]. This section reviews the basic definitions of harmonic Maass forms and mock modular forms. A good reference for background material is [1, §4].

A weak Maass form of weight k on a congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a smooth function $f : \mathfrak{h} \to \mathbb{C}$ which transforms like a modular form of weight k, is an eigenfunction of the weight k Laplacian

$$\Delta_k := -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + iky\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right),$$

and has at most linear exponential growth at cusps.

If $\Delta_k f = 0$, then f is called a harmonic (weak) Maass form of manageable growth. Let $H_k^!(\Gamma)$ denote the space of weight k harmonic Maass forms of manageable growth on Γ. If $\Gamma = \Gamma_0(N)$ or $\Gamma_1(N)$, then any $f(x+iy) \in H_k^!(\Gamma)$ admits a Fourier expansion at ∞ of the form

$$f(z) = \sum_{n \geq n^+} c^+(n)e(nz) + c^-(0)y^{1-k} + \sum_{\substack{n \geq n^- \\ n \neq 0}} c^-(n)\Gamma(1-k, 4\pi ny)e(-nz) \tag{2-1}$$

(cf. [1, Lemma 4.3]), where $\Gamma(\beta, y) := \int_y^\infty t^{\beta-1}e^{-t}\,dt$ is the incomplete gamma function. In the case $k = 1$, the term $c^-(0)y^{1-k}$ is replaced with $c^-(0)\log y$. The first sum in the Fourier expansion (2-1) of $f(z)$ is called the holomorphic part, and the rest of the right-hand side of (2-1) is the nonholomorphic part. Any function which arises as the holomorphic part of a harmonic Maass form of manageable growth is called a mock modular form.

Fourier expansions of analogous shape exist for each cusp of Γ. To describe this precisely, we assume henceforth that $k \in \frac{1}{2}\mathbb{Z}$ and $\Gamma \subset \Gamma_0(4)$ and restrict to Maass forms with the theta multiplier system υ_θ.

That is, where $\theta(z) := \sum_{n\in\mathbb{Z}} e(n^2 z)$ denotes the Jacobi theta function, we assume that

$$f(\gamma z) = \left(\frac{\theta(\gamma z)}{\theta(z)}\right)^{2k} f(z)$$

for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$. This may also be written $f(\gamma z) = \upsilon_\theta(\gamma)^{2k}(cz+d)^k f(z)$, in which υ_θ is defined by $\upsilon_\theta\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right) := \epsilon_d^{-1}\left(\frac{c}{d}\right)$, where $\epsilon_d = 1$ for $d \equiv 1 \bmod 4$ and $\epsilon_d = i$ for $d \equiv 3 \bmod 4$. For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}(2,\mathbb{R})$ with $\det\gamma > 0$, we define the weight k slash operator by

$$f|_\gamma(z) := \left(\frac{\theta(\gamma z)}{\theta(z)}\right)^{-2k} f(\gamma z).$$

Finally, for each cusp $\mathfrak{a}$ of Γ, let $\Gamma_\mathfrak{a} = \langle \pm t_\mathfrak{a}\rangle \subset \Gamma$ denote the stabilizer of $\mathfrak{a}$. Let $\sigma_\mathfrak{a}$ denote a scaling matrix for $\mathfrak{a}$, i.e., a matrix in $\mathrm{GL}(2,\mathbb{R})$ for which $t_\mathfrak{a} = \sigma_\mathfrak{a}\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\sigma_\mathfrak{a}^{-1}$. Define the cusp parameter $\varkappa_\mathfrak{a} \in [0,1)$ so that $e(\varkappa_\mathfrak{a}) = \upsilon_\theta(t_\mathfrak{a})$. If $\varkappa_\mathfrak{a} = 0$, the cusp $\mathfrak{a}$ is called singular; otherwise, $\mathfrak{a}$ is called nonsingular.

Given all this notation, $f(z)$ admits a Fourier expansion at each cusp $\mathfrak{a}$ of Γ, given by

$$\begin{aligned} f_\mathfrak{a}(z) := f|_{\sigma_\mathfrak{a}}(z) = \sum_{n\geq n^+} c_\mathfrak{a}^+(n) e((n+\varkappa_\mathfrak{a})z) \\ + c_\mathfrak{a}^-(0) y^{1-k} + \sum_{\substack{n\geq n^- \\ n\neq\varkappa_\mathfrak{a}}} c_\mathfrak{a}^-(n)\Gamma(1-k, 4\pi(n-\varkappa_\mathfrak{a})y) e(-(n-\varkappa_\mathfrak{a})z), \end{aligned} \tag{2-2}$$

where $c_\mathfrak{a}^-(0)y^{1-k}$ appears only when $\varkappa_\mathfrak{a} = 0$. When $k = 1$ and $\varkappa_\mathfrak{a} = 0$, we replace this term by $c_\mathfrak{a}^-(0)\log y$. Since we work most commonly with the Fourier expansion at $\mathfrak{a} = \infty$, we retain the shorthand $c^\pm(n) := c_\infty^\pm(n)$.

2A. *The shadow operator ξ_k.* This section follows [1, §5.1]. Recall the Maass lowering operator L_k defined by $L_k = -iy^2\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)$. We define as well the shadow operator $\xi_k = y^{k-2}\overline{L_k}$. By [1, Theorem 5.10], ξ_k maps $H_k^!(\Gamma_0(N))$ surjectively to $M_{2-k}^!(\Gamma_0(N))$, the space of weakly holomorphic modular forms of weight $2-k$. This map is given by

$$\xi_k(f(z)) = (1-k)\overline{c^-(0)} - (4\pi)^{1-k}\sum_{\substack{n\geq n^- \\ n\neq 0}} \overline{c^-(n)} n^{1-k} e(nz). \tag{2-3}$$

The form $\xi_k f$ is called the shadow of f.

2B. *Harmonic Maass forms of polynomial growth.* Generically, the coefficient series $\{c_\mathfrak{a}^\pm(n)\}$ grow superpolynomially as $n \to \infty$. In the remainder of this article, we restrict to the special case in which the coefficients are polynomially bounded in n. This is equivalent to the property that $f(z)$ have no poles at cusps, or that $n^\pm \pm \varkappa_\mathfrak{a} \geq 0$ in (2-1) for all $\mathfrak{a}$.

Let $H_k^\sharp(\Gamma_0(N))$ denote the subspace of $H_k^!(\Gamma_0(N))$ consisting of forms with at most polynomial growth at cusps. We remark that the space $H_k^\sharp$ features prominently in [31], where it serves as a natural setting to study L-functions attached to mock modular forms. Note that $H_k^\sharp$ is a subspace of the space of (not necessarily cuspidal) Maass wave forms of weight k.

The shadow operator maps $\xi_k : H_k^\sharp(\Gamma_0(N)) \to M_{2-k}(\Gamma_0(N))$. In particular, ξ_k annihilates $H_k^\sharp(\Gamma_0(N))$ for $k > 2$. In other words, $H_k^\sharp = M_k$ for $k > 2$, so the space $H_k^\sharp$ is most interesting for $k \le 2$.

Though exact growth rates for the coefficients $c_{\mathfrak{a}}^{\pm}(n)$ are not known, adequate on-average bounds are known from the Rankin–Selberg method as applied to Maass forms (including noncuspidal Maass forms) in [26]. Specializing to the case of harmonic Maass forms and translating notation, we present the following result.

Lemma 2.1 (cf. [26, Theorem 5.2]). *Fix $f(z) \in H_k^\sharp(\Gamma_0(N))$ with $k \in \frac{1}{2}\mathbb{Z}$ and $k \neq 1$. If f has Fourier expansion (2-2), then*

$$\sum_{n \le X} \frac{|c_{\mathfrak{a}}^{\pm}(n)|^2}{(n \pm \varkappa_{\mathfrak{a}})^{k-1}} = \begin{cases} c_{f,\mathfrak{a}}^{\pm} X + O_f\big(X^{\frac{3}{5}} \log X\big) & \text{if } f \text{ is cuspidal,} \\ c_{f,\mathfrak{a}}^{\pm} X^{1+|k-1|} + O_f\big(X^{1+|k-1|-\frac{2+4|k-1|}{5+8|k-1|}} \log X\big) & \text{else,} \end{cases}$$

for some constants $c_{f,\mathfrak{a}}^{\pm}$.

3. Shifted convolutions via inner products

In this section, we show that shifted convolution Dirichlet series of the form (1-5) can be recognized in terms of Petersson inner products. To begin, we treat a generic form $f(z) \in H_k^\sharp(\Gamma_0(N))$ with Fourier expansion (2-1). We define the ℓ-th Poincaré series $P_\ell(z,s)$ of weight 0 on $\Gamma_0(N)$ by

$$P_\ell(z,s) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \operatorname{Im}(\gamma z)^s e(\ell \gamma z).$$

For s with $\operatorname{Re} s$ sufficiently large, the Rankin–Selberg unfolding method gives

$$\begin{aligned} \langle y^k |f|^2, P_\ell(\cdot, \bar{s})\rangle &= \int_0^\infty \int_0^1 y^{s+k} |f(z)|^2 \overline{e(\ell z)} \frac{dx\,dy}{y^2} \\ &= \sum_{n_1 = n_2 + \ell} \int_0^\infty y^{s+k-1} c(n_1, y) \overline{c(n_2, y)} e^{-2\pi h y} \frac{dy}{y}, \end{aligned} \tag{3-1}$$

in which $c(n,y)$ denotes the n-th Fourier coefficient of $f(z)$ at the cusp $\mathfrak{a} = \infty$. In other words, $c(n,y) = c^+(n)e^{-2\pi n y}$ for $n \ge 1$, $c(0,y) = c^+(0) + c^-(0)y^{1-k}$, and $c(n,y) = c^-(-n)\Gamma(1-k, -4\pi n y)e^{-2\pi n y}$ for $n \le -1$.

The contribution of $n_1, n_2 > 0$ to the inner product is a standard shifted convolution Dirichlet series:

$$I_\ell^+(s) := \sum_{n_1 = n_2 + \ell} c^+(n_1)\overline{c^+(n_2)} \int_0^\infty y^{s+k-1} e^{-2\pi(n_1+n_2+\ell)y} \frac{dy}{y} = \frac{\Gamma(s+k-1)}{(4\pi)^{s+k-1}} \sum_{n_2=1}^\infty \frac{c^+(n_2+\ell)\overline{c^+(n_2)}}{(n_2+\ell)^{s+k-1}}.$$

Since $f \in H_k^\sharp$, the Dirichlet series in the line above converges absolutely in some right half-plane. More precisely, Lemma 2.1 gives convergence in $\operatorname{Re} s > 1 + |k-1|$, extending to $\operatorname{Re} s > 1$ in the cuspidal case.

The net contributions from $(n_1,n_2)=(\ell,0)$ and $(0,-\ell)$ total

$$I_\ell^0(s):= \frac{c^+(\ell)\overline{c^-(0)}\Gamma(s)}{(4\pi\ell)^s}+\frac{c^+(\ell)\overline{c^+(0)}\Gamma(s+k-1)}{(4\pi\ell)^{s+k-1}}+\frac{c^-(0)\overline{c^-(\ell)}\Gamma(s-k+1)}{(4\pi\ell)^s s}+\frac{c^+(0)\overline{c^-(\ell)}\Gamma(s)}{(4\pi\ell)^{s+k-1}(s+k-1)}.$$

The function $I_\ell^0(s)$ is meromorphic in $s\in\mathbb{C}$ and analytic in $\operatorname{Re} s>|k-1|$. Note that $I_\ell^0(s)$ vanishes when f is cuspidal.

There is another finite collection of "cross terms" when $n_1>0$ and $n_2<0$, which contributes

$$\begin{aligned} I_\ell^\times(s) &:= \sum_{m=1}^{\ell-1} c^+(\ell-m)\overline{c^-(m)}\int_0^\infty y^{s+k-1}e^{-4\pi(\ell-m)y}\Gamma(1-k,4\pi my)\frac{dy}{y} \\ &= \frac{\Gamma(s)}{s+k-1}\sum_{m=1}^{\ell-1}\frac{c^+(\ell-m)\overline{c^-(m)}}{(4\pi m)^{s+k-1}}\,{}_2F_1\left(\begin{matrix} s,s+k-1 \\ s+k\end{matrix}\,\middle|\, 1-\frac{\ell}{m}\right), \end{aligned}$$

in which we've evaluated the integral via [9, 6.455(1)]. The function $I_\ell^\times(s)$ is analytic in the right half-plane $\operatorname{Re} s>\max(0,1-k)$ and has an obvious meromorphic continuation to $s\in\mathbb{C}$.

Lastly, we record the contribution of $n_1,n_2<0$, which can be written

$$\begin{aligned} I_\ell^-(s) &:= \sum_{n=1}^\infty \frac{c^-(n)\overline{c^-(n+\ell)}}{(4\pi)^{s+k-1}}G_k(s,n,n+\ell), \quad \text{with} \\ G_k(s,n,n+\ell) &:= \int_0^\infty y^{s+k-1}\Gamma(1-k,ny)\Gamma(1-k,(n+\ell)y)e^{ny}\frac{dy}{y}. \end{aligned} \tag{3-2}$$

The two asymptotic expressions $\Gamma(\beta,y)=\Gamma(\beta)-y^\beta/\beta+O_\beta(y^{\beta+1})$ as $y\to 0$ and $\Gamma(\beta,y)=e^{-y}y^{\beta-1}(1+O_\beta(y^{-1}))$ as $y\to\infty$ imply that $G_k(s,n,n+\ell)$ converges absolutely when $\operatorname{Re} s>|k-1|$. In this region, $G_k(s,n,n+\ell)\ll G_k(\operatorname{Re} s,n,n)\ll_{\operatorname{Re} s} n^{-\operatorname{Re} s-k+1}$ by change of variable. Thus $I_\ell^-(s)$ converges to an analytic function in $\operatorname{Re} s>1+|k-1|$, extending to the domain $\operatorname{Re} s>1$ in the cuspidal case.

We conclude that the unfolding procedure is valid in $\operatorname{Re} s>1+|k-1|$, and that in this region we have the decomposition

$$\langle y^k|f|^2, P_\ell(\cdot,\bar{s})\rangle = I_\ell^+(s)+I_\ell^0(s)+I_\ell^\times(s)+I_\ell^-(s). \tag{3-3}$$

3A. ***Application to $\mathcal{H}(z)$.*** The formulas in this section simplify considerably for the specific form $\mathcal{H}\in H^\sharp_{3/2}(\Gamma_0(4))$ defined in (1-4). Recall from (1-5) the definition of the shifted convolution Dirichlet series

$$D_\ell(s):=\sum_{n\geq 1}\frac{H(n)H(n+\ell)}{(n+\ell)^{s+\frac12}}, \tag{3-4}$$

which converges absolutely in $\operatorname{Re} s>\frac32$. By (1-4), the coefficients $c^-(n)$ of $\mathcal{H}$ may be written in terms of $r_1(n)$, the number of representations of n as the square of an integer. Simplifying the various terms at right in (3-3) produces the formula

$$\begin{aligned}\langle y^{\frac{3}{2}}|\mathcal{H}|^2, P_\ell(\cdot,\bar{s})\rangle &= \frac{\Gamma\big(s+\frac{1}{2}\big)}{(4\pi)^{s+\frac{1}{2}}}D_\ell(s)+\frac{H(\ell)\Gamma(s)}{8\pi(4\pi\ell)^s}-\frac{H(\ell)\Gamma\big(s+\frac{1}{2}\big)}{12(4\pi\ell)^{s+\frac{1}{2}}}\\ &\quad+\frac{r_1(\ell)\Gamma\big(s-\frac{1}{2}\big)}{128\pi^2(4\pi\ell)^{s-\frac{1}{2}}s}-\frac{r_1(\ell)\Gamma(s)}{192\pi(4\pi\ell)^s\big(s+\frac{1}{2}\big)}\\ &\quad+\frac{\Gamma(s)}{s+\frac{1}{2}}\sum_{m=1}^{\ell-1}\frac{H(\ell-m)r_1(m)}{16\pi(4\pi m)^s}\,{}_2F_1\left(\begin{matrix}s,s+\frac{1}{2}\\ s+\frac{3}{2}\end{matrix}\middle|\,1-\frac{\ell}{m}\right)\\ &\quad+\sum_{\substack{m_1^2-m_2^2=\ell\\ m_1,m_2\geq 1}}\frac{m_1m_2}{4(4\pi)^{s+\frac{3}{2}}}G_{\frac{3}{2}}(s,m_2^2,m_1^2).\end{aligned}\tag{3-5}$$

The contribution of $I_\ell^-(s)$ in the fourth line of (3-5) is a finite sum, since $m_1^2-m_2^2=\ell$ has finitely many solutions. This phenomenon generalizes to any $f\in H^\sharp_{3/2}(\Gamma_0(N))$ for which $M_{1/2}(\Gamma_0(N))$ is one-dimensional, since in that case $\xi_{3/2}f$ is necessarily a twisted theta function by [30, Theorem A]. Thus, in departure from the general case, we conclude that $I_\ell^-(s)$ is analytic in $\operatorname{Re}s>\frac{1}{2}$.

Secondly, we remark that the contribution of $I_\ell^\times(s)$ bears some resemblance to one side of the Eichler–Selberg class number relation (cf. (1-3)). More specifically,

$$\operatorname*{Res}_{s=0}\frac{\Gamma(s)}{s+\frac{1}{2}}\sum_{m=1}^{\ell-1}\frac{H(\ell-m)r_1(m)}{16\pi(4\pi m)^s}\,{}_2F_1\left(\begin{matrix}s,s+\frac{1}{2}\\ s+\frac{3}{2}\end{matrix}\middle|\,1-\frac{\ell}{m}\right)=\frac{1}{4\pi}\sum_{m^2<\ell}H(\ell-m^2),$$

which is essentially one of the sums described in [1, §10.3]. It would be interesting to know if the methods in this paper could be used to produce new class number relations.

4. Automorphic regularization

To produce a meromorphic continuation for the Dirichlet series $D_\ell(s)$, we first show that the inner product $\langle y^k|\mathcal{H}|^2, P_\ell(\cdot,\bar{s})\rangle$ has a meromorphic continuation to a larger domain. This latter continuation involves the spectral decomposition of $P_\ell(z,s)$ with respect to the hyperbolic Laplacian and is complicated by the fact that $y^{3/2}|\mathcal{H}(z)|^2\notin L^2(\Gamma_0(4)\backslash\mathfrak{h})$. To rectify this, we modify $y^{3/2}|\mathcal{H}(z)|^2$ by subtracting a linear combination of automorphic forms chosen to neutralize growth at the cusps of $\Gamma_0(N)$.

We define the weight 0 Eisenstein series attached to cusp $\mathfrak{a}$ of $\Gamma_0(N)$ by

$$E_\mathfrak{a}(z,s)=\sum_{\gamma\in\Gamma_\mathfrak{a}\backslash\Gamma_0(N)}\operatorname{Im}(\sigma_\mathfrak{a}^{-1}\gamma z)^s.$$

These Eisenstein series have Fourier expansion at the cusp $\mathfrak{b}$ of the form

$$\begin{aligned}&E_\mathfrak{a}(\sigma_\mathfrak{b}z,w)\\ &=\delta_{[\mathfrak{a}=\mathfrak{b}]}y^w+\pi^{\frac{1}{2}}\frac{\Gamma\big(w-\frac{1}{2}\big)}{\Gamma(w)}\varphi_{\mathfrak{ab}0}(w)y^{1-w}+\frac{2\pi^wy^{\frac{1}{2}}}{\Gamma(w)}\sum_{n\neq 0}\varphi_{\mathfrak{ab}n}(w)|n|^{w-\frac{1}{2}}K_{w-\frac{1}{2}}(2\pi|n|y)e(nx),\end{aligned}\tag{4-1}$$

in which $\delta_{[\cdot]}$ denotes the Kronecker delta, $K_\nu(y)$ is the K-Bessel function, and the coefficients $\varphi_{\mathfrak{ab}n}(w)$ are described in [4], for example.

As in the previous section, we first consider a general $f(z) \in H_k^\sharp(\Gamma_0(N))$, with $k \in \frac{1}{2}\mathbb{Z}$ ($k \neq 1$), specializing to $f = \mathcal{H}$ when convenient. Let $F_\mathfrak{a}(z) := y^k |f|_{\sigma_\mathfrak{a}}(z)|^2 = \mathrm{Im}(\sigma_\mathfrak{a} z)^k |f(\sigma_\mathfrak{a} z)|^2$. If $\varkappa_\mathfrak{a} \neq 0$, then $F_\mathfrak{a}(z)$ decays exponentially as $y \to \infty$ by the Fourier expansion (2-2), and no regularization is required. Otherwise, when $\varkappa_\mathfrak{a} = 0$, (2-2) implies that

$$F_\mathfrak{a}(z) = y^k |c_\mathfrak{a}^+(0) + c_\mathfrak{a}^-(0) y^{1-k}|^2 + O(y^{-M}) \tag{4-2}$$

as $y \to \infty$ for all $M > 0$. It therefore suffices to regularize growth of sizes y^k, y^1, and y^{2-k} at the singular cusps.

For $k > 1$, the Eisenstein series $E_\mathfrak{a}(z, k)$ counteracts growth of size y^k at $\mathfrak{a}$, while for $k < 1$ we utilize $E_\mathfrak{a}(z, 2-k)$ to address y^{2-k}. Unfortunately, this technique fails to regularize growth of size y^1, since $E_\mathfrak{a}(z, w)$ has a pole at $w = 1$. In this case, we instead subtract a multiple of the constant term in the Laurent expansion of $E_\mathfrak{a}(z, w)$ at $w = 1$, which we denote $\widetilde{E}_\mathfrak{a}(z, 1)$, and which satisfies $\widetilde{E}_\mathfrak{a}(\sigma_\mathfrak{b} z, 1) = \delta_{[\mathfrak{a}=\mathfrak{b}]} y - \pi \log y \operatorname{Res}_{w=1} \varphi_{\mathfrak{ab}0}(w) + \tilde{c}_{\mathfrak{ab}} + O(e^{-2\pi y})$ for some constant $\tilde{c}_{\mathfrak{ab}}$. Thus, for example,

$$\mathcal{V}_f(z) := F(z) - \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} |c_\mathfrak{a}^-(0)|^2 E_\mathfrak{a}(z, 2-k) - 2 \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} \mathrm{Re}(c_\mathfrak{a}^+(0)\overline{c_\mathfrak{a}^-(0)}) \widetilde{E}_\mathfrak{a}(z, 1) \tag{4-3}$$

satisfies $\mathcal{V}_f(\sigma_\mathfrak{a} z) = O(y^k + \log y)$ as $y \to \infty$ when $k < 1$. If $k < \frac{1}{2}$, it follows that $\mathcal{V}_f \in L^2(\Gamma_0(N)\backslash\mathfrak{h})$. The case $k > \frac{3}{2}$ may be treated analogously.

The situation is more complicated in weights $k = \frac{1}{2}$ and $k = \frac{3}{2}$, as here we must regularize terms of size $y^{1/2}$. The obvious choice for regularizing $y^{1/2}$ is to subtract a multiple of $E_\mathfrak{a}(z, \frac{1}{2})$, but this term equals 0 since the completed Eisenstein series $\zeta^*(2w) E_\mathfrak{a}(z, w)$ is analytic at $w = \frac{1}{2}$. Likewise, it is not possible to regularize with a linear combination of terms of the form $\lim_{w \to \frac{1}{2}} \zeta^*(2w) E_\mathfrak{a}(z, w)$, as these grow as $y^{1/2} \log y$ near $\mathfrak{a}$.

In weight $k = \frac{3}{2}$, the growth of size $y^{1/2}$ comes from the nonholomorphic part (cf. (4-2)). In particular, we can regularize all cusp growth of size $y^{1/2}$ simultaneously by subtracting an appropriate multiple of $y^{1/2} |\xi_{3/2} f|^2$. Specifically, we define

$$\mathcal{V}_f(z) := F(z) - \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} |c_\mathfrak{a}^+(0)|^2 E_\mathfrak{a}(z, \tfrac{3}{2}) - 2 \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} \mathrm{Re}(c_\mathfrak{a}^+(0)\overline{c_\mathfrak{a}^-(0)}) \widetilde{E}_\mathfrak{a}(z, 1) - 4 y^{\frac{1}{2}} |\xi_{\frac{3}{2}} f(z)|^2. \tag{4-4}$$

Then $\mathcal{V}_f(\sigma_\mathfrak{a} z) = O(\log y)$ at each cusp by (2-3), so $\mathcal{V}_f \in L^2(\Gamma_0(N)\backslash\mathfrak{h})$.

In weight $k = \frac{1}{2}$, we may likewise attempt to regularize by subtracting a function of the form $y^{1/2}|g(z)|^2$, where $g \in M_{1/2}(\Gamma_0(N))$. However, there is no guarantee that a modular form with compatible cusp growth need exist. If $f \in H_{1/2}^\sharp(\Gamma_0(N))$ is chosen, we may test for the existence of a compatible g using the basis for $M_{1/2}(\Gamma_0(N))$ described in [30, Theorem A]. Since we do not require $k = \frac{1}{2}$ for our principal application, we leave the question of the existence of a compatible g as an interesting open problem.

4A. ***Automorphic regularization of $\mathcal{H}(z)$.*** In practical terms, the problem of regularizing $\mathcal{H}(z)$ reduces to the problem of computing the constant Fourier coefficients $c_{\mathfrak{a}}^{\pm}(0)$ at each singular cusp $\mathfrak{a}$ of $\Gamma_0(4)$. In this section, we determine these coefficients, as summarized in the following proposition.

Proposition 4.1. *Let $\mathcal{H} \in H_{3/2}^{\sharp}(\Gamma_0(4))$ denote Zagier's nonholomorphic Eisenstein series from (1-4). The cusp $\mathfrak{a} = \frac{1}{2}$ is nonsingular for υ_θ; for the other cusps, $\mathcal{H}(z)$ has a Fourier expansion of the form (2-2), in which*

$$c_\infty^{+}(0) = -\frac{1}{12}, \qquad c_\infty^{-}(0) = \frac{1}{8\pi}, \qquad c_0^{+}(0) = \frac{1}{24}, \qquad c_0^{-}(0) = -\frac{1}{8\pi}.$$

Consequently, the function

$$\mathcal{V}_{\mathcal{H}}(z) := y^{\frac{3}{2}}|\mathcal{H}(z)|^2 - \frac{1}{144}E_\infty\big(z, \tfrac{3}{2}\big) - \frac{1}{576}E_0\big(z, \tfrac{3}{2}\big) + \frac{1}{48\pi}\widetilde{E}_\infty(z, 1) + \frac{1}{96\pi}\widetilde{E}_0(z, 1) - \frac{1}{64\pi^2}y^{\frac{1}{2}}|\theta(z)|^2$$

lies in $L^2(\Gamma_0(4)\backslash\mathfrak{h})$.

Proof. To verify that $\mathfrak{a} = \frac{1}{2}$ is nonsingular, we first note that $\Gamma_{1/2}$ is generated by $t_{1/2} = \left(\begin{smallmatrix} -1 & 1 \\ -4 & 3 \end{smallmatrix}\right)$. Since $\upsilon_\theta(t_{1/2}) = i$, we have $\varkappa_{1/2} = \frac{1}{4} \neq 0$.

As for the singular cusps, we clearly have $c_\infty^{+}(0) = -\frac{1}{12}$ and $c_\infty^{-}(0) = \frac{1}{8\pi}$ from the Fourier expansion (1-4). To understand the behavior of $\mathcal{H}(z)$ near $\mathfrak{a} = 0$, we follow [13] and relate $\mathcal{H}(z)$ to certain Eisenstein series of weight $\frac{3}{2}$. Specifically, we introduce the Eisenstein series

$$E_{\frac{3}{2},s}(z) := \sum_{\substack{m>0,\, n\in\mathbb{Z} \\ (m,2n)=1}} \sum \frac{\big(\frac{n}{m}\big)\epsilon_m}{(mz+n)^{3/2}|mz+n|^{2s}},$$

as well as a second Eisenstein series $F_{3/2,s}(z) := z^{-3/2}|z|^{-2s}E_{3/2,s}(-1/4z)$. Though $E_{3/2,s}$ converges only for $\operatorname{Re} s > \frac{1}{4}$, [13, Theorem 2] implies that $E_{3/2,s}$ and $F_{3/2,s}$ have meromorphic continuation to $s \in \mathbb{C}$ and that

$$\mathcal{F}_s(z) := -\tfrac{1}{96}\big((1-i)E_{3/2,s}(z) - iF_{3/2,s}(z)\big)$$

satisfies $\mathcal{F}_0(z) = \mathcal{H}(z)$.

To investigate $\mathcal{H}(z)$ near the cusp 0, we compute a partial Fourier expansion of $\mathcal{F}_s|_{\sigma_0}(z)$, where $\sigma_0 = \left(\begin{smallmatrix} 0 & -1 \\ 4 & 0 \end{smallmatrix}\right)$. The functional equation $\theta(-1/4z) = (-2iz)^{1/2}\theta(z)$ implies that the weight $k = \frac{3}{2}$ slash operator satisfies

$$\begin{aligned} \mathcal{F}_s|_{\sigma_0}(z) &= (-2iz)^{-\frac{3}{2}}\mathcal{F}_s\big(-\tfrac{1}{4z}\big) \\ &= -\tfrac{1}{96}(-2iz)^{-\frac{3}{2}}\Big((1-i)z^{\frac{3}{2}}|z|^{2s}F_{3/2,s}(z) - i\big(-\tfrac{1}{4z}\big)^{-\frac{3}{2}}|4z|^{2s}E_{3/2,s}(z)\Big) \\ &= -\tfrac{1}{192}i|z|^{2s}F_{3/2,s}(z) + \tfrac{1}{48}(1-i)|z|^{2s}2^{4s}E_{3/2,s}(z). \end{aligned}$$

Thus $\mathcal{H}|_{\sigma_0}(z)$ has a Fourier expansion which may be read from the Fourier coefficients of $E_{3/2,0}(z)$ and $F_{3/2,0}(z)$. Since we require only the constant Fourier coefficient of $\mathcal{H}|_{\sigma_0}$, it suffices to consider the constant Fourier coefficients of $E_{3/2,s}$ and $F_{3/2,s}$.

By [13, p.93-94], the constant Fourier coefficient of $E_{3/2,s}(z)$ equals

$$\alpha_0(s,y)\sum_{m\text{ odd}}\frac{\epsilon_m\sum_{j(m)}\left(\frac{j}{m}\right)}{m^{2s+\frac{3}{2}}}=\alpha_0(s,y)\sum_{m\text{ odd}}\frac{\epsilon_{m^2}\varphi(m^2)}{(m^2)^{2s+\frac{3}{2}}}$$
$$=\alpha_0(s,y)\sum_{m\text{ odd}}\frac{\varphi(m)}{m^{4s+2}}=\alpha_0(s,y)\frac{\zeta(4s+1)}{\zeta(4s+2)}\cdot\frac{1-2^{-4s-1}}{1-2^{-4s-2}},$$

in which $\varphi(m)$ denotes the totient function and $\alpha_0(s,y)$ is defined by

$$\alpha_0(s,y)=\int_{\operatorname{Im}w=y}w^{-\frac{3}{2}}|w|^{-2s}\,dw=-\frac{2(1+i)\sqrt{\pi}\,s\Gamma\left(2s+\frac{1}{2}\right)}{\Gamma(2s+2)}y^{-\frac{1}{2}-2s}.$$

By taking the limit as $s\to 0$, we conclude that the constant Fourier coefficient of $E_{3/2,0}(z)$ equals $-(2+2i)/\pi\cdot y^{-1/2}$.

Similarly, formulas on [13, p.94] show that $F_{3/2,s}$ has constant Fourier coefficient

$$2^{2s+3}i+(1+i)\alpha_0(s,y)\sum_{m\text{ even}}\frac{m^{-1/2}\sum_{j(2m)}\left(\frac{m}{j}\right)e\left(\frac{j}{8}\right)}{(m/2)^{2s+1}}.$$

The sum $\sum_{j \bmod 2m}\left(\frac{m}{j}\right)e(j/8)$ vanishes for m even unless $m=2n^2$, in which case it equals $\sqrt{2}\varphi(2n^2)$. Thus the constant Fourier coefficient equals

$$2^{2s+3}i+(1+i)\alpha_0(s,y)\sum_{n\geq 1}\frac{\varphi(2n^2)}{n^{4s+3}}=2^{2s+3}i+(1+i)\alpha_0(s,y)\sum_{n\geq 1}\frac{\varphi(2n)}{n^{4s+2}}$$
$$=2^{2s+3}i+(1+i)\alpha_0(s,y)2^{4s+2}\frac{\zeta(4s+1)}{\zeta(4s+2)}\left(1-\frac{1-2^{-4s-1}}{1-2^{-4s-2}}\right).$$

By taking the limit as $s\to 0$, we conclude that the constant Fourier coefficient of $F_{3/2,0}(z)$ equals $8i-8i/(\pi y^{1/2})$. It follows that the constant Fourier coefficient of $\mathcal{H}|_{\sigma_0}(z)$ equals

$$\frac{-i}{192}\left(8i-\frac{8i}{\pi\sqrt{y}}\right)+\frac{1-i}{48}\left(\frac{-2-2i}{\pi y^{1/2}}\right)=\frac{1}{24}-\frac{1}{8\pi\sqrt{y}},$$

hence $c_0^+(0)=\frac{1}{24}$ and $c_0^-(0)=-\frac{1}{8\pi}$. □

5. Inner products involving regularization terms

As before, we fix $k\in\frac{1}{2}\mathbb{Z}$ and $f(z)\in H_k^\sharp(\Gamma_0(N))$ and define $F(z)=y^k|f(z)|^2$. In Section 4, we showed that $F(z)$ differed from an L^2 function $\mathcal{V}_f(z)$ by a sum involving Eisenstein series and theta functions, at least when $k\notin\left\{\frac{1}{2},1\right\}$. In this section, we relate $\langle\mathcal{V}_f,P_\ell(\cdot,\bar{s})\rangle$ to $\langle F,P_\ell(\cdot,\bar{s})\rangle$ by accounting for the contribution of these regularization terms.

To compute the inner products of the form $\langle E_{\mathfrak{a}}(\cdot, w), P_\ell(\cdot, \bar{s})\rangle$, we recall the Fourier expansion of $E_{\mathfrak{a}}(z, w)$ from (4-1). We unfold the inner product using the Poincaré series as in (3-1) to produce

$$\langle E_{\mathfrak{a}}(\cdot, w), P_\ell(\cdot, \bar{s})\rangle = \frac{2\pi^{w+\frac{1}{2}}}{(4\pi\ell)^{s-\frac{1}{2}}} \ell^{w-\frac{1}{2}} \varphi_{\mathfrak{a}\infty\ell}(w) \frac{\Gamma(s+w-1)\Gamma(s-w)}{\Gamma(s)\Gamma(w)}, \tag{5-1}$$

provided that $\operatorname{Re} s > \frac{1}{2} + |\operatorname{Re} w - \frac{1}{2}|$ to begin. We write $\varphi_{\mathfrak{a}\infty n}(w) = \varphi_{\mathfrak{a}n}(w)$ for brevity and remark that formulas for these coefficients appear in [4].

The functions $\varphi_{\mathfrak{a}n}(w)$ have meromorphic continuation in w. For $n \neq 0$, they are analytic at $w = 1$. By considering the Laurent expansion of each side of (5-1) at $w = 1$, we obtain

$$\langle \widetilde{E}_{\mathfrak{a}}(\cdot, 1), P_\ell(\cdot, \bar{s})\rangle = \frac{\pi\varphi_{\mathfrak{a}\ell}(1)}{(4\pi\ell)^{s-1}} \Gamma(s-1),$$

in which $\widetilde{E}_{\mathfrak{a}}(z, 1)$ is the constant term of the Laurent expansion of $E_{\mathfrak{a}}(z, w)$ at $w = 1$ (as defined immediately before (4-3)).

Lastly, we consider inner products of the form $\langle y^{\frac{1}{2}}|g(z)|^2, P_\ell(\cdot, \bar{s})\rangle$, in which $g \in M_{1/2}(\Gamma_0(N))$. Suppose that $g(z) = \sum b(n)e(nz)$. Then

$$\langle y^{\frac{1}{2}}|g(z)|^2, P_\ell(\cdot, \bar{s})\rangle = \frac{\Gamma\big(s-\frac{1}{2}\big)}{(4\pi)^{s-\frac{1}{2}}} \sum_{n\geq 0} \frac{b(n+\ell)\overline{b(n)}}{(n+\ell)^{s-\frac{1}{2}}}.$$

By [30, Theorem A], $M_{1/2}(\Gamma_0(N))$ is spanned by theta functions of the form $\sum \chi_t(n)e(n^2tz)$, with t square-free and satisfying $4\operatorname{cond}(\chi_t)^2 t \mid N$, where $\operatorname{cond}(\chi)$ denotes the conductor of χ, $\chi_t(n) = \big(\frac{t}{n}\big)$ if $t \equiv 1 \bmod 4$, and $\chi_t(n) = \big(\frac{4t}{n}\big)$ otherwise. We note that $\operatorname{cond}(\chi_t) = t$ if $t \equiv 1 \bmod 4$ and $4t$ otherwise. In particular, $\{b(n)\}$ is supported on integers of the form n^2t, where $n \in \mathbb{Z}$ and $4t^3 \mid N$. Thus

$$\langle y^{\frac{1}{2}}|g(z)|^2, P_\ell(\cdot, \bar{s})\rangle = \frac{\Gamma\big(s-\frac{1}{2}\big)}{(4\pi)^{s-\frac{1}{2}}} \sum_{\substack{t_1^3, t_2^3 \mid \frac{N}{4} \\ t_i \text{ square-free}}} \sum_{n_1^2 t_1 = n_2^2 t_2 + \ell} \frac{b(n_1^2 t_1)\overline{b(n_2^2 t_2)}}{(n_1^2 t_1)^{s-\frac{1}{2}}}.$$

If $M_{1/2}(\Gamma_0(N))$ is one-dimensional (for example, if $\frac{N}{4}$ is cube-free), then $t_1 = t_2 = 1$. In this case, the inner sum $n_1^2 = n_2^2 + \ell$ has finitely many solutions. Otherwise, the sum may be infinite (depending on ℓ). Since the solution set (n_1, n_2) of $n_1^2 t_1 = n_2^2 t_2 + \ell$ is exponentially sparse in any case, the sum above always converges for $\operatorname{Re} s > \frac{1}{2}$. Thus $\langle y^{\frac{1}{2}}|g(z)|^2, P_\ell(\cdot, \bar{s})\rangle$ is analytic in $\operatorname{Re} s > \frac{1}{2}$ no matter the dimension of $M_{1/2}(\Gamma_0(N))$.

Remark 5.1. In fact, $\langle y^{\frac{1}{2}}|g(z)|^2, P_\ell(\cdot, \bar{s})\rangle$ has a meromorphic continuation to all $s \in \mathbb{C}$. To see this, note that the series above is essentially supported on positive integers x satisfying the generalized Pell equation $t_1x^2 - t_2y^2 = \ell$. When solutions exist, they lie in finitely many classes of linear recurrences. Splitting $\langle y^{\frac{1}{2}}|g(z)|^2, P_\ell(\cdot, \bar{s})\rangle$ along this subdivision, and splitting further to ignore the effect of the characters χ_{t_1} and χ_{t_2}, it suffices to continue series of the form $\sum_{m\geq 1} A_m^{-s}$, where $\{A_m\}$ satisfies a degree two linear recurrence. Fortunately, such results are known; see for example [29], which treats a much more general case.

At this point, it is straightforward to relate the inner products $\langle F, P_\ell(\cdot, \bar{s})\rangle$ and $\langle \mathcal{V}_f, P_\ell(\cdot,\bar{s})\rangle$. We record our results in the following proposition.

Proposition 5.2. *Let $f \in H_k^\sharp(\Gamma_0(N))$ and set $F(z) = y^k|f(z)|^2$. For $k = \frac{3}{2}$,*

$$\begin{aligned}\langle F, P_\ell(\cdot,\bar{s})\rangle = \langle \mathcal{V}_f, P_\ell(\cdot,\bar{s})\rangle &+ \frac{\sqrt{\pi}\,\Gamma\big(s+\frac{1}{2}\big)\Gamma\big(s-\frac{3}{2}\big)}{(4\pi\ell)^{s-\frac{3}{2}}\Gamma(s)} \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} |c_\mathfrak{a}^+(0)|^2 \varphi_{\mathfrak{a}\ell}\big(\tfrac{3}{2}\big)\\ &+ \frac{2\pi\Gamma(s-1)}{(4\pi\ell)^{s-1}} \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} \mathrm{Re}\big(c_\mathfrak{a}^+(0)\overline{c_\mathfrak{a}^-(0)}\big)\varphi_{\mathfrak{a}\ell}(1)\\ &+ \frac{4\Gamma\big(s-\frac{1}{2}\big)}{(4\pi)^{s-\frac{1}{2}}} \sum_{\substack{t_1^3,t_2^3 \mid \frac{N}{4}\\ t_i \text{ square-free}}} \sum_{n_1^2 t_1 = n_2^2 t_2 + \ell} \frac{a_{\xi f}(n_1^2 t_1)\overline{a_{\xi f}(n_2^2 t_2)}}{(n_1^2 t_1)^{s-\frac{1}{2}}},\end{aligned}$$

in which $\mathcal{V}_f$ is defined as in (4-4) and $a_{\xi f}(n)$ denotes the n-th Fourier coefficient of $\xi_{3/2} f(z)$. For $k < \frac{1}{2}$ in $\frac{1}{2}\mathbb{Z}$, we have instead

$$\begin{aligned}\langle F, P_\ell(\cdot,\bar{s})\rangle = \langle \mathcal{V}_f, P_\ell(\cdot,\bar{s})\rangle &+ \frac{2\pi\Gamma(s+1-k)\Gamma(s+k-2)}{(4\pi\ell)^{s-\frac{1}{2}}(\pi\ell)^{k-\frac{3}{2}}\Gamma(s)\Gamma(2-k)} \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} |c_\mathfrak{a}^-(0)|^2\varphi_{\mathfrak{a}\ell}(2-k)\\ &+ \frac{2\pi\Gamma(s-1)}{(4\pi\ell)^{s-1}} \sum_{\mathfrak{a}:\varkappa_\mathfrak{a}=0} \mathrm{Re}\big(c_\mathfrak{a}^+(0)\overline{c_\mathfrak{a}^-(0)}\big)\varphi_{\mathfrak{a}\ell}(1),\end{aligned}$$

in which $\mathcal{V}_f$ is defined as in (4-3).

As a corollary, we specify the contribution of correction terms in the regularization $\mathcal{V}_\mathcal{H}(z)$ of $\mathcal{H}(z)$.

Corollary 5.3. *Let ℓ_o denote the odd-part of ℓ. In $\mathrm{Re}\, s > \frac{3}{2}$, we have*

$$\begin{aligned}\langle y^{\frac{3}{2}}|\mathcal{H}|^2, P_\ell(\cdot,\bar{s})\rangle = \langle \mathcal{V}_\mathcal{H}, P_\ell(\cdot,\bar{s})\rangle &+ \frac{\sqrt{\pi}\,\Gamma\big(s+\frac{1}{2}\big)\Gamma\big(s-\frac{3}{2}\big)}{(4\pi\ell)^{s-\frac{3}{2}}\Gamma(s)} \cdot \frac{2\sigma_{-2}\big(\frac{\ell}{4}\big)-\sigma_{-2}\big(\frac{\ell}{2}\big)+\sigma_{-2}(\ell_o)}{4032\,\zeta(3)}\\ &- \frac{\Gamma(s-1)}{(4\pi\ell)^{s-1}} \cdot \frac{2\sigma_{-1}\big(\frac{\ell}{4}\big)-\sigma_{-1}\big(\frac{\ell}{2}\big)+\sigma_{-1}(\ell_o)}{288\,\zeta(2)} + \frac{\Gamma\big(s-\frac{1}{2}\big)}{32\pi^{s+\frac{3}{2}}} \sum_{\substack{d\mid\ell\\ d\equiv\frac{\ell}{d} \bmod 2}} \big(d+\tfrac{\ell}{d}\big)^{1-2s}.\end{aligned}$$

Proof. Since $\xi_{3/2}\mathcal{H}(z) = -\frac{1}{16\pi}\theta(z)$ by (2-3) and the Fourier expansion (1-4), we have

$$a_{\xi\mathcal{H}}(n) = -\frac{1}{16\pi} r_1(n).$$

Propositions 4.1 and 5.2 then give

$$\langle y^{\frac{3}{2}}|\mathcal{H}|^2, P_\ell(\cdot,\bar{s})\rangle = \langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot,\bar{s})\rangle + \frac{\sqrt{\pi}\,\Gamma\big(s+\frac{1}{2}\big)\Gamma\big(s-\frac{3}{2}\big)}{(4\pi\ell)^{s-\frac{3}{2}}\Gamma(s)}\Big(\tfrac{1}{144}\varphi_{\infty\ell}\big(\tfrac{3}{2}\big)+\tfrac{1}{576}\varphi_{0\ell}\big(\tfrac{3}{2}\big)\Big)$$
$$-\frac{\Gamma(s-1)}{(4\pi\ell)^{s-1}}\Big(\tfrac{1}{48}\varphi_{\infty\ell}(1)+\tfrac{1}{96}\varphi_{0\ell}(1)\Big)+\frac{\Gamma\big(s-\frac{1}{2}\big)}{4(4\pi)^{s+\frac{3}{2}}}\sum_{n\geq 0}\frac{r_1(n+\ell)r_1(n)}{(n+\ell)^{s-\frac{1}{2}}}. \quad (5\text{-}2)$$

To simplify further, we give explicit descriptions of the Fourier coefficients $\varphi_{\mathfrak{a}\ell}(w)$. Conveniently, the formulas we require appear in [15, §3.3]:

$$\varphi_{0\ell}(w)=\frac{\sigma^{(2)}_{1-2w}(\ell)}{4^w\zeta^{(2)}(2w)},\quad \varphi_{\infty\ell}(w)=\frac{2^{2-4w}\sigma_{1-2w}\big(\frac{\ell}{4}\big)-2^{1-4w}\sigma_{1-2w}\big(\frac{\ell}{2}\big)}{\zeta^{(2)}(2w)}, \quad (5\text{-}3)$$

in which $\zeta^{(2)}(s)=(1-2^{-s})\zeta(s)$, $\sigma^{(2)}_\nu$ denotes the sum-of-divisors function with its 2-factor removed, and we adopt the convention that $\sigma_\nu(m)=0$ for $m\notin\mathbb{Z}$. By Euler products, $\sigma^{(2)}_\nu(\ell)=\sigma_\nu(\ell_o)$, where ℓ_0 is the odd part of ℓ.

Finally, we note that the series in (5-2) may be written as a divisor sum:

$$\sum_{n\geq 0}\frac{r_1(n+\ell)r_1(n)}{(n+\ell)^{s-\frac{1}{2}}}=\sum_{\substack{n_1,n_2\in\mathbb{Z}\\ n_2^2-n_1^2=\ell}}|n_2|^{1-2s}=2^{2s-1}\sum_{\substack{d_1,d_2\in\mathbb{Z}\\ d_1d_2=\ell\\ d_1\equiv d_2 \bmod 2}}|d_1+d_2|^{1-2s}=2^{2s}\sum_{\substack{d\mid\ell\\ d\equiv\frac{\ell}{d} \bmod 2}}\Big(d+\frac{\ell}{d}\Big)^{1-2s},$$

which completes the proof. □

6. Spectral expansion and rightmost poles

As before, fix $f\in H_k^\sharp(\Gamma_0(N))$ with $k\notin\{\frac{1}{2},1\}$ and define $F(z)=y^k|f(z)|^2$. In this section, we show that the inner product $\langle F, P_\ell(\cdot,\bar{s})\rangle$ admits meromorphic continuation to $s\in\mathbb{C}$. By Proposition 5.2, it suffices to consider the inner products $\langle \mathcal{V}_f, P_\ell(\cdot,\bar{s})\rangle$ instead, as the regularization terms contribute explicit terms which are meromorphic in $\mathbb{C}$, either by inspection or as a consequence of Remark 5.1.

Selberg's spectral theorem (cf. [18, Theorem 15.5]) gives the following spectral expansion of $P_\ell(z,s)$:

$$P_\ell(z,s)=\sum_j\langle P_\ell(\cdot,s),\mu_j\rangle\mu_j(z)+\sum_{\mathfrak{a}}\frac{V_N}{4\pi}\int_{-\infty}^{\infty}\big\langle P_\ell(\cdot,s),E_{\mathfrak{a}}\big(\cdot,\tfrac{1}{2}+it\big)\big\rangle E_{\mathfrak{a}}\big(z,\tfrac{1}{2}+it\big)\,dt, \quad (6\text{-}1)$$

in which $\mathfrak{a}$ varies through the cusps of $\Gamma_0(N)$, $V_N=\frac{\pi}{3}\cdot N\prod_{p\mid N}(1+1/p)$ denotes the volume of $\Gamma_0(N)\backslash\mathfrak{h}$, and $\{\mu_j\}$ is an orthonormal Hecke eigenbasis for the space of weight 0 Maass cusp forms on $\Gamma_0(N)$. These Maass forms have Fourier expansions at all cusps, which we write in the form

$$\mu_{j\mathfrak{a}}(z):=\mu_j|_{\sigma_{\mathfrak{a}}}(z)=y^{\frac{1}{2}}\sum_{n\neq 0}\rho_{j\mathfrak{a}}(n)K_{it_j}(2\pi|n|y)e(nx). \quad (6\text{-}2)$$

We next record two useful lemmas regarding the growth of the coefficients $\rho_{j\mathfrak{a}}(m)$ on average. The first of them concerns the average growth of $\rho_{j\mathfrak{a}}(m)$ with respect to m and is taken from [17].

Lemma 6.1 [17, (8.7)]. *Let μ_j be an L^2-normalized Maass cusp form on $\Gamma_0(N)$ with Fourier expansion at $\mathfrak{a}$ of the form (6-2). Then*

$$\sum_{m\leq M} |\rho_{j\mathfrak{a}}(m)|^2 \ll_N (M+|t_j|)e^{\pi|t_j|}.$$

Our second lemma is a spectral average generalizing [24, Theorem 6].

Lemma 6.2. *Let $\{\mu_j\}$ denote an orthonormal basis of Maass cusp forms for $\Gamma_0(N)$, with Fourier expansions given by (6-2). For $\ell > 0$ and any $\epsilon > 0$,*

$$\sum_{|t_j|\leq X} \frac{|\rho_j(\ell)|^2}{\cosh \pi t_j} = \frac{X^2}{\pi^2} + O_{N,\epsilon}(X\log X + X\ell^{\epsilon} + \ell^{\frac{1}{2}+\epsilon}).$$

Proof. For level $N = 1$, this result is [24, Theorem 6]. More generally, we adapt [24, §6], replacing the level 1 trace formula with one on $\Gamma_0(N)$, as found in [4, Lemma 4.7], for example. To carry out this generalization, we require the Kloosterman sum estimate $S_{\infty\infty}(\ell,\ell,c) \ll (\ell,c)^{1/2}c^{1/2}d(c)$ from [4, Lemma 2.6] as well as the Eisenstein series coefficient estimate

$$\varphi_{\mathfrak{a}\ell}\big(\tfrac{1}{2}+it\big) \ll_N \frac{d(\ell)}{|\zeta(1+2it)|} \ll_{N,\epsilon} d(\ell)\log t, \tag{6-3}$$

To see (6-3), one may represent $E_{\mathfrak{a}}(z,s)$ in terms of Eisenstein series attached to characters via [35, Theorem 6.1] and apply the Fourier coefficient formulas in [35, Proposition 4.1], then apply [32, (3.11.10)]. $\square$

Continuing on, substitution of (6-1) into $\langle \mathcal{V}_f, P_h(\cdot,\bar{s})\rangle$ produces

$$\langle \mathcal{V}_f, P_\ell(\cdot,\bar{s})\rangle = \sum_j \langle \mu_j, P_\ell(\cdot,\bar{s})\rangle\langle \mathcal{V}_f, \mu_j(z)\rangle + \frac{V_N}{4\pi}\sum_{\mathfrak{a}}\int_{-\infty}^{\infty} \big\langle E_{\mathfrak{a}}\big(\cdot,\tfrac{1}{2}+it\big), P_\ell(\cdot,\bar{s})\big\rangle\big\langle \mathcal{V}_f, E_{\mathfrak{a}}\big(z,\tfrac{1}{2}+it\big)\big\rangle\,dt, \tag{6-4}$$

which we call the *spectral expansion of* $\langle \mathcal{V}_f, P_\ell(\cdot,\bar{s})\rangle$. We will refer to the terms at right in (6-4) as the *discrete spectrum* and *continuous spectrum*, respectively. To make this more explicit, we apply (5-1) and the formula

$$\langle \mu_j, P_\ell(\cdot,\bar{s})\rangle = \frac{\rho_j(\ell)\sqrt{\pi}}{(4\pi\ell)^{s-\frac{1}{2}}}\frac{\Gamma\big(s-\frac{1}{2}-it_j\big)\Gamma\big(s-\frac{1}{2}+it_j\big)}{\Gamma(s)},$$

which follows from [9, 6.621(3)]. We conclude that $\langle \mathcal{V}_f, P_\ell(\cdot,\bar{s})\rangle$ admits a spectral decomposition of the form $\Sigma_{\mathrm{disc}}(s) + \Sigma_{\mathrm{cont}}(s)$, in which

$$\Sigma_{\mathrm{disc}}(s) := \frac{\sqrt{\pi}}{(4\pi\ell)^{s-\frac{1}{2}}\Gamma(s)}\sum_j \rho_j(\ell)\Gamma\big(s-\tfrac{1}{2}+it_j\big)\Gamma\big(s-\tfrac{1}{2}-it_j\big)\langle \mathcal{V}_f, \mu_j\rangle,$$

$$\Sigma_{\mathrm{cont}}(s) := \frac{V_N}{2}\sum_{\mathfrak{a}}\int_{-\infty}^{\infty} \frac{\varphi_{\mathfrak{a}\ell}\big(\frac{1}{2}+it\big)\Gamma\big(s-\frac{1}{2}+it\big)\Gamma\big(s-\frac{1}{2}-it\big)}{(4\pi\ell)^{s-\frac{1}{2}}(\pi\ell)^{-it}\Gamma(s)\Gamma\big(\frac{1}{2}+it\big)}\big\langle \mathcal{V}_f, E_{\mathfrak{a}}\big(\cdot,\tfrac{1}{2}+it\big)\big\rangle\,dt.$$

This spectral expansion is initially defined for $\operatorname{Re} s > 1 + |k-1|$, provided all expressions converge. Fortunately, convergence is not an issue:

Lemma 6.3. *The functions $\Sigma_{\mathrm{disc}}(s)$ and $\Sigma_{\mathrm{cont}}(s)$ converge for all $s \in \mathbb{C}$ away from their poles.*

Proof. In the discrete spectrum $\Sigma_{\mathrm{disc}}(s)$, this follows from Lemma 6.2, Stirling's approximation (providing decay of size $e^{-\pi|t_j|}$ for fixed s), and the trivial estimate $|\langle \mathcal{V}_f, \mu_j\rangle| \ll \|\mathcal{V}_f\|^{1/2} \cdot \|\mu_j\|^{1/2} \ll_f 1$. Here, we've used that $\|\mu_j\| = 1$ (by definition of μ_j) and that $\mathcal{V}_f \in L^2(\Gamma_0(N)\backslash\mathfrak{h})$ by our work in Section 4. (This estimate is *very* weak, and will be improved in Section 8.)

In the continuous spectrum, convergence follows from Stirling's approximation, weak upper bounds for $\big\langle \mathcal{V}_f, E_{\mathfrak{a}}\big(\cdot, \tfrac{1}{2}+it\big)\big\rangle$ derived from the Rankin–Selberg method and Phragmén–Lindelöf convexity principle, and (6-3). □

Thus $\Sigma_{\mathrm{disc}}(s)$ defines a meromorphic function on the entire complex plane, with potential poles at $s = \frac{1}{2} \pm it_j - m$ for integer $m \geq 0$ and any spectral type t_j. It is analytic in the right half-plane $\operatorname{Re} s > \frac{1}{2} + \Theta$, where $\Theta \leq 7/64$ denotes partial progress towards the Ramanujan–Petersson conjecture [21].

The continuous spectrum $\Sigma_{\mathrm{cont}}(s)$ also has meromorphic continuation to all s, though the precise continuation to $\operatorname{Re} s > -M$ involves both $\Sigma_{\mathrm{cont}}(s)$ and $O(M)$ residue terms extracted through contour shifts of the integral in Σ_{cont}. For a discussion of the continuation process in a similar case, we refer the reader to [14, §4] or [15, §3.3.2]. The continuous spectrum is clearly analytic in $\operatorname{Re} s > \frac{1}{2}$.

Thus $\langle \mathcal{V}_f, P_\ell(\cdot, \bar{s})\rangle$, originally defined for $\operatorname{Re} s > 1 + |k-1|$, extends meromorphically to a function on the entire complex plane. Since it is analytic in $\operatorname{Re} s > \frac{1}{2} + \Theta$, any pole of $\langle F, P_\ell(\cdot, \bar{s})\rangle$ in $\operatorname{Re} s > \frac{1}{2} + \Theta$ occurs as a pole of the explicit regularization factors presented in Proposition 5.2.

Remark 6.4. In Section 3, we gave the general decomposition

$$\langle F, P_\ell(\cdot, \bar{s})\rangle = I_\ell^{+}(s) + I_\ell^{0}(s) + I_\ell^{\times}(s) + I_\ell^{-}(s).$$

The two terms $I_\ell^0(s)$ and $I_\ell^\times(s)$ are finite sums and inherit meromorphic continuation to $s \in \mathbb{C}$ from the continuations of G_k and the ${}_2F_1$-hypergeometric function. Thus the continuation of $\langle F, P_\ell(\cdot, \bar{s})\rangle$ implies a continuation for $I_\ell^+(s) + I_\ell^-(s)$. It is possible, albeit challenging, to establish the meromorphic continuations of $I_\ell^+(s)$ and $I_\ell^-(s)$ as separate entities. Here, the idea is to first continue $I_\ell^-(s)$ by relating it to the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{a_{\xi f}(n)\overline{a_{\xi f}(n+\ell)}}{(n+\ell)^{s-w-k} n^{w+1}},$$

which admits meromorphic continuation through relation to the triple inner product $\langle y^{2-k}|\xi_k f|^2, P_\ell(\cdot, \bar{s})\rangle$. Establishing this continuation is not so difficult when $\operatorname{Re} w < -1$, but in practice we require $\operatorname{Re} w$ as large as $-k$ (to evaluate a particular contour integral representation), and this creates major complications in weights $k < 1$.

Fortunately, these problems disappear altogether for $f = \mathcal{H}$, as the series $I_\ell^-(s)$ defines a *finite* sum in this case (cf. Section 3A). To simplify the exposition in this work, we narrow our typical focus from $f \in H_k^\sharp(\Gamma_0(N))$ to $f = \mathcal{H}$. Still, the construction of the meromorphic continuations of $I_\ell^\pm(s)$ and estimates for shifted convolutions of generic mock modular forms of polynomial growth are of independent interest and will appear in future work.

6A. ***Classifying the rightmost poles of $D_\ell(s)$.*** As an application of our work thus far, we classify the rightmost poles of the shifted convolution Dirichlet series $D_\ell(s)$ from (3-4). We prove the following theorem.

Theorem 6.5. *The Dirichlet series $D_\ell(s)$ is analytic in the right half-plane* $\operatorname{Re} s > \frac{3}{2}$ *and extends meromorphically to all $s \in \mathbb{C}$. If $\ell \equiv 2 \bmod 4$, then $D_\ell(s) = 0$ identically. Otherwise, $D_\ell(s)$ has two simple poles in the right half-plane* $\operatorname{Re} s > \frac{1}{2}$, *at $s = \frac{3}{2}$ and $s = 1$, with residues*

$$\operatorname*{Res}_{s=\frac{3}{2}} D_\ell(s) = \frac{\pi^2}{126\,\zeta(3)}\bigl(2\sigma_{-2}\bigl(\tfrac{\ell}{4}\bigr) - \sigma_{-2}\bigl(\tfrac{\ell}{2}\bigr) + \sigma_{-2}(\ell_o)\bigr),$$

$$\operatorname*{Res}_{s=1} D_\ell(s) = -\frac{1}{3\pi}\bigl(2\sigma_{-1}\bigl(\tfrac{\ell}{4}\bigr) - \sigma_{-1}\bigl(\tfrac{\ell}{2}\bigr) + \sigma_{-1}(\ell_o)\bigr).$$

The function $D_\ell(s)$ is otherwise analytic in $\operatorname{Re} s > \frac{1}{2}$.

Proof. Equation (3-5) relates $D_\ell(s)$ to $\langle y^{3/2}|\mathcal{H}|^2, P_\ell(\cdot, \bar{s})\rangle$ and Corollary 5.3 relates $\langle y^{3/2}|\mathcal{H}|^2, P_\ell(\cdot, \bar{s})\rangle$ to $\langle \mathcal{V}_\mathcal{H}, P_\ell(\cdot, \bar{s})\rangle$. When combined, this produces

$$\begin{aligned} D_\ell(s) = {} & \frac{\pi^{\frac{5}{2}}\Gamma\bigl(s-\frac{3}{2}\bigr)}{\ell^{s-\frac{3}{2}}\Gamma(s)} \cdot \frac{2\sigma_{-2}\bigl(\frac{\ell}{4}\bigr) - \sigma_{-2}\bigl(\frac{\ell}{2}\bigr) + \sigma_{-2}(\ell_o)}{252\,\zeta(3)} - \frac{\pi^{\frac{3}{2}}\Gamma(s-1)}{\ell^{s-1}\Gamma\bigl(s+\frac{1}{2}\bigr)} \cdot \frac{2\sigma_{-1}\bigl(\frac{\ell}{4}\bigr) - \sigma_{-1}\bigl(\frac{\ell}{2}\bigr) + \sigma_{-1}(\ell_o)}{36\,\zeta(2)} \\ & + \frac{(4\pi)^{s+\frac{1}{2}}\langle \mathcal{V}_\mathcal{H}, P_\ell(\cdot, \bar{s})\rangle}{\Gamma\bigl(s+\frac{1}{2}\bigr)} - \frac{H(\ell)\Gamma(s)}{4\sqrt{\pi}\,\ell^s\,\Gamma\bigl(s+\frac{1}{2}\bigr)} + \frac{H(\ell)}{12\ell^{s+\frac{1}{2}}} \\ & - \frac{r_1(\ell)}{32\pi\ell^{s-\frac{1}{2}}s\bigl(s-\frac{1}{2}\bigr)} + \frac{r_1(\ell)\Gamma(s)}{96\sqrt{\pi}\,\ell^s\,\Gamma\bigl(s+\frac{3}{2}\bigr)} \\ & - \frac{\Gamma(s)}{\Gamma\bigl(s+\frac{3}{2}\bigr)} \sum_{m=1}^{\ell-1} \frac{H(\ell-m)r_1(m)}{8\sqrt{\pi}\,m^s} \, {}_2F_1\biggl(\begin{matrix} s, s+\frac{1}{2} \\ s+\frac{3}{2} \end{matrix}\biggm| 1 - \frac{\ell}{m}\biggr) \\ & - \sum_{\substack{m_1^2-m_2^2=\ell \\ m_1,m_2\geq 1}} \frac{m_1 m_2 G_{3/2}(s, m_2^2, m_1^2)}{16\pi\,\Gamma\bigl(s+\frac{1}{2}\bigr)} + \frac{2^{2s-4}}{\pi\bigl(s-\frac{1}{2}\bigr)} \sum_{\substack{d\mid\ell \\ d\equiv\frac{\ell}{d} \bmod 2}} \bigl(d + \tfrac{\ell}{d}\bigr)^{1-2s}. \end{aligned} \tag{6-5}$$

Recall that $\langle \mathcal{V}_\mathcal{H}, P_\ell(\cdot, \bar{s})\rangle$ is analytic in $\operatorname{Re} s > \frac{1}{2} + \Theta$. By Huxley's resolution of the Selberg eigenvalue conjecture in low level [16], the inner product is in fact analytic in $\operatorname{Re} s > \frac{1}{2}$. Thus, by previous comments, all but the first two terms at right above are analytic in $\operatorname{Re} s > \frac{1}{2}$. Computation of residues completes the proof. □

Since $D_\ell(s)$ has nonnegative coefficients, the Wiener–Ikehara theorem (see [25, Corollary 8.8], for example) immediately produces the following:

Corollary 6.6. *For fixed ℓ, as $X \to \infty$ we have*

$$\sum_{n \le X} H(n)H(n+\ell) \sim \frac{\pi^2 X^2}{252\,\zeta(3)}\big(2\sigma_{-2}\big(\tfrac{\ell}{4}\big) - \sigma_{-2}\big(\tfrac{\ell}{2}\big) + \sigma_{-2}(\ell_o)\big).$$

7. Bounding $D_\ell(s)$ in vertical strips

To quantify the rate of convergence in Corollary 6.6, we require additional information about the meromorphic properties of $D_\ell(s)$. Specifically, we require uniform estimates for the growth of $D_\ell(s)$ with respect to $|\mathrm{Im}\, s|$ in vertical strips outside the domain of absolute convergence.

It suffices to produce growth estimates for each component of the decomposition of $D_\ell(s)$ given in (6-5). In this section, we produce uniform estimates for every term besides $(4\pi)^{s+\frac{1}{2}}\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot\,,\bar{s})\rangle / \Gamma\big(s+\frac{1}{2}\big)$, which requires more involved techniques.

Proposition 7.1. *Fix s with $\mathrm{Re}\, s > 0$. Away from poles of $D_\ell(s)$, we have*

$$D_\ell(s) \ll_\epsilon \ell^{-\mathrm{Re}\, s+\epsilon} + \ell^{\frac{3}{2}-\mathrm{Re}\, s+\epsilon}|s|^{-\frac{3}{2}} + \left|\frac{\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot\,,\bar{s})\rangle}{\Gamma\big(s+\frac{1}{2}\big)}\right|$$

for all $\epsilon > 0$.

The proof requires a few lemmas, starting with a simple upper bound for the Hurwitz class number.

Lemma 7.2. *We have $H(\ell) \ll_\epsilon \ell^{\frac{1}{2}+\epsilon}$ for all $\epsilon > 0$.*

Proof. The moment estimate (1-1) implies $\tilde{h}(-\ell) \ll_\epsilon \ell^{\frac{1}{2}+\epsilon}$ for all $\epsilon > 0$. Since $h(-\ell) = \sum_{d^2 \mid \ell} \tilde{h}(-\ell/d^2)$, we have $h(-\ell) \ll \sum_{d \ge 1} (\ell/d^2)^{\frac{1}{2}+\epsilon} \ll \ell^{\frac{1}{2}+\epsilon}$. The same bound holds for $H(\ell) = h(-\ell) + O(1)$. ☐

We also require uniform estimates for the ${}_2F_1$-hypergeometric function and the function $G_{3/2}$, which are provided by the following two lemmas.

Lemma 7.3. *For $1 \le m \le \ell - 1$ and $\mathrm{Re}\, s > 0$, we have*

$${}_2F_1\left(\begin{matrix} s, s+\frac{1}{2} \\ s+\frac{3}{2}\end{matrix}\,\middle|\, 1-\frac{\ell}{m}\right) \ll \left(\frac{m}{\ell}\right)^{\mathrm{Re}\, s}.$$

Proof. Following [9, 9.131(1)] and the Euler integral [9, 9.111],

$$\begin{aligned} {}_2F_1\left(\begin{matrix} s, s+\frac{1}{2} \\ s+\frac{3}{2}\end{matrix}\,\middle|\, 1-\frac{\ell}{m}\right) &= \left(\frac{\ell}{m}\right)^{1-s} {}_2F_1\left(\begin{matrix} \frac{3}{2}, 1 \\ s+\frac{3}{2}\end{matrix}\,\middle|\, 1-\frac{\ell}{m}\right) \\ &= \left(\frac{\ell}{m}\right)^{1-s}\big(s+\tfrac{1}{2}\big)\int_0^1 \frac{(1-t)^{s-\frac{1}{2}}\,dt}{\big(1-\big(1-\frac{\ell}{m}\big)t\big)^{3/2}} \end{aligned} \tag{7-1}$$

in the region $\mathrm{Re}\, s > -\frac{1}{2}$. In this form, we recognize that the hypergeometric function at right in (7-1) is bounded by ${}_2F_1\big(\frac{3}{2}, 1, \frac{3}{2}\,\big|\, 1-\frac{\ell}{m}\big)$ when $\mathrm{Re}\, s > 0$. To conclude, note that ${}_2F_1\big(\frac{3}{2}, 1, \frac{3}{2}\,\big|\, 1-\frac{\ell}{m}\big) = \frac{m}{\ell}$ by [9, (9.121)]. ☐

Lemma 7.4. *Fix $\epsilon > 0$. In the region $\operatorname{Re} s > 0$, the function $G_{3/2}(s, n, n+\ell)$ defined in (3-2) satisfies*

$$G_{\frac{3}{2}}(s, n, n+\ell) \ll \frac{|s|^{\operatorname{Re} s - 2 + \epsilon}}{(n+\ell)^{\operatorname{Re} s}} e^{-\frac{\pi}{2}|\operatorname{Im} s|} \left(\frac{|s|^{\frac{1}{2}}}{\sqrt{n+\ell}} + \frac{1}{\sqrt{n}} \right).$$

Proof. We begin with the contour integral representation [5, (8.6.12)]

$$\Gamma(1-k, y) e^y = -\frac{y^{-k}}{\Gamma(k)} \cdot \frac{\pi}{2\pi i} \int_C \frac{\Gamma(w+k) y^{-w}}{\sin(\pi w)} \, dw, \tag{7-2}$$

where C is a contour separating the poles of $\Gamma(w+k)$ from those at $w = 0, 1, \ldots$ arising from $1/\sin(\pi w)$. Here we require $k \notin -\mathbb{N}$. For $k > 0$, we may take C as a vertical line with $\operatorname{Re} w = -\epsilon$. We apply (7-2) and the Mellin transform [5, (8.14.4)] to $G_k(s, n, n+\ell)$ to write

$$\begin{aligned} G_k &= -\frac{\pi n^{-k}}{\Gamma(k)} \frac{1}{2\pi i} \int_{(-\epsilon)} \frac{\Gamma(w+k)}{\sin(\pi w) n^w} \left(\int_0^\infty y^{s-1-w} \Gamma(1-k, (n+\ell)y) \frac{dy}{y} \right) dw \\ &= -\frac{\pi}{\Gamma(k)} \frac{1}{2\pi i} \int_{(-\epsilon)} \frac{\Gamma(w+k)\Gamma(s-w-k)\csc(\pi w)}{n^{k+w}(n+\ell)^{s-w-1}(s-w-1)} dw, \end{aligned} \tag{7-3}$$

provided $\operatorname{Re} s > \max(1, k)$ to begin. Shifting the contour of integration to $\operatorname{Re} w = -\max(1, k) - \epsilon$ passes finitely many poles from $\csc(\pi w)$ and gives a meromorphic continuation of G_k to $\operatorname{Re} s > 0$ when $k > 0$.

We now specialize to $k = \frac{3}{2}$. The contour shift in (7-3) to $\operatorname{Re} w = -\frac{3}{2} - \epsilon$ passes a single pole at $w = -1$, with residue

$$\frac{2\Gamma\left(s - \frac{1}{2}\right)}{\sqrt{n}(n+\ell)^s s} \ll \frac{1}{\sqrt{n}(n+\ell)^{\operatorname{Re} s}} |s|^{\operatorname{Re} s - 2} e^{-\frac{\pi}{2}|\operatorname{Im} s|}.$$

Stirling shows that the integrand decays exponentially in $|\operatorname{Im} w|$, for any s. We may therefore truncate the integral to $|\operatorname{Im} w| \leq \frac{1}{2}|\operatorname{Im} s|$. In this range, the estimates $|s - w - k| \asymp |s|$ and $e^{-\frac{\pi}{2}|\operatorname{Im}(s-w)|} \ll e^{\frac{\pi}{2}|\operatorname{Im} w| - \frac{\pi}{2}|\operatorname{Im} s|}$ allow us to extract the s-dependence of the integrand. Hence the shifted integral (7-3) is $O((n+\ell)^{-\operatorname{Re} s - \frac{1}{2}} |s|^{\operatorname{Re} s - \frac{3}{2} + \epsilon} e^{-\frac{\pi}{2}|\operatorname{Im} s|})$, which completes the proof. □

Proof of Proposition 7.1. Lemma 7.2 and the divisor estimates $\sigma_{-2}(\ell) \ll 1$ and $\sigma_{-1}(\ell) \ll \ell^\epsilon$ imply that the terms at right in the first three lines of (6-5) (excluding the term containing $\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot, \bar{s}) \rangle$) are

$$O_{\operatorname{Re} s, \epsilon}\big(\ell^{\frac{3}{2} - \operatorname{Re} s} |s|^{-\frac{3}{2}} + \ell^{\frac{1}{2} - \operatorname{Re} s + \epsilon} |s|^{-\frac{1}{2}} + \ell^{-\operatorname{Re} s + \epsilon}\big). \tag{7-4}$$

By factoring this upper bound in the form $\ell^{-\operatorname{Re} s + \epsilon} |s|^{-\frac{3}{2}} (\ell^{\frac{3}{2}} + \ell^{\frac{1}{2}} |s| + |s|^{\frac{3}{2}})$, we observe that the second summand is always dominated by the first or third term, and may be ignored.

It remains to estimate the three terms in the last two lines of (6-5). We first consider the divisor sum. In the right half-plane $\operatorname{Re} s > \frac{1}{2}$, we bound $|d + \ell/d|^{1 - 2\operatorname{Re} s} \ll \ell^{\frac{1}{2} - \operatorname{Re} s}$, so the divisor sum is $O(\ell^{\frac{1}{2} - \operatorname{Re} s + \epsilon} |s|^{-1})$, which is nondominant. Otherwise, if $\operatorname{Re} s < \frac{1}{2}$, we bound $|d + \ell/d|^{1 - 2\operatorname{Re} s} \ll \ell^{1 - 2\operatorname{Re} s}$, so the full divisor sum is $O(\ell^{1 - 2\operatorname{Re} s + \epsilon} |s|^{-1})$. This term is dominated by the second term of (7-4) when $\operatorname{Re} s > 0$.

We next consider the contribution of the hypergeometric term in (6-5). By Lemma 7.3, Stirling's formula, and then Lemma 7.2, this term is

$$\ll_{\operatorname{Re} s} \ell^{-\operatorname{Re} s}|s|^{-\frac{3}{2}} \sum_{m=1}^{\ell-1} H(\ell-m) r_1(m) \ll_{\operatorname{Re} s,\epsilon} \ell^{1-\operatorname{Re} s+\epsilon}|s|^{-\frac{3}{2}}$$

in the region $\operatorname{Re} s > 0$. Note that this term is dominated by the first error term in (7-4).

Finally, we consider the term in (6-5) involving $G_{3/2}(s, m_2^2, m_1^2)$. By Lemma 7.4, this term is

$$O_{\operatorname{Re} s}\biggl(|s|^{-2+\epsilon} \sum_{m_1^2-m_2^2=\ell} \frac{m_1 m_2}{m_1^{2\operatorname{Re} s}}\biggl(\frac{|s|^{\frac{1}{2}}}{m_1}+\frac{1}{m_2}\biggr)\biggr) \tag{7-5}$$

in the region $\operatorname{Re} s > 0$. The contribution of $1/m_2$ in the parenthetical is

$$\ll_{\operatorname{Re} s} |s|^{-2+\epsilon} \sum_{m_1^2-m_2^2=\ell} \frac{1}{m_1^{2\operatorname{Re} s-1}} \ll_{\operatorname{Re} s} |s|^{-2+\epsilon}\ell^{\epsilon}\bigl(\ell^{\frac{1}{2}-\operatorname{Re} s}+\ell^{1-2\operatorname{Re} s}\bigr),$$

in which we've used that $\sqrt{\ell} \le m_1 \le \ell$ and that the sum has at most $d(\ell)$ terms. Since $m_1 \ge m_2$, the contribution of the other term in the parenthetical of (7-5) is at most $|s|^{1/2}$ times larger. Both upper bounds are majorized by the contribution of the divisor sum in (6-5). □

8. Noncuspidal spectral inner products

To bound $\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot,\bar{s})\rangle$ in vertical strips, we apply the spectral expansion $\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot,\bar{s})\rangle = \Sigma_{\mathrm{disc}}(s) + \Sigma_{\mathrm{cont}}(s)$ computed in Section 6. In the discrete spectrum, Stirling's approximation, dyadic subdivision, Cauchy–Schwarz, and Lemma 6.2 reduce our task to bounding the inner products $\langle \mathcal{V}_{\mathcal{H}}, \mu_j\rangle$. Since Maass cusp forms are orthogonal to Eisenstein series and to norm-squares of theta functions (cf. [27, Remark 2]), this is equivalent to bounding the unregularized inner products $\langle y^{3/2}|\mathcal{H}|^2, \mu_j\rangle$.

While good estimates for inner products of the form $\langle y^k|f|^2, \mu_j\rangle$ are known when f is a holomorphic cusp form or Maass cusp form (at least on average), the noncuspidal nature of $\mathcal{H}$ meters the applicability of prior results. Fortunately, it is possible to modify work of Jutila [19; 20] in the Maass cusp form case to address the case of harmonic Maass forms. Working in a somewhat general setting, we prove the following theorem.

Theorem 8.1. *Fix $f \in H_k^\sharp(\Gamma_0(N))$ with $k \in \frac{1}{2}+\mathbb{Z}$. Let $\mu_j(z)$ be an L^2-normalized Hecke–Maass cusp form of weight 0 on $\Gamma_0(N)$, with spectral type $t_j \in \mathbb{R}$. For all $\epsilon > 0$, we have*

$$\langle y^k|f|^2, \mu_j\rangle \ll \bigl(|t_j|^{2k-1+\epsilon}+|t_j|^{3-2k+\epsilon}\bigr)e^{-\frac{\pi}{2}|t_j|}.$$

Our proof of this follows the general method of [19; 20]. Very roughly, this plan involves two steps:

a. We relate $\langle y^k|f|^2, \mu_j\rangle$, which is an integral over $\Gamma_0(N)\backslash\mathfrak{h}$, to an "unfolded" integral over $\Gamma_\infty\backslash\mathfrak{h}$, by introducing an Eisenstein series as an unfolding object. This technique was developed in [19, §2] for f a level 1 holomorphic or Maass cusp form, and we adapt it to the case of $f \in H_k^\sharp(\Gamma_0(N))$.

b. The unfolded integral can be understood as an integral transform of a sum involving Fourier coefficients of f and μ_j at various cusps. We truncate the sums and integrals and apply estimates for the Fourier coefficients of f and μ_j to bound the truncations.

We remark that [20] also applies the spectral large sieve, to produce a fairly sharp upper bound for the spectral average $\sum_{|t_j|\sim T} |\langle y^k|f|^2, \mu_j\rangle|^2 e^{\pi|t_j|}$ (when f is a Maass cusp form). To simplify parts of our argument when f is noncuspidal, we do not apply the spectral large sieve and instead produce bounds for individual $\langle y^k|f|^2, \mu_j\rangle$. It would be interesting to determine if our growth estimates for $D_\ell(s)$ could be improved by replacing Theorem 8.1 with an appropriate spectral average.

Though not the main focus of this work, we remark that Theorem 8.1 has applications to modular forms of half-integral weight, since $M_k(\Gamma_0(N)) \subset H_k^\sharp(\Gamma_0(N))$. For convenient reference, we present this as a corollary.

Corollary 8.2. *Fix $k \in \frac{1}{2} + \mathbb{Z}$ and $f \in M_k(\Gamma_0(N))$. Let $\mu_j(z)$ be an L^2-normalized Hecke–Maass cusp form of weight 0 on $\Gamma_0(N)$, with spectral type $t_j \in \mathbb{R}$. For all $\epsilon > 0$, we have*

$$\langle y^k|f|^2, \mu_j\rangle \ll \big(|t_j|^{2k-1+\epsilon} + |t_j|^{1+\epsilon}\big)e^{-\frac{\pi}{2}|t_j|}.$$

We remark that Corollary 8.2 improves certain technical results in [22]. In particular, we improve the t_j-dependence of [22, Proposition 14] in any case that our result applies.

8A. ***Jutila's extension of the Rankin–Selberg method.*** The material in this section adapts [19, §2] from $\mathrm{SL}(2,\mathbb{Z})$ to $\Gamma_0(N)$. Let $\phi(z)$ be an L^2 function on $\Gamma_0(N)\backslash\mathfrak{h}$ satisfying $\phi(z) = O(y^{-\delta})$ for some $\delta > 0$ as $y \to \infty$ and let $E_\infty(z,s)$ denote the weight 0 Eisenstein series at the cusp ∞ of $\Gamma_0(N)$. Since $E_\infty(z,s)$ has a simple pole at $s = 1$ with residue $\frac{3}{\pi}\cdot[\Gamma_0(N) : \mathrm{SL}(2,\mathbb{Z})]^{-1} = V_N^{-1}$, we have

$$\iint_{\Gamma_0(N)\backslash\mathfrak{h}} \phi(z)\frac{dx\,dy}{y^2} = V_N \iint_{\Gamma_0(N)\backslash\mathfrak{h}} \phi(z) \lim_{s\to 1^+}(s-1)E_\infty(z,s)\frac{dx\,dy}{y^2}.$$

We now interchange the limit and integral, which can be justified by expanding $E_\infty(z,s)$ in a (rapidly converging) Fourier series and noting that the pole at $s = 1$ appears only within the constant phase. The growth estimate $\phi(z) = O(y^{-\delta})$ gives convergence in this surviving term and justifies the exchange. Then, since $\mathrm{Re}\, s > 1$, the method of unfolding provides

$$\begin{aligned}\iint_{\Gamma_0(N)\backslash\mathfrak{h}} \phi(z)\frac{dx\,dy}{y^2} &= V_N \lim_{s\to 1^+}(s-1)R(\phi,s), \quad \text{with} \\ R(\phi,s) &:= \int_0^\infty\int_0^1 \phi(z)y^{s-1}\frac{dx\,dy}{y},\end{aligned} \tag{8-1}$$

in which $R(\phi,s)$ is the typical Rankin–Selberg transform of ϕ.

We define $R^*(\phi,s) = \zeta^*(2s)R(\phi,s)$ and $R_0^*(\phi,s) = s(s-1)R^*(\phi,s)$, so that (8-1) equals

$$\frac{\pi}{6}V_N \operatorname*{Res}_{s=1} R^*(\phi,s) = \frac{\pi}{6}V_N R_0^*(\phi,1).$$

Note that $R_0^*(\phi, s)$ is entire, in part because $\phi \in L^2$. By the residue theorem,

$$R_0^*(\phi, 1) = \frac{1}{2\pi i} \int_{\mathcal{O}} g(s) \frac{R_0^*(\phi, s)}{s-1} ds,$$

in which $\mathcal{O}$ is a contour encircling $s = 1$ once counterclockwise and $g(s)$ is a rapidly decaying holomorphic function satisfying $g(1) = 1$. We bend $\mathcal{O}$ into a rectangle connecting $a \pm iT$ and $1 - a \pm iT$, then let $T \to \infty$ and use decay in g to render the horizontal components of $\mathcal{O}$ negligible. It follows that

$$\begin{aligned} R_0^*(\phi, 1) &= \frac{1}{2\pi i} \int_{(a)} g(s) \frac{R_0^*(\phi, s)}{s-1} ds - \frac{1}{2\pi i} \int_{(1-a)} g(s) \frac{R_0^*(\phi, s)}{s-1} ds \\ &= \frac{1}{2\pi i} \int_{(a)} \left(\frac{g(s)}{s-1} R_0^*(\phi, s) + \frac{g(1-s)}{s} R_0^*(\phi, 1-s) \right) ds. \end{aligned} \tag{8-2}$$

We now apply the functional equation of the Eisenstein series on $\Gamma_0(N)$ to relate $R^*(\phi, 1-s)$ to a sum of Rankin–Selberg transforms at the other cusps of $\Gamma_0(N)$. This takes the form

$$R^*(\phi, 1-s) = \sum_{\mathfrak{a}} \gamma_{\mathfrak{a}}(s) R^*(\phi_{\mathfrak{a}}, s), \tag{8-3}$$

in which $\gamma_{\mathfrak{a}}(s)$ is an entry of the scattering matrix for $\Gamma_0(N)$ and $\phi_{\mathfrak{a}} = \phi|_{\sigma_{\mathfrak{a}}}$ under the weight 0 slash operator. Exact formulas for $\gamma_{\mathfrak{a}}$ may be obtained by combining [35, Theorem 6.1] and [35, Proposition 4.2]. We have $\gamma_{\mathfrak{a}}(s) = O(1)$ in fixed vertical strips away from poles. By applying (8-3) to (8-2), we conclude that

$$R_0^*(\phi, 1) = \frac{1}{2\pi i} \int_{(a)} \left(\frac{g(s) R_0^*(\phi, s)}{s-1} + \frac{g(1-s)}{s} \sum_{\mathfrak{a}} \gamma_{\mathfrak{a}}(s) R_0^*(\phi_{\mathfrak{a}}, s) \right) ds.$$

In our application, we take $\phi(z) = \phi_j(z) = y^k |f(z)|^2 \overline{\mu_j(z)}$, where μ_j is a Maass cusp form on $\Gamma_0(N)$ with $\|\mu_j\| = 1$. We conclude that

$$\langle F, \mu_j \rangle = \frac{V_N}{12i} \sum_{\mathfrak{a}} \int_{(a)} \left(\delta_{[\mathfrak{a}=\infty]} s g(s) + (s-1) g(1-s) \gamma_{\mathfrak{a}}(s) \right) \zeta^*(2s) R(\phi_{j\mathfrak{a}}, s)\, ds, \tag{8-4}$$

which generalizes [19, (2.10)]. This expression lets us determine $\langle F, \mu_j \rangle$ while only sampling $R(\phi_{j\mathfrak{a}}, s)$ on the line $\operatorname{Re} s = a \gg 1$. We also note that the pole of $\zeta^*(2s)$ at $s = \frac{1}{2}$ is canceled by $R(\phi_{j\mathfrak{a}}, s)$; hence the only poles of the integrand in $\operatorname{Re} s > 0$ are those of $R(\phi_{j\mathfrak{a}}, s)$.

Remark 8.3. Following [19, (2.8)], we take $g(s) = \exp\bigl(1 - \cos \frac{s-1}{B}\bigr)$, for some large $B > 0$. This choice implies $|g(s)| \ll \exp\bigl(-\frac{1}{2}\exp(|\operatorname{Im} s|/B)\bigr)$ in the vertical strip $|\operatorname{Re} s - 1| \leq \pi B/3$. In particular, the contour integral (8-4) converges if $R(\phi_{j\mathfrak{a}}, s)$ grows at most exponentially in $|\operatorname{Im} s|$. This will be established in Remark 8.10.

To bound the Rankin–Selberg transform

$$R(\phi_{j\mathfrak{a}}, s) = \int_0^\infty \int_0^1 y^{s+k} |f_{\mathfrak{a}}(z)|^2 \overline{\mu_{j\mathfrak{a}}(z)} \frac{dx\, dy}{y^2},$$

we represent $f_{\mathfrak{a}}$ and $\mu_{j\mathfrak{a}}$ as Fourier series, as described in (2-2) and (6-2), then execute the x-integral. This expresses $R(\phi_{j\mathfrak{a}}, s)$ as a triple sum over integers (n_1, n_2, n_3) subject to the relation $n_1 - n_2 = n_3$. As in Section 3, we group these terms based on the signs of n_1 and n_2, so that

$$R(\phi_{j\mathfrak{a}}, s) = I^{+}_{j\mathfrak{a}}(s) + I^{-}_{j\mathfrak{a}}(s) + I^{\times}_{j\mathfrak{a}}(s) + I^{0}_{j\mathfrak{a}}(s),$$

denoting the subsums in which (n_1, n_2) are both positive, are both negative, have mixed sign, or contain a zero, respectively. By changing variables to introduce $m := |n_1 - n_2| = |n_3|$ and grouping similar terms, we write

$$\begin{aligned}
I^{+}_{j\mathfrak{a}}(s) &= \sum_{m, n+\varkappa_{\mathfrak{a}}>0} 2\operatorname{Re}\big(c^{+}_{\mathfrak{a}}(n+m)\overline{c^{+}_{\mathfrak{a}}(n)\rho_{j\mathfrak{a}}(m)}\big)\,\varphi^{+}_{j}(m, n+\varkappa_{\mathfrak{a}}, s),\\
I^{-}_{j\mathfrak{a}}(s) &= \sum_{m,n\geq 1} 2\operatorname{Re}\big(c^{-}_{\mathfrak{a}}(n)\overline{c^{-}_{\mathfrak{a}}(n+m)\rho_{j\mathfrak{a}}(m)}\big)\,\varphi^{-}_{j}(m, n-\varkappa_{\mathfrak{a}}, s),\\
I^{\times}_{j\mathfrak{a}}(s) &= \sum_{m=1}^{\infty}\sum_{n=1-\lceil\varkappa_{\mathfrak{a}}\rceil}^{m-1} 2\operatorname{Re}\big(c^{+}_{\mathfrak{a}}(n)\overline{c^{-}_{\mathfrak{a}}(m-n)\rho_{j\mathfrak{a}}(m)}\big)\,\varphi^{\times}_{j}(m, n+\varkappa_{\mathfrak{a}}, s),
\end{aligned}$$

in which the functions φ^{+}_{j}, φ^{-}_{j}, and $\varphi^{\times}_{j}$ are defined by

$$\varphi^{+}_{j}(m,n,s) := \int_0^\infty y^{s+k-\frac{1}{2}} e^{-2\pi(2n+m)y} K_{it_j}(2\pi m y)\frac{dy}{y}, \tag{8-5}$$

$$\varphi^{-}_{j}(m,n,s) := \int_0^\infty y^{s+k-\frac{1}{2}} e^{2\pi(2n+m)y}\Gamma(1-k, 4\pi n y)\Gamma(1-k, 4\pi(n+m)y) K_{it_j}(2\pi m y)\frac{dy}{y}, \tag{8-6}$$

$$\varphi^{\times}_{j}(m,n,s) := \int_0^\infty y^{s+k-\frac{1}{2}} e^{2\pi(m-2n)y}\Gamma(1-k, 4\pi(m-n)y) K_{it_j}(2\pi m y)\frac{dy}{y}. \tag{8-7}$$

Here we have assumed without loss of generality that $\rho_{j\mathfrak{a}}(-m) = \overline{\rho_{j\mathfrak{a}}(m)}$ for Maass cusp forms of weight 0. Lastly, for singular cusps, we define

$$\begin{aligned}
I^{0}_{j\mathfrak{a}}(s) = \sum_{m>0} & 2\operatorname{Re}\big(c^{+}_{\mathfrak{a}}(m)\overline{c^{+}_{\mathfrak{a}}(0)\rho_{j\mathfrak{a}}(m)}\big)\varphi^{+}_{j}(m,0,s)\\
&+ \sum_{m>0} 2\operatorname{Re}\big(c^{+}_{\mathfrak{a}}(m)\overline{c^{-}_{\mathfrak{a}}(0)\rho_{j\mathfrak{a}}(m)}\big)\varphi^{+}_{j}(m,0,s-k+1)\\
&+ \sum_{m>0} 2\operatorname{Re}\big(c^{-}_{\mathfrak{a}}(m)\overline{c^{+}_{\mathfrak{a}}(0)\rho_{j\mathfrak{a}}(m)}\big)\varphi^{\times}_{j}(m,0,s)\\
&+ \sum_{m>0} 2\operatorname{Re}\big(c^{-}_{\mathfrak{a}}(m)\overline{c^{-}_{\mathfrak{a}}(0)\rho_{j\mathfrak{a}}(m)}\big)\varphi^{\times}_{j}(m,0,s-k+1).
\end{aligned} \tag{8-8}$$

For nonsingular cusps, we set $I^{0}_{j\mathfrak{a}}(s) = 0$, as the corresponding summands vanish or otherwise incorporate into $I^{+}_{j\mathfrak{a}}(s)$.

Remark 8.4. These decompositions mirror [19; 20], except that we separate $I^{+}_{j\mathfrak{a}}$ from $I^{-}_{j\mathfrak{a}}$ and introduce $I^{0}_{j\mathfrak{a}}$ to account for noncuspidality. In fact, φ_j^{+} exactly matches an unnamed function from [19, p. 449]. Our functions φ_j^{-} and $\varphi_j^{\times}$ can be viewed as variants of the functions φ_j^{+} and φ_j^{-} from [20, (3.4)], respectively.

8B. ***Representations and estimates for φ_j^{+}, $\varphi_j^{\times}$, and φ_j^{-}.*** We now record some useful information about the functions φ_j^{+}, $\varphi_j^{\times}$, and φ_j^{-}. We first consider φ_j^{+}, leveraging earlier work of Jutila.

Lemma 8.5 [19, §3]. *Define $\lambda = \lambda(m,n) := \sqrt{1-m^2/(2n+m)^2}$ and set $p := s+k-\frac{1}{2}$. The function $\varphi_j^{+}(m,n,s)$ defined in (8-5) is analytic in* $\operatorname{Re} p > 0$ *and may be written in either of the forms*

$$\varphi_j^{+}(m,n,s) = \frac{\sqrt{\pi}\, m^{it_j}\,\Gamma(p+it_j)\Gamma(p-it_j)}{(4\pi)^p(2n+m)^{p+it_j}\,\Gamma\big(p+\frac{1}{2}\big)}(1+\lambda)^{-p-it_j}\,{}_2F_1\Big(p, p+it_j, 2p\,\Big|\,\frac{2\lambda}{1+\lambda}\Big), \tag{8-9}$$

$$\begin{aligned}\varphi_j^{+}(m,n,s) = \frac{2^{-1-2p}\pi^{-p}}{(n(n+m))^{p/2}}\Bigg(&\Big(\frac{1-\lambda}{1+\lambda}\Big)^{\frac{it_j}{2}}\Gamma(-it_j)\Gamma(p+it_j)\,{}_2F_1\Big(p,1-p,1+it_j\,\Big|\,\frac{\lambda-1}{2\lambda}\Big)\\ &+\Big(\frac{1-\lambda}{1+\lambda}\Big)^{-\frac{it_j}{2}}\Gamma(it_j)\Gamma(p-it_j)\,{}_2F_1\Big(p,1-p,1-it_j\,\Big|\,\frac{\lambda-1}{2\lambda}\Big)\Bigg).\end{aligned} \tag{8-10}$$

Proof. These identities are implicit in [19, (3.16)–(3.21)]. □

In the special case $n=0$, we have $\lambda=0$ and (8-9) implies that

$$\varphi_j^{+}(m,0,s) = \frac{\sqrt{\pi}\,\Gamma(p+it_j)\Gamma(p-it_j)}{(4\pi m)^p\,\Gamma\big(p+\frac{1}{2}\big)}, \tag{8-11}$$

which can also be seen directly via [9, 6.621(3)]. For $n \neq 0$, we don't expect simplification but can still produce upper bounds. For example, in the text surrounding [20, (4.5)], Jutila applies (8-9) to produce

$$\varphi_j^{+}(m,n,s) \ll_{\operatorname{Re} p} \frac{|\Gamma(p+it_j)\Gamma(p-it_j)|}{(2n+m)^{\operatorname{Re} p}(1+\lambda)^{\operatorname{Re} p}|\Gamma(p)|}\log(2n+m), \tag{8-12}$$

valid for $\operatorname{Re} p > 0$. An upper bound derived from the representation (8-10) is presented in the following lemma.

Lemma 8.6 (cf. [20, p. 452]). *Fix $t_j \in \mathbb{R}$ and $\epsilon > 0$. Suppose that $\lambda \neq 0$. For any s in a fixed vertical strip away from poles,*

$$\varphi_j^{+}(m,n,s) \ll_\epsilon \frac{|t_j|^{\operatorname{Re} p-1}}{(n(n+m))^{\frac{\operatorname{Re} p}{2}}}\Bigg(1+\Big|\frac{1+|s|^2}{\lambda t_j}\Big|^{1+|\operatorname{Re} p|+\epsilon}\Bigg)\frac{e^{\frac{\pi}{2}|\operatorname{Im} s|}}{e^{\pi|t_j|}}. \tag{8-13}$$

Proof. For $p \notin \mathbb{Z}$ and nonpositive $z \in \mathbb{C}$, consider the integral representation

$${}_2F_1\Big(p,1-p,1+it_j\,\Big|\,z\Big) = \int_B \frac{\Gamma(1+it_j)\Gamma(p+w)\Gamma(1-p+w)\Gamma(-w)}{\Gamma(p)\Gamma(1-p)\Gamma(1+it_j+w)}(-z)^w\,dw, \tag{8-14}$$

in which the contour B separates the poles of $\Gamma(p+w)\Gamma(1-p+w)$ from those of $\Gamma(-w)$ [5, (15.6.6)]. We suppose that $\operatorname{Re} p > 0$ and shift the contour B to the line $\operatorname{Re} w = \operatorname{Re} p + \epsilon$. This shift passes poles and extracts residues at $w = 0, 1, \dots, \lfloor \mathrm{Re}\, p + \epsilon \rfloor$, totaling

$$\sum_{v=0}^{\lfloor \mathrm{Re}\, p+\epsilon \rfloor} \frac{(p)_v (1-p)_v}{v!\,(1+it_j)_v} z^v \ll 1 + \left| \frac{(1+|p|^2)z}{t_j} \right|^{\mathrm{Re}\, p+\epsilon},$$

in which $(\alpha)_v := \Gamma(v+\alpha)/\Gamma(\alpha)$ denotes the Pochhammer symbol. The same upper bound holds for the shifted integral, by Stirling's approximation. We apply this estimate for $z = \frac{\lambda-1}{2\lambda} \ll \lambda^{-1}$, then apply Stirling's approximation to the other factors of (8-10) to complete the case $\mathrm{Re}\, p > 0$. The case $\mathrm{Re}\, p < 0$ then follows using the invariance of (8-14) under $p \leftrightarrow 1-p$. □

We conclude our discussion of φ_j^+ by presenting a uniform upper bound for the size of its residues.

Lemma 8.7. *Fix $t_j \in \mathbb{R}$. For each integer $r \geq 0$, we have*

$$\operatorname*{Res}_{s=\frac{1}{2}-k\pm it_j-r} \varphi_j^+(m,n,s) \ll_r (n+m)^r |t_j|^{-\frac{1}{2}} e^{-\frac{\pi}{2}|t_j|}.$$

Proof. Stirling's approximation and (8-10) give

$$\operatorname*{Res}_{s=\frac{1}{2}-k+it_j-r} \varphi_j^+(m,n,s) \ll_r \frac{(n(m+n))^{\frac{r}{2}}}{|t_j|^{1/2} e^{\frac{\pi}{2}|t_j|}} \cdot \left| {}_2F_1\left(\begin{matrix} it_j - r, 1+r-it_j, \\ 1-it_j \end{matrix} \middle| \frac{\lambda-1}{2\lambda} \right) \right|.$$

The transformation ${}_2F_1(a,b,c,z) = (1-z)^{-a}\,{}_2F_1(a, c-b, c, \frac{z}{z-1})$ (cf. [9, 9.131(1)]) relates the hypergeometric function above to the finite sum

$$\left(\frac{\lambda+1}{2\lambda}\right)^{-it_j+r} {}_2F_1\left(\begin{matrix} it_j - r, -r, \\ 1-it_j \end{matrix} \middle| \frac{1-\lambda}{1+\lambda} \right) \ll \lambda^{-r} \sum_{v=0}^{r} \frac{(it_j-r)_v (-r)_v}{(1-it_j)_v v!} \left(\frac{1-\lambda}{1+\lambda}\right)^v,$$

which is $O_r(\lambda^{-r})$, uniformly in t_j. The claim now follows from the estimate $\lambda^2 \asymp n/(n+m)$, and the computation for $s = \frac{1}{2} - k - it_j - r$ is identical. □

To understand $\varphi_j^\times$ and φ_j^-, we express them as contour integral transforms of φ_j^+. The following lemma consolidates relevant information about $\varphi_j^\times$.

Lemma 8.8. *The function $\varphi_j^\times(m,n,s)$ defined in (8-7) admits meromorphic continuation to $s \in \mathbb{C}$, with poles at $s = -\frac{1}{2} \pm it_j - r$ and $s = \frac{1}{2} - k \pm it_j - r$, for $r \in \mathbb{Z}_{\geq 0}$. If $t_j \in \mathbb{R}$ and $\mathrm{Re}\, s > 0$ away from poles, we have*

$$\begin{aligned} &\varphi_j^\times(m,n,s) \\ &\ll \frac{1}{(m-n)^k\sqrt{m}} \left(\frac{|s-it_j|\cdot|s+it_j|}{m|s|} \right)^{\mathrm{Re}\, s-1} |s|^{-\frac{1}{2}} \left(1 + \left(\frac{m|s|}{(m-n)|s-it_j|\cdot|s+it_j|} \right)^{|k|+\epsilon} \right) e^{-\pi|t_j| + \frac{\pi}{2}|\mathrm{Im}\, s|} \\ &\qquad + \delta_{[\mathrm{Re}\, s < \frac{1}{2}-k]} (m+n+|s|+|t_j|)^A e^{-2\pi|t_j| + \frac{3\pi}{2}|\mathrm{Im}\, s|} \end{aligned} \tag{8-15}$$

for all $\epsilon > 0$ and for some $A > 0$ depending only on k.

Proof. The integral representation (7-2) implies that

$$\varphi_j^{\times}(m,n,s) = \frac{-\pi}{\Gamma(k)} \cdot \frac{1}{2\pi i} \int_C \frac{\Gamma(w+k)}{(4\pi(m-n))^{w+k}\sin(\pi w)} \left(\int_0^\infty y^{s-\frac{1}{2}-w} e^{-2\pi my} K_{it_j}(2\pi my) \frac{dy}{y} \right) dw$$
$$= \frac{-\pi}{\Gamma(k)} \cdot \frac{1}{2\pi i} \int_C \frac{\Gamma(w+k)\varphi_j^{+}(m,0,s-k-w)}{(4\pi(m-n))^{w+k}\sin(\pi w)} dw,$$

where C is a contour separating the poles of $\Gamma(w+k)$ from those at $w=0,1,\dots,$ arising from $1/\sin(\pi w)$. To begin, we require $\operatorname{Re} s > \frac{1}{2} + \max\{\operatorname{Re} w : w \in C\}$. To consider general s, we shift the contour C left, passing poles from $\Gamma(w+k)$ and $1/\sin(\pi w)$ and extracting residues involving $\varphi_j^{+}(m,0,s)$ at shifted arguments. By (8-11), these residues contribute poles at the poles of $\Gamma\big(s+k-\frac{1}{2}\pm it_j\big)$ and $\Gamma\big(s+\frac{1}{2}\pm it_j\big)$.

To produce growth estimates, we then shift C rightwards, to the contour $\operatorname{Re} w = |k|+\epsilon$. This extracts a sum of residues equal to

$$\sum_{q=0}^{|k|-\frac{1}{2}} \frac{(-1)^q \Gamma(q+k)}{\Gamma(k)(4\pi(m-n))^{k+q}} \cdot \varphi_j^{+}(m,0,s-k-q)$$
$$+ \sum_{r=0}^{\lfloor |k|+\epsilon-\operatorname{Re} s+\frac{1}{2}\rfloor} \sum_{\pm} \operatorname*{Res}_{w=s-\frac{1}{2}\pm it_j+r} \frac{\pi\Gamma(w+k)\varphi_j^{+}(m,0,s-k-w)}{(4\pi(m-n))^{w+k}\sin(\pi w)\Gamma(k)}.$$

Stirling's approximation and Lemma 8.7 show that the exponential decay in the residues in the second line is $e^{-\frac{3\pi}{2}|\operatorname{Im} s\pm it_j|-\frac{\pi}{2}|t_j|} \ll e^{-2\pi|t_j|+\frac{3\pi}{2}|\operatorname{Im} s|}$, while the worst polynomial growth is $O((m+n+|s|+|t_j|)^A)$ for some $A>0$ depending linearly on $|k|$ and $\operatorname{Re} s$. Since $\operatorname{Re} s \in \big(0,|k|+\frac{1}{2}\big)$ when these terms appear, we may take the constant A to depend on k alone.

Exponential decay in $|\operatorname{Im} w|$ within the integrand bounds the shifted contour integral to at most a constant multiple of the integrand near $|k|+\epsilon$. Stirling's approximation and (8-11) then complete the proof of (8-15). □

The corresponding properties of φ_j^{-} may be obtained in a similar (though more complicated) way and are summarized in the following lemma.

Lemma 8.9. *The function $\varphi_j^{-}(m,n,s)$ defined in (8-6) admits meromorphic continuation to $s\in\mathbb{C}$, with poles at $s=k-\frac{3}{2}-m\pm it_j$ and $s=\frac{1}{2}-k-m\pm it_j$, for $m\in\mathbb{Z}_{\geq 0}$. If $t_j\in\mathbb{R}$ and $\operatorname{Re} s>3|k|+1$, then for all $\epsilon>0$ we have*

$$\varphi_j^{-}(m,n,s) \tag{8-16}$$
$$\ll \frac{\log(m+n)}{(n(m+n))^{k+\frac{1}{2}}} \left(\frac{|s-it_j||s+it_j|}{(2n+m)(1+\lambda)|s|} \right)^{\operatorname{Re} s-k-1} \left(1 + \left(\frac{(n+m)|s|^2}{n|s-it_j|^2|s+it_j|^2} \right)^{|k|+\epsilon} \right) e^{-\pi|t_j|+\frac{\pi}{2}|\operatorname{Im} s|}.$$

Proof. Using (7-2), we write $\varphi_j^-(m,n,s)$ as a double contour integral,

$$\begin{aligned}\varphi_j^-(m,n,s) &= \frac{\pi^2}{\Gamma(k)^2(2\pi i)^2}\int_{C_1}\int_{C_2}\frac{\Gamma(w_1+k)\Gamma(w_2+k)}{(4\pi n)^{w_1+k}(4\pi(n+m))^{w_2+k}\sin(\pi w_1)\sin(\pi w_2)} \\ &\qquad\times\left(\int_0^\infty y^{s-k-\frac{1}{2}-w_1-w_2}e^{-2\pi(2n+m)y}K_{it_j}(2\pi my)\frac{dy}{y}\right)dw_2\,dw_1 \\ &= \frac{\pi^2}{(2\pi i)^2}\int_{C_1}\int_{C_2}\frac{\Gamma(w_1+k)\Gamma(w_2+k)\varphi_j^+(m,n,s-2k-w_1-w_2)\,dw_2\,dw_1}{\Gamma(k)^2(4\pi n)^{w_1+k}(4\pi(n+m))^{w_2+k}\sin(\pi w_1)\sin(\pi w_2)},\end{aligned}$$

where C_1 and C_2 are instances of the contour C described in Lemma 8.8 and $\operatorname{Re} s > k+\frac{1}{2}+2\max\{\operatorname{Re} w : w\in C\}$ to begin. As in Lemma 8.8, shifting the contours left produces residues which determine the poles of φ_j^-. To produce growth estimates, we shift C_1 and C_2 to the lines $\operatorname{Re} w_1 = \operatorname{Re} w_2 = |k|+\epsilon$, extracting a series of single contour integrals and a double sum of residues from $1/\sin(\pi w_1)\sin(\pi w_2)$ equal to

$$\sum_{q_1,q_2=0}^{|k|-\frac{1}{2}}\frac{(-1)^{q_1+q_2}\Gamma(q_1+k)\Gamma(q_2+k)\varphi_j^+(m,n,s-2k-q_1-q_2)}{\Gamma(k)^2(4\pi n)^{k+q_1}(4\pi(m+n))^{k+q_2}}. \tag{8-17}$$

Exponential decay in vertical strips implies that the contour integrals are bounded by their values near the near axis, whereby the bound (8-12) and Stirling's approximation gives (8-16). □

Remark 8.10. The upper bounds for φ_j^+, $\varphi_j^\times$, and φ_j^- given in (8-12), (8-15), and (8-16) imply that $R(\phi_{j\mathfrak{a}},s)$ satisfies a bound of the form

$$R(\phi_{j\mathfrak{a}},s)\ll(|s|+|t_j|)^A e^{-\frac{\pi}{2}|t_j|+\frac{\pi}{2}|\operatorname{Im} s|}, \tag{8-18}$$

for sufficiently large $\operatorname{Re} s$ and some $A>0$. Indeed, such a bound holds for each of $I_{j\mathfrak{a}}^+(s)$, $I_{j\mathfrak{a}}^\times(s)$, $I_{j\mathfrak{a}}^-(s)$, and $I_{j\mathfrak{a}}^0(s)$, by dyadic subdivision of their defining sums, polynomial growth bounds on $c_{\mathfrak{a}}^\pm(n)$, and a bound for $\rho_{j\mathfrak{a}}(n)$ such as Lemma 6.1.

Note that (8-18) implies that the contour integral (8-4) for $\langle F,\mu_j\rangle$ converges for $\operatorname{Re} s$ sufficiently large. More specifically, it implies that $\langle F,\mu_j\rangle \ll |t_j|^A e^{-\frac{\pi}{2}|t_j|}$ for some $A>0$. These coarse estimates also show that the integral in (8-4) may be truncated to $|\operatorname{Im} s| = c\log(1+|t_j|)$ for some $c>0$ while introducing negligible error. We assume this henceforth.

We conclude this section with an upper bound for φ_j^- obtained via (8-13). We assume $\lambda \gg |t_j|^{-1-\epsilon}$. We also assume that $|s|\ll \log|t_j|$, which holds without loss of generality by Remark 8.10. The bound (8-13) implies that the contribution of the residues (8-17) is

$$O\left(\frac{|t_j|^{\operatorname{Re} s-k-\frac{3}{2}+\epsilon}}{(n(m+n))^{\frac{1}{2}\operatorname{Re} s+\frac{k}{2}-\frac{1}{4}}}e^{-\pi|t_j|+\frac{\pi}{2}|\operatorname{Im} s|}\sum_{q_1,q_2}^{|k|-\frac{1}{2}}\left(\frac{m+n}{n}\right)^{\frac{q_1-q_2}{2}}|t_j|^{-q_1-q_2}\right).$$

Since $\lambda \asymp \sqrt{n}/\sqrt{n+m}$, the estimate $\lambda \gg |t_j|^{-1-\epsilon}$ implies that $|t_j|^{2+2\epsilon}\gg \frac{m+n}{n}$. Thus, up to $|t_j|^\epsilon$ factors, the (q_1,q_2)-sum is dominated by the $q_1=q_2=0$ term. Our estimate for the $q_1=q_2=0$ term likewise acts as a bound for the shifted double contour and any of the single contour integrals associated to residues from $1/\sin(\pi w_1)$ or $1/\sin(\pi w_2)$.

The contribution of the residues from the poles of $\varphi_j^+(m,n,s-2k-w_1-w_2)$ (as either single contour integrals or residues from single contour integrals) is $O\big((m+n+|s|+|t_j|)^A e^{-2\pi|t_j|+\frac{3\pi}{2}|\mathrm{Im}\,s|}\big)$ for some $A>0$, by Lemma 8.7 and Stirling's approximation. (At this level of precision it suffices to consider only the exponential factor in Stirling's approximation.) Thus

$$\varphi_j^-(m,n,s) \ll \frac{|t_j|^{\mathrm{Re}\,s-k-\frac{3}{2}+\epsilon}}{(n(m+n))^{\frac{1}{2}\mathrm{Re}\,s+\frac{k}{2}-\frac{1}{4}}} e^{-\pi|t_j|+\frac{\pi}{2}|\mathrm{Im}\,s|} + \delta_{[\mathrm{Re}\,s<|k-1|-\frac{1}{2}]}(m+n+|s|+|t_j|)^A e^{-2\pi|t_j|+\frac{3\pi}{2}|\mathrm{Im}\,s|}. \tag{8-19}$$

8C. *Sum truncation.* For some (m,n), the functions φ_j^+, $\varphi_j^\times$, and φ_j^- may be made arbitrarily small by taking $\mathrm{Re}\,s$ very large. For example, (8-15) implies that $\varphi_j^\times(m,n,s)$ decays with respect to $|t_j|$ as $\mathrm{Re}\,s\to\infty$ provided $m<|t_j|^{2+\delta}$, for any fixed $\delta>0$. In other words, we may truncate $I_{j\mathfrak{a}}^\times(s)$ to $m\ll|t_j|^{2+\delta}$ in our estimate for $\langle F,\mu_j\rangle$, with a negligible error. Likewise, (8-12) and (8-15) imply that $I_{j\mathfrak{a}}^0(s)$ may be truncated to $m\ll|t_j|^{2+\delta}$.

We claim that $I_{j\mathfrak{a}}^+(s)$ and $I_{j\mathfrak{a}}^-(s)$ may be truncated to $n(m+n)\ll|t_j|^{2+\delta}$ at the cost of negligible error. To prove this, we follow [19, (3.25)] and subdivide cases based on whether $\lambda\ll|t_j|^{-1}$.

a. If $\lambda\ll|t_j|^{-1}$ and $n(m+n)\gg|t_j|^{2+\delta}$, then $\lambda^2\ll|t_j|^{-2}$, so that $(2n+m)^2\gg n(n+m)|t_j|^2$ after simplifying. The lower bound $n(m+n)\gg|t_j|^{2+\delta}$ implies that $2n+m\gg|t_j|^{2+\delta/2}$, hence $n+m\gg|t_j|^{2+\delta/2}$. In this case, (8-12) and (8-16) produce arbitrary polynomial improvements in $|t_j|$ as $\mathrm{Re}\,s\to\infty$.

b. If $\lambda\gg|t_j|^{-1}$ and $n(m+n)\gg|t_j|^{2+\delta}$, we instead argue using the upper bounds (8-13) and (8-19).

Let $J_{j\mathfrak{a}}^+$ and $J_{j\mathfrak{a}}^-$ denote the truncations of $I_{j\mathfrak{a}}^+$ and $I_{j\mathfrak{a}}^-$ to $n(m+n)\ll|t_j|^{2+\delta}$. Likewise, define $J_{j\mathfrak{a}}^\times$ and $J_{j\mathfrak{a}}^0$ as the truncations of $I_{j\mathfrak{a}}^\times$ and $I_{j\mathfrak{a}}^0$ to $m\ll|t_j|^{2+\delta}$.

8D. *Estimation of the truncated sums.* To complete our estimation of the inner product $\langle F,\mu_j\rangle$, we bound the sums $J_{j\mathfrak{a}}^+(s)$, $J_{j\mathfrak{a}}^-(s)$, $J_{j\mathfrak{a}}^\times(s)$, and $J_{j\mathfrak{a}}^0(s)$ on the line $\mathrm{Re}\,s=\delta$, where δ is the same constant used to define the truncation conditions. We assume that $|\mathrm{Im}\,s|=O(\log|t_j|)$, by Remark 8.10.

We first consider $J_{j\mathfrak{a}}^+(s)$, which we truncated to $n(m+n)\ll|t_j|^{2+\delta}$. We subdivide into dyadic intervals, with $m\sim M$ and $n\sim L$. On each dyadic subsum, we estimate $\varphi_j^+(m,n+\varkappa_\mathfrak{a},s)$ using (8-13), which outperforms (8-12) in these regimes. Since $L(L+M)\ll|t_j|^{2+\delta}$ and $\lambda^2\asymp n/(n+m)$, we have $\lambda\gg|t_j|^{-1-\delta}$. This observation, and the free assumption $s=O(\log|t_j|)$, shows that (8-13) bounds a given dyadic sum by

$$|t_j|^{k-\frac{3}{2}+O(\delta)}e^{-\pi|t_j|+\frac{\pi}{2}|\mathrm{Im}\,s|}\sum_{\substack{m\sim M\\ n\sim L}}\frac{|c_\mathfrak{a}^+(n+m)c_\mathfrak{a}^+(n)\rho_{j\mathfrak{a}}(m)|}{(L(L+M))^{\frac{k}{2}-\frac{1}{4}}}. \tag{8-20}$$

In the sum within (8-20), we apply Cauchy–Schwarz, Lemma 2.1, and Lemma 6.1 to compute

$$\sum_{n\sim L}|c_{\mathfrak{a}}^{+}(n)|\sum_{m\sim M}|c_{\mathfrak{a}}^{+}(n+m)\rho_{j\mathfrak{a}}(m)| \ll \sum_{n\sim L}|c_{\mathfrak{a}}^{+}(n)|\biggl(\sum_{q\sim M+L}|c_{\mathfrak{a}}^{+}(q)|^2\biggr)^{\frac{1}{2}}\biggl(\sum_{m\sim M}|\rho_{j\mathfrak{a}}(m)|^2\biggr)^{\frac{1}{2}}$$
$$\ll L^{\frac{1}{2}}\bigl(L(L+M)^{k+|k-1|}\bigr)^{\frac{1}{2}}\bigl(M+|t_j|\bigr)^{\frac{1}{2}}e^{\frac{\pi}{2}|t_j|}. \tag{8-21}$$

Thus the (L,M)-dependence in (8-20) is $L^{\frac{1}{2}}M^{\frac{1}{2}}(L(L+M))^{\frac{1}{4}+\frac{1}{2}|k-1|}$ or $L^{\frac{1}{2}}(L(L+M))^{\frac{1}{4}+\frac{1}{2}|k-1|}$. In the first case, the dominant dyadic interval takes $M\sim|t_j|^2$ and $L\sim 1$, while in the second case we dominate by the $M\sim 1$ and $L\sim|t_j|$ subintervals (up to $|t_j|^{\delta}$ factors). Either way, we conclude that

$$J_{j\mathfrak{a}}^{+}(s)\ll|t_j|^{k+|k-1|+O(\delta)}e^{-\frac{\pi}{2}|t_j|+\frac{\pi}{2}|\operatorname{Im}s|}. \tag{8-22}$$

Our treatment of $J_{j\mathfrak{a}}^{-}(s)$ is essentially the same. We again subdivide into dyadic intervals, with $m\sim M$ and $n\sim L$, then apply (8-19). The contribution from $(m+n+|s|+|t_j|)^{A}e^{-2\pi|t_j|+\frac{3\pi}{2}|\operatorname{Im}s|}$ within $\varphi_j^{-}(m,n,s)$ is clearly $O(|t_j|^{A'}e^{-2\pi|t_j|})$ for some $A'>0$, which will be exponentially nondominant. Otherwise, (8-19) bounds a given dyadic interval by

$$|t_j|^{-k-\frac{3}{2}+O(\delta)}e^{-\pi|t_j|+\frac{\pi}{2}|\operatorname{Im}s|}\sum_{\substack{m\sim M\\ n\sim L}}\frac{|c_{\mathfrak{a}}^{-}(n+m)c_{\mathfrak{a}}^{-}(n)\rho_{j\mathfrak{a}}(m)|}{(L(L+M))^{\frac{k}{2}-\frac{1}{4}}},$$

which matches the $J_{j\mathfrak{a}}^{+}(s)$ case except that we've multiplied by $|t_j|^{-2k}$ and replaced $c_{\mathfrak{a}}^{+}$ with $c_{\mathfrak{a}}^{-}$. By Lemma 2.1, the change $c_{\mathfrak{a}}^{+}\mapsto c_{\mathfrak{a}}^{-}$ does not worsen our estimate. We conclude that

$$J_{j\mathfrak{a}}^{-}(s)\ll|t_j|^{|k-1|-k+O(\delta)}e^{-\frac{\pi}{2}|t_j|+\frac{\pi}{2}|\operatorname{Im}s|}. \tag{8-23}$$

We next consider $J_{j\mathfrak{a}}^{\times}(s)$. By applying (8-15) and disregarding the nondominant contribution of $(m+n+|s|+|t_j|)^{A}e^{-2\pi|t_j|+\frac{3\pi}{2}|\operatorname{Im}s|}$, we find that

$$J_{j\mathfrak{a}}^{\times}(s)\ll|t_j|^{-2+O(\delta)}e^{-\pi|t_j|+\frac{\pi}{2}|\operatorname{Im}s|}\sum_{m<|t_j|^{2+\delta}}\frac{|\rho_{j\mathfrak{a}}(m)|}{m^{-1/2}}\sum_{n=1-\lceil\varkappa_{\mathfrak{a}}\rceil}^{m-1}\frac{|c_{\mathfrak{a}}^{+}(n)c_{\mathfrak{a}}^{-}(m-n)|}{(m-n-\varkappa_{\mathfrak{a}})^k},$$

under the standing assumptions on s. To estimate the sums, we map $n\mapsto m-n$ in the n-sum, restrict m to a dyadic interval $m\sim M$, swap the order of summation, and apply Cauchy–Schwarz and Lemma 2.1:

$$\sum_{m\sim M}\frac{|\rho_{j\mathfrak{a}}(m)|}{m^{-1/2}}\sum_{n=1}^{m}\frac{|c_{\mathfrak{a}}^{+}(m-n)c_{\mathfrak{a}}^{-}(n)|}{(n-\varkappa_{\mathfrak{a}})^k}\ll M^{\frac{1}{2}}\sum_{n\le 2M}\frac{|c_{\mathfrak{a}}^{-}(n)|}{n^k}\biggl(\sum_{m\sim M}|\rho_{j\mathfrak{a}}(m)|^2\biggr)^{\frac{1}{2}}\biggl(\sum_{m\sim M}|c_{\mathfrak{a}}^{+}(m)|^2\biggr)^{\frac{1}{2}}$$
$$\ll M^{\frac{1}{2}k+\frac{1}{2}|k-1|+\frac{1}{2}}(M+|t_j|)^{\frac{1}{2}}e^{\frac{\pi}{2}|t_j|}\sum_{n\le 2M}\frac{|c_{\mathfrak{a}}^{-}(n)|}{n^k}.$$

The remaining n-sum has size $O(M^{\frac{1}{2}-\frac{1}{2}k+\frac{1}{2}|k-1|}\log M)$ by dyadic subdivision, Cauchy–Schwarz, and

Lemma 2.1. The largest overall contribution to $J^{\times}_{j\mathfrak{a}}$ appears when $M \sim |t_j|^{2+\delta}$, which gives the estimate

$$J^{\times}_{j\mathfrak{a}}(s) \ll e^{-\frac{\pi}{2}|t_j|+\frac{\pi}{2}|\mathrm{Im}\, s|} \begin{cases} |t_j|^{2k-1+O(\delta)} & \text{if } k > 1, \\ |t_j|^{3-2k+O(\delta)} & \text{if } k < 1. \end{cases} \tag{8-24}$$

Finally, we consider $J^0_{j\mathfrak{a}}(s)$, which we treat according to the four-part decomposition of $I^0_{j\mathfrak{a}}(s)$ in (8-8). Applying (8-12) and (8-15) and ignoring the contribution of $(m+n+|s|+|t_j|)^A e^{-2\pi|t_j|+\frac{3\pi}{2}|\mathrm{Im}\, s|}$ in (8-15) (since it is nondominant) produces

$$J^0_{j\mathfrak{a}}(s) \ll e^{-\pi|t_j|+\frac{\pi}{2}|\mathrm{Im}\, s|}|t_j|^{O(\delta)} \times \sum_{m \ll |t_j|^{2+\delta}} \biggl(\frac{|c^+_{\mathfrak{a}}(m)\rho_{j\mathfrak{a}}(m)|}{m^{k-\frac{1}{2}}|t_j|^{2-2k}} + \frac{|c^+_{\mathfrak{a}}(m)\rho_{j\mathfrak{a}}(m)|}{m^{\frac{1}{2}}} + \frac{|c^-_{\mathfrak{a}}(m)\rho_{j\mathfrak{a}}(m)|}{m^{k-\frac{1}{2}}|t_j|^{2}} + \frac{|c^-_{\mathfrak{a}}(m)\rho_{j\mathfrak{a}}(m)|}{m^{\frac{1}{2}}|t_j|^{2k}} \biggr).$$

For each term in the parenthetical, we subdivide dyadically on m, then apply Cauchy–Schwarz, Lemma 2.1, and Lemma 6.1. In each term, the largest dyadic contribution has $m \sim |t_j|^{2+\delta}$. The first two terms contribute $O(|t_j|^{k+|k-1|+O(\delta)}e^{\frac{\pi}{2}|t_j|})$, while the last two are $O(|t_j|^{|k-1|-k+O(\delta)}e^{\frac{\pi}{2}|t_j|})$. We conclude that

$$J^0_{j\mathfrak{a}}(s) \ll e^{-\frac{\pi}{2}|t_j|+\frac{\pi}{2}|\mathrm{Im}\, s|} \cdot \begin{cases} |t_j|^{2k-1+O(\delta)} & \text{if } k > 1, \\ |t_j|^{1+O(\delta)} & \text{if } k = \frac{1}{2}, \\ |t_j|^{1-2k+O(\delta)} & \text{if } k < 0. \end{cases} \tag{8-25}$$

By combining the upper bounds derived in this section, we complete our estimation of $\langle F, \mu_j\rangle$ and prove Theorem 8.1:

Proof of Theorem 8.1. We estimate (8-4), truncating the contour to $|\mathrm{Im}\, s| \le c\log(1+|t_j|)$ with negligible error by Remark 8.10. We write $R(\phi_{j\mathfrak{a}}, s) = I^+_{j\mathfrak{a}}(s) + I^-_{j\mathfrak{a}}(s) + I^{\times}_{j\mathfrak{a}}(s) + I^0_{j\mathfrak{a}}(s)$, truncating each term in the decomposition as described in Section 8C. Within the truncated contour, we shift to $\mathrm{Re}\, s = \delta$ (with negligible error) and apply (8-22), (8-23), (8-24), and (8-25) to produce

$$\langle F, \mu_j\rangle \ll \sum_{\mathfrak{a}} e^{-\frac{\pi}{2}|t_j|}\bigl(|t_j|^{2k-1+O(\delta)} + |t_j|^{3-2k+O(\delta)}\bigr) \times \int_{(\delta)} \bigl|\delta_{[\mathfrak{a}=\infty]} s g(s) + (s-1)g(1-s)\gamma_{\mathfrak{a}}(s)\bigr| \cdot |\zeta^*(2s)| e^{\frac{\pi}{2}|\mathrm{Im}\, s|}\, ds.$$

The integral is $O_{\mathfrak{a},\delta}(1)$, and the proof follows by taking δ near 0. □

9. Bounding $D_\ell(s)$ in vertical strips, part II

In Section 7, we proved Proposition 7.1, which reduced the problem of bounding $D_\ell(s)$ to the problem of bounding $\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot, \bar{s})\rangle$. In this section, we estimate the latter to prove the following theorem.

Theorem 9.1. *Fix $\epsilon > 0$ small. In the vertical strip $\mathrm{Re}\, s \in \bigl(\frac{1}{2}+\epsilon, \frac{3}{2}+\epsilon\bigr)$ away from poles of $D_\ell(s)$, we have*

$$D_\ell(s) \ll_\epsilon \ell^\epsilon |s|^\epsilon \cdot \bigl(|s|^{\frac{5}{2}} + \ell^{\frac{1}{4}}|s|^2 + \ell|s|^{-\frac{3}{2}}\bigr)^{\frac{3}{2}-\mathrm{Re}\, s}.$$

The proof follows the decomposition of $\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot, \bar{s})\rangle$ into discrete and continuous spectra.

9A. ***Growth of the discrete spectrum*** Σ_{disc}. For convenience, recall that the discrete spectrum equals

$$\Sigma_{\mathrm{disc}}(s) := \frac{\sqrt{\pi}}{(4\pi\ell)^{s-\frac12}\Gamma(s)} \sum_j \rho_j(\ell)\Gamma\big(s-\tfrac12+it_j\big)\Gamma\big(s-\tfrac12-it_j\big)\langle \mathcal{V}_{\mathcal{H}}, \mu_j\rangle.$$

By the comments at the start of Section 8, we may replace $\mathcal{V}_{\mathcal{H}}$ here with the unregularized form $y^{3/2}|\mathcal{H}(z)|^2$. Then, by Theorem 8.1 and Stirling,

$$\Sigma_{\mathrm{disc}}(s) \ll \ell^{\frac12-\mathrm{Re}\,s}|s|^{\frac12-\mathrm{Re}\,s}e^{\frac{\pi}{2}|\mathrm{Im}\,s|}\sum_j \frac{|\rho_j(\ell)|}{\cosh\frac{\pi}{2}t_j}\cdot|t_j|^{2+\epsilon}|s+it_j|^{\mathrm{Re}\,s-1}|s-it_j|^{\mathrm{Re}\,s-1}e^{-\pi\max(|t_j|,|\mathrm{Im}\,s|)}.$$

Here we have used that $t_j\in\mathbb{R}$ for Maass forms on $\Gamma_0(4)$.

By Lemma 6.2, the mass in the t_j-sum in $\Sigma_{\mathrm{disc}}(s)$ concentrates to within $|t_j|<|\mathrm{Im}\,s|$. Thus

$$\Sigma_{\mathrm{disc}}(s) \ll \ell^{\frac12-\mathrm{Re}\,s}\frac{|s|^{\frac52-\mathrm{Re}\,s+\epsilon}}{e^{\frac{\pi}{2}|\mathrm{Im}\,s|}}\sum_{|t_j|<|\mathrm{Im}\,s|}\frac{|\rho_j(\ell)|}{\cosh\frac{\pi}{2}t_j}|s+it_j|^{\mathrm{Re}\,s-1}|s-it_j|^{\mathrm{Re}\,s-1}. \tag{9-1}$$

Lemma 6.2 implies a short-interval second moment estimate of the form

$$\sum_{X\le|t_j|\le X+1}\frac{|\rho_j(\ell)|^2}{\cosh\pi t_j} \ll_{N,\epsilon} X^{1+\epsilon}\ell^{\epsilon}+\ell^{\frac12+\epsilon}. \tag{9-2}$$

By dividing the range of summation in (9-1) into subintervals of length 1 and applying Cauchy–Schwarz and (9-2) to each subinterval, we find

$$\Sigma_{\mathrm{disc}}(s) \ll_{\mathrm{Re}\,s,\epsilon} \ell^{\frac12-\mathrm{Re}\,s+\epsilon}|s|^{2+\epsilon}\big(|s|^{\mathrm{Re}\,s}+1\big)\big(|s|^{\frac12}+\ell^{\frac14}\big)e^{-\frac{\pi}{2}|\mathrm{Im}\,s|}. \tag{9-3}$$

9B. ***Growth of the continuous spectrum*** Σ_{cont}. Recall that the continuous spectrum equals

$$\Sigma_{\mathrm{cont}} = \frac{V_N}{2}\sum_{\mathfrak{a}}\int_{-\infty}^{\infty}\frac{\varphi_{\mathfrak{a}\ell}\big(\frac12+it\big)\Gamma\big(s-\frac12+it\big)\Gamma\big(s-\frac12-it\big)}{(4\pi\ell)^{s-\frac12}(\pi\ell)^{-it}\Gamma(s)\Gamma\big(\frac12+it\big)}\big\langle \mathcal{V}_{\mathcal{H}}, E_{\mathfrak{a}}\big(\cdot,\tfrac12+it\big)\big\rangle\,dt$$

in $\mathrm{Re}\,s>\frac12$. To bound the growth of $\Sigma_{\mathrm{cont}}(s)$ with respect to $|\mathrm{Im}\,s|$ in this region, we must control the growth of both $\varphi_{\mathfrak{a}\ell}\big(\frac12+it\big)$ and $\big\langle \mathcal{V}_{\mathcal{H}}, E_{\mathfrak{a}}\big(\cdot,\frac12+it\big)\big\rangle$. Sufficient estimates for $\varphi_{\mathfrak{a}\ell}\big(\frac12+it\big)$ appear in (6-3).

To estimate $\big\langle \mathcal{V}_{\mathcal{H}}, E_{\mathfrak{a}}\big(\cdot,\frac12+it\big)\big\rangle$, we apply the Phragmén–Lindelöf convexity principle to $\langle \mathcal{V}_{\mathcal{H}}, E_{\mathfrak{a}}(\cdot,\bar{w})\rangle$, studying the latter outside the critical strip. We prove the following result.

Proposition 9.2. *For all* $\epsilon>0$, $\big\langle \mathcal{V}_{\mathcal{H}}, E_{\mathfrak{a}}\big(\cdot,\frac12+it\big)\big\rangle \ll_\epsilon (1+|t|)^{\frac52+\epsilon}e^{-\frac{\pi}{2}|t|}$.

Proof. To begin, we interpret $\langle \mathcal{V}_{\mathcal{H}}, E_{\mathfrak{a}}(\cdot,\bar{w})\rangle$ via the Rankin–Selberg method. More precisely, we interpret the inner product using Zagier's extension of the Rankin–Selberg method to functions with polynomial growth at cusps, as generalized to congruence subgroups by Gupta [37; 6].

Recall from (4-4) that $\mathcal{V}_{\mathcal{H}}(z)$ differs from $y^{3/2}|\mathcal{H}(z)|^2$ by a linear combination of the functions $E_{\mathfrak{b}}\big(z,\frac32\big)$, $\widetilde{E}_{\mathfrak{b}}(z,1)$, and $y^{1/2}|\theta(z)|^2$. It follows that

$$\mathcal{V}_{\mathcal{H}}(\sigma_{\mathfrak{a}}z) = \psi_{\mathfrak{a}}(y)+O(y^{-M})$$

for all $M > 0$ as $y \to \infty$, in which $\psi_{\mathfrak{a}}(y)$ is a linear combination of $y^{-1/2}$ (from $E_{\mathfrak{b}}(z, \frac{3}{2})$), $\log y$, and y^0 (both from $\widetilde{E}_{\mathfrak{b}}(z, 1)$). We define the Rankin–Selberg transform $R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w)$ by

$$R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w) := \int_0^\infty \int_0^1 y^w \big(\mathcal{V}_{\mathcal{H}}(\sigma_{\mathfrak{a}} z) - \psi_{\mathfrak{a}}(y)\big) \frac{dx\,dy}{y^2}. \tag{9-4}$$

We write $\mathcal{V}_{\mathcal{H}}(z)$ as a Fourier series and execute the x-integral in (9-4), extracting the constant Fourier coefficient. This produces

$$\begin{aligned} R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w) := \int_0^\infty y^{w+\frac{1}{2}} \sum_{n+\varkappa_{\mathfrak{a}}>0} |c_{\mathfrak{a}}^+(n)|^2 e^{-4\pi(n+\varkappa_{\mathfrak{a}})y} \frac{dy}{y} \\ + \int_0^\infty y^{w+\frac{1}{2}} \sum_{n\geq 1} |c_{\mathfrak{a}}^-(n)|^2 \Gamma\big(-\tfrac{1}{2}, 4\pi(n-\varkappa_{\mathfrak{a}})y\big)^2 e^{4\pi(n-\varkappa_{\mathfrak{a}})y} \frac{dy}{y} \\ - \frac{1}{64\pi^2} \int_0^\infty y^{w-\frac{1}{2}} \sum_{n+\varkappa_{\mathfrak{a}}>0} |r_{\mathfrak{a}}(n)|^2 e^{-4\pi(n+\varkappa_{\mathfrak{a}})y} \frac{dy}{y}, \end{aligned}$$

where $\theta|_{\sigma_{\mathfrak{a}}}(z) = \sum_{n\geq 0} r_{\mathfrak{a}}(n) e((n+\varkappa_{\mathfrak{a}})z)$. Note that the constant Fourier coefficients of $E_{\mathfrak{b}}\big(z, \frac{3}{2}\big)$ and $\widetilde{E}_{\mathfrak{b}}(z, 1)$ cancel with corresponding terms in $\psi_{\mathfrak{a}}(y)$ and do not appear above. It follows that

$$\begin{aligned} R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w) = \frac{\Gamma\big(w+\frac{1}{2}\big)}{(4\pi)^{w+\frac{1}{2}}} \sum_{n>-\varkappa_{\mathfrak{a}}} \frac{|c_{\mathfrak{a}}^+(n)|^2}{(n+\varkappa_{\mathfrak{a}})^{w+\frac{1}{2}}} - \frac{\Gamma\big(w-\frac{1}{2}\big)}{4(4\pi)^{w+\frac{3}{2}}} \sum_{n>-\varkappa_{\mathfrak{a}}} \frac{|r_{\mathfrak{a}}(n)|^2}{(n+\varkappa_{\mathfrak{a}})^{w-\frac{1}{2}}} \\ + \sum_{n\geq 1} \frac{|c_{\mathfrak{a}}^-(n)|^2}{(4\pi(n-\varkappa_{\mathfrak{a}}))^{w+\frac{1}{2}}} \int_0^\infty y^{w+\frac{1}{2}} \Gamma\big(-\tfrac{1}{2}, y\big)^2 e^y \frac{dy}{y}. \end{aligned}$$

Lemma 2.1 implies that the two Dirichlet series converge in $\operatorname{Re} w > \frac{3}{2}$. Note that the integral above equals $G_{3/2}(w, 1, 1)$ as defined in (3-2), so by the comments following (3-2), the second line above converges for $\operatorname{Re} w > \frac{3}{2}$.

To estimate the growth of $R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w)$ on the line $\operatorname{Re} w = \frac{3}{2} + \epsilon$, we must quantify the growth of $G_{3/2}(w, 1, 1)$ with respect to $|\operatorname{Im} w|$. This was computed in Lemma 7.4; away from poles, we have

$$G_{\frac{3}{2}}(w, 1, 1) \ll_\epsilon |w|^{\operatorname{Re} w - \frac{3}{2} + \epsilon} e^{-\frac{\pi}{2}|\operatorname{Im} w|}.$$

It follows that $R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w) \ll |w|^{\frac{3}{2}+\epsilon} e^{-\frac{\pi}{2}|\operatorname{Im} w|}$ on the line $\operatorname{Re} w = \frac{3}{2} + \epsilon$.

The estimate $\zeta^*(2-2w) R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, 1-w) \ll \sum_{\mathfrak{b}} \zeta^*(2w) R_{\mathfrak{b}}(\mathcal{V}_{\mathcal{H}}, w)$ (cf. (8-3)) can be used to produce bounds in a left half-plane. In particular, we find $R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w) \ll |w|^{\frac{7}{2}+\epsilon} e^{-\frac{\pi}{2}|\operatorname{Im} w|}$ on $\operatorname{Re} w = -\frac{1}{2} - \epsilon$. The Phragmén–Lindelöf convexity principle then implies

$$R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, \tfrac{1}{2} + it) \ll (1 + |t|)^{\frac{5}{2}+\epsilon} e^{-\frac{\pi}{2}|t|}.$$

for real t. To complete the proof, we note that $R_{\mathfrak{a}}(\mathcal{V}_{\mathcal{H}}, w) = \langle \mathcal{V}_{\mathcal{H}}, E_{\mathfrak{a}}(\cdot, \bar{w})\rangle$ within the critical strip $\operatorname{Re} w \in (0, 1)$ by [15, Proposition A.3]. (The constant Θ defined therein equals 0, since $\psi_{\mathfrak{a}}(y)$ is a linear combination of $\log y$, y^0, and $y^{-1/2}$ for each $\mathfrak{a}$.) $\square$

By Proposition 9.2, (6-3), and Stirling's approximation, we have

$$\Sigma_{\mathrm{cont}}(s) \ll \frac{\ell^{\frac{1}{2}-\mathrm{Re}\,s+\epsilon} e^{\frac{\pi}{2}|\mathrm{Im}\,s|}}{|s|^{\mathrm{Re}\,s-\frac{1}{2}}} \int_{-\infty}^{\infty} |s+it|^{\mathrm{Re}\,s-1} |s-it|^{\mathrm{Re}\,s-1} \frac{(1+|t|)^{\frac{5}{2}+\epsilon}}{e^{\pi \max(|\mathrm{Im}\,s|,|t|)}} dt.$$

The mass of the integral above concentrates in $|t| < |\mathrm{Im}\,s|$; restricting to this range, we find that

$$\begin{aligned}\Sigma_{\mathrm{cont}}(s) &\ll \frac{\ell^{\frac{1}{2}-\mathrm{Re}\,s+\epsilon} |s|^{3-\mathrm{Re}\,s+\epsilon}}{e^{\frac{\pi}{2}|\mathrm{Im}\,s|}} \int_{-|\mathrm{Im}\,s|}^{|\mathrm{Im}\,s|} |s+it|^{\mathrm{Re}\,s-1} |s-it|^{\mathrm{Re}\,s-1} dt \\ &\ll \frac{\ell^{\frac{1}{2}-\mathrm{Re}\,s+\epsilon} |s|^{2+\epsilon}}{e^{\frac{\pi}{2}|\mathrm{Im}\,s|}} \bigl(|s|^{\mathrm{Re}\,s}+1\bigr), \end{aligned} \tag{9-5}$$

at least in the region $\mathrm{Re}\,s > \frac{1}{2}$ (where Σ_{cont} has this one-term description).

9C. *Growth of $D_\ell(s)$*. In $\mathrm{Re}\,s > \frac{1}{2}$, the upper bound for $\Sigma_{\mathrm{cont}}(s)$ from (9-5) is dominated by the upper bound for $\Sigma_{\mathrm{disc}}(s)$ from (9-3). It follows that

$$\frac{\langle \mathcal{V}_{\mathcal{H}}, P_\ell(\cdot,\bar{s})\rangle}{\Gamma\bigl(s+\frac{1}{2}\bigr)} \ll_\epsilon \ell^{\frac{1}{2}-\mathrm{Re}\,s+\epsilon} |s|^{2+\epsilon} \bigl(|s|^{\frac{1}{2}} + \ell^{\frac{1}{4}}\bigr) \tag{9-6}$$

in this region. By combining this estimate with Proposition 7.1 and the convexity principle, we complete our proof of Theorem 9.1.

Proof of Theorem 9.1. For $\mathrm{Re}\,s > \frac{3}{2}$, the upper bound

$$D_\ell(s) \ll \biggl(\sum_{n\ge 1} \frac{H(n)^2}{(n+\ell)^{\mathrm{Re}\,s+\frac{1}{2}}}\biggr)^{\frac{1}{2}} \biggl(\sum_{n\ge 1} \frac{H(n+\ell)^2}{(n+\ell)^{\mathrm{Re}\,s+\frac{1}{2}}}\biggr)^{\frac{1}{2}} \ll \sum_{n\ge 1} \frac{H(n)^2}{n^{\mathrm{Re}\,s+\frac{1}{2}}} \ll 1$$

implies that the result holds on $\mathrm{Re}\,s = \frac{3}{2}+\epsilon$, for $\epsilon > 0$. The result also holds on the line $\mathrm{Re}\,s = \frac{1}{2}+\epsilon$, by Proposition 7.1 and (9-6). The full theorem now follows by the convexity principle. □

10. Applying a truncated Perron formula

To prove our main arithmetic result, Theorem 1.1, we apply a truncated Perron formula to $D_\ell(s)$. Fix $\epsilon > 0$. For X nonintegral, we have

$$\begin{aligned}&\sum_{n\le X} H(n)H(n-\ell) \\ &= \frac{1}{2\pi i}\int_{2+\epsilon-iT}^{2+\epsilon+iT} D_\ell\bigl(s-\tfrac{1}{2}\bigr)\frac{X^s}{s}ds + O\biggl(\frac{X^{2+\epsilon}}{T} + \sum_{n=X/2}^{2X} |H(n)H(n-\ell)| \min\Bigl(1, \frac{X}{T|X-n|}\Bigr)\biggr)\end{aligned} \tag{10-1}$$

by [25, Corollary 5.3]. By Lemma 7.2, the error term in (10-1) is

$$O\biggl(\frac{X^{2+\epsilon}}{T} + X^{1+\epsilon}\sum_{n=X/2}^{2X} \min\Bigl(1, \frac{X}{T|X-n|}\Bigr)\biggr) = O\biggl(\frac{X^{2+\epsilon}}{T}\biggr).$$

To estimate the integral in (10-1), we shift the contour from $\operatorname{Re} s = 2+\epsilon$ to $\operatorname{Re} s = 1+\epsilon$. By Theorem 6.5, this extracts two residues, which total

$$\tfrac{1}{2}X^2 \operatorname*{Res}_{s=\frac{3}{2}} D_\ell(s) + \tfrac{2}{3}X^{\frac{3}{2}} \operatorname*{Res}_{s=1} D_\ell(s).$$

Shifting the truncated contour introduces error terms from horizontal contour integrals, which by Theorem 9.1 are bounded by

$$O\Big(\int_{1+iT+\epsilon}^{2+iT+\epsilon} D_\ell\big(s-\tfrac{1}{2}\big)\frac{X^s}{s}\,ds\Big) \ll \frac{(\ell T)^\epsilon}{T}\int_{1+\epsilon}^{2+\epsilon} \big(T^{\frac{5}{2}} + \ell^{\frac{1}{4}}T^2 + \ell T^{-\frac{3}{2}}\big)^{2-\sigma} X^\sigma\,d\sigma$$
$$\ll (\ell X T)^{\epsilon'}\Big(\frac{X^2}{T} + XT^{\frac{3}{2}} + \ell^{\frac{1}{4}}XT + \ell X T^{-\frac{5}{2}}\Big).$$

Once the contour is shifted to $\operatorname{Re} s = 1+\epsilon$, we separate the contribution of the discrete spectrum $\Sigma_{\mathrm{disc}}(s)$ from the rest of $D_\ell(s)$. The estimates from Proposition 7.1 and (9-5) imply that the non-Σ_{disc} terms contribute

$$O\bigg(\int_{1-iT+\epsilon}^{1+iT+\epsilon} (\ell|s|)^\epsilon\bigg(|s|^2 + \ell^{-\frac{1}{2}} + \frac{\ell}{|s|^{\frac{3}{2}}}\bigg)\frac{X^{1+\epsilon}}{|s|}\,ds\bigg) \ll (\ell X T)^\epsilon\big(X\ell + XT^2\big).$$

To bound the contribution of $\Sigma_{\mathrm{disc}}\big(s-\frac{1}{2}\big)/\Gamma(s)$, we shift the contour farther left, to $\operatorname{Re} s = \epsilon$. This shift introduces an error term (from the horizontal contours), which has size

$$O\big((\ell X T)^\epsilon \cdot (T^{\frac{3}{2}}X + \ell^{\frac{1}{4}}TX + T^2 + \ell^{\frac{1}{4}}T^{\frac{3}{2}})\big),$$

by (9-3) as well as a finite sum of residues equal to

$$\mathfrak{R} := \sum_{|t_j|<T}\bigg(\frac{X^{1+it_j}}{\Gamma(2+it_j)}\operatorname*{Res}_{s=\frac{1}{2}+it_j}\Sigma_{\mathrm{disc}}(s) + \frac{X^{1-it_j}}{\Gamma(2-it_j)}\operatorname*{Res}_{s=\frac{1}{2}-it_j}\Sigma_{\mathrm{disc}}(s)\bigg).$$

The contribution of Σ_{disc} on the contour $\operatorname{Re} s = \epsilon$ is $O((\ell X T)^\epsilon \cdot (\ell T^3 + \ell^{\frac{5}{4}}T^{\frac{5}{2}}))$ by (9-3). Evaluating the residues in $\mathfrak{R}$ and bounding in absolute values gives

$$\mathfrak{R} \ll X\sum_{|t_j|<T}\frac{|\rho_j(\ell)\langle \mathcal{V}_{\mathcal{H}}, \mu_j\rangle|}{|t_j|^2} \ll XT^\epsilon \sum_{|t_j|<T}\frac{|\rho_j(\ell)|}{\cosh\frac{\pi}{2}t_j} \ll XT^{1+\epsilon}\bigg(\sum_{|t_j|<T}\frac{|\rho_j(\ell)|^2}{\cosh \pi t_j}\bigg)^{\frac{1}{2}},$$

in which we've applied Theorem 8.1 and Cauchy–Schwarz. Lemma 6.2 then implies that $\mathfrak{R} \ll_\epsilon X(\ell T)^\epsilon(T^2 + \ell^{\frac{1}{4}}T^{\frac{3}{2}})$.

Putting everything together and omitting obviously nondominant errors, we conclude that

$$\sum_{n\le X} H(n)H(n-\ell)$$
$$= \tfrac{1}{2}X^2\operatorname*{Res}_{s=\frac{3}{2}} D_\ell(s) + \tfrac{2}{3}X^{\frac{3}{2}}\operatorname*{Res}_{s=1} D_\ell(s) + O_\epsilon\bigg((\ell X T)^\epsilon\bigg(\frac{X^2}{T} + X\big(T^2 + \ell^{\frac{1}{4}}T^{\frac{3}{2}} + \ell\big) + \ell T^3 + \ell^{\frac{5}{4}}T^{\frac{5}{2}}\bigg)\bigg).$$

When $\ell \ll X^{2/3}$, these errors are minimized by setting $T = X^{1/3}$, producing a collected error of size $O(X^{\frac{5}{3}+\epsilon})$. In the range $X^{2/3} \ll \ell \ll X$, we choose any $T \in [X/\ell, X^{\frac{2}{5}}\ell^{-\frac{1}{10}}]$, producing a collected error of size $O(X^{1+\epsilon}\ell)$. Using the residue formulas from Theorem 6.5, we conclude that

$$\sum_{n\le X} H(n)H(n-\ell) = \frac{\pi^2 X^2}{252\,\zeta(3)}\big(2\sigma_{-2}\big(\tfrac{\ell}{4}\big) - \sigma_{-2}\big(\tfrac{\ell}{2}\big) + \sigma_{-2}(\ell_o)\big) + O_\epsilon\big(X^{\frac{5}{3}+\epsilon} + X^{1+\epsilon}\ell\big).$$

Theorem 1.1 then follows by assuming $\ell \ll X$ and mapping $X \mapsto X + \ell$.

Remark 10.1. The error terms in Theorem 1.1 may be improved dramatically if the sharp cutoff $n \le X$ is replaced by a smooth cutoff. To this effect, fix a smooth function $w(x)$ with inverse Mellin transform $W(s)$. We have

$$\sum_{n\ge 1} H(n)H(n-\ell)\, w\Big(\frac{n}{X}\Big) = \frac{1}{2\pi i}\int_{(2+\epsilon)} D_\ell\big(s-\tfrac{1}{2}\big)W(s)X^s\,ds,$$

provided both sides converge. If $W(s)$ decays exponentially in $|\operatorname{Im} s|$, we may shift the contour of integration left to $\operatorname{Re} s = 1+\epsilon$ by Theorem 9.1. This extracts two residues, and the shifted contour integral contributes $O((X\ell)^{1+\epsilon})$ by Theorem 9.1. We conclude that

$$\sum_{n\ge 1} H(n)H(n-\ell)w\big(\tfrac{n}{X}\big) = W(2)X^2 \operatorname*{Res}_{s=\frac{3}{2}} D_\ell(s) + W\big(\tfrac{3}{2}\big)X^{\frac{3}{2}} \operatorname*{Res}_{s=1} D_\ell(s) + O_\epsilon((X\ell)^{1+\epsilon}),$$

which offers some evidence in support of the conjecture (1-6).

Acknowledgments

The author was supported by the Additional Funding Programme for Mathematical Sciences, delivered by EPSRC (EP/V521917/1) and the Heilbronn Institute for Mathematical Research. This work was also supported by the Swedish Research Council under grant no. 2021-06594 while the author was in residence at Institut Mittag-Leffler in Djursholm, Sweden.

References

[1] K. Bringmann, A. Folsom, K. Ono, and L. Rolen, *Harmonic Maass forms and mock modular forms: theory and applications*, Amer. Math. Soc. Colloq. Publ. **64**, Amer. Math. Soc., Providence, RI, 2017. MR

[2] J. H. Bruinier and J. Funke, "On two geometric theta lifts", *Duke Math. J.* **125**:1 (2004), 45–90. MR

[3] H. Cohen, *A course in computational algebraic number theory*, Grad. Texts in Math. **138**, Springer, 1993. MR

[4] J.-M. Deshouillers and H. Iwaniec, "Kloosterman sums and Fourier coefficients of cusp forms", *Invent. Math.* **70**:2 (1982), 219–288. MR

[5] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain (editors), "NIST digital library of mathematical functions", electronic reference, Nat. Inst. Standards Tech., 2020, available at http://dlmf.nist.gov. Release 1.0.26.

[6] S. Dutta Gupta, "The Rankin–Selberg method on congruence subgroups", *Illinois J. Math.* **44**:1 (2000), 95–103. MR

[7] D. M. Goldfeld, "The class number of quadratic fields and the conjectures of Birch and Swinnerton-Dyer", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **3**:4 (1976), 624–663. MR

[8] A. Good, "Cusp forms and eigenfunctions of the Laplacian", *Math. Ann.* **255**:4 (1981), 523–548. MR

[9] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 8th ed., Elsevier, Amsterdam, 2015. MR

[10] A. Granville and K. Soundararajan, "The distribution of values of $L(1, \chi_d)$", *Geom. Funct. Anal.* **13**:5 (2003), 992–1028. MR

[11] B. Gross and D. Zagier, "Points de Heegner et dérivées de fonctions L", *C. R. Acad. Sci. Paris Sér. I Math.* **297**:2 (1983), 85–87. MR

[12] H. Heilbronn, "On the class number in imaginary quadratic fields", *Q. J. Math.* **5**:1 (1934), 150–160.

[13] F. Hirzebruch and D. Zagier, "Intersection numbers of curves on Hilbert modular surfaces and modular forms of Nebentypus", *Invent. Math.* **36** (1976), 57–113. MR

[14] J. Hoffstein and T. A. Hulse, "Multiple Dirichlet series and shifted convolutions", *J. Number Theory* **161** (2016), 457–533. MR

[15] T. A. Hulse, C. I. Kuan, D. Lowry-Duda, and A. Walker, "Second moments in the generalized Gauss circle problem", *Forum Math. Sigma* **6** (2018), art. id. e24. MR

[16] M. N. Huxley, "Introduction to Kloostermania", pp. 217–306 in *Elementary and analytic theory of numbers* (Warsaw, 1982), edited by H. Iwaniec, Banach Center Publ. **17**, PWN, Warsaw, 1985. MR

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

[18] H. Iwaniec and E. Kowalski, *Analytic number theory*, Amer. Math. Soc. Colloq. Publ. **53**, Amer. Math. Soc., Providence, RI, 2004. MR

[19] M. Jutila, "The additive divisor problem and its analogs for Fourier coefficients of cusp forms, I", *Math. Z.* **223**:3 (1996), 435–461. MR

[20] M. Jutila, "The additive divisor problem and its analogs for Fourier coefficients of cusp forms, II", *Math. Z.* **225**:4 (1997), 625–637. MR

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

[22] E. M. Kıral, "Subconvexity for half integral weight L-functions", *Math. Z.* **281**:3-4 (2015), 689–722. MR

[23] V. V. Kumaraswamy, "On correlations between class numbers of imaginary quadratic fields", *Acta Arith.* **185**:3 (2018), 211–231. MR

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

[25] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory, I: Classical theory*, Cambridge Stud. Adv. Math. **97**, Cambridge Univ. Press, 2007. MR

[26] W. Müller, "The Rankin–Selberg method for non-holomorphic automorphic forms", *J. Number Theory* **51**:1 (1995), 48–86. MR

[27] P. D. Nelson, "The spectral decomposition of $|\theta|^2$", *Math. Z.* **298**:3-4 (2021), 1425–1447. MR

[28] P. Sarnak, "Integrals of products of eigenfunctions", *Int. Math. Res. Not.* **1994**:6 (1994), 251–260. MR

[29] Á. Serrano Holgado and L. M. Navas Vicente, "The zeta function of a recurrence sequence of arbitrary degree", *Mediterr. J. Math.* **20**:4 (2023), art. id. 224. MR

[30] J.-P. Serre and H. M. Stark, "Modular forms of weight $1/2$", pp. 27–67 in *Modular functions of one variable, VI* (Bonn, Germany, 1976), edited by J.-P. Serre and D. B. Zagier, Lecture Notes in Math. **627**, Springer, 1977. MR

[31] K. D. Shankhadhar and R. K. Singh, "An analogue of Weil's converse theorem for harmonic Maass forms of polynomial growth", *Res. Number Theory* **8**:2 (2022), art. id. 36. MR

[32] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., Oxford Univ. Press, 1986. MR

[33] A. I. Vinogradov and L. A. Takhtadzhyan, "The zeta function of the additive divisor problem and spectral expansion of the automorphic Laplacian", *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* (*LOMI*) **134** (1984), 84–116. In Russian; translated in *J. Soviet Math.* **36**:1 (1987), 57–78. MR

[34] D. Wolke, "Moments of class numbers, III", *J. Number Theory* **4** (1972), 523–531. MR

[35] M. P. Young, "Explicit calculations with Eisenstein series", *J. Number Theory* **199** (2019), 1–48. MR

[36] D. Zagier, "Nombres de classes et formes modulaires de poids $3/2$", *C. R. Acad. Sci. Paris Sér. A-B* **281**:21 (1975), 883–886. MR

[37] D. Zagier, "The Rankin–Selberg method for automorphic functions which are not of rapid decay", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **28**:3 (1981), 415–437. MR

Communicated by Roger Heath-Brown
Received 2024-04-05 Revised 2024-10-21 Accepted 2024-12-23

alexander.walker@ucl.ac.uk *Department of Mathematics, University College London, London WC1E 6BT, United Kingdom*

ALGEBRA AND NUMBER THEORY 19:12 (2025)
https://doi.org/10.2140/ant.2025.19.2471

On two definitions of wave-front sets for p-adic groups

Cheng-Chiang Tsai

The wave-front set for an irreducible admissible representation of a p-adic reductive group is the set of maximal nilpotent orbits which appear in the local character expansion. By a result of Mœglin and Waldspurger, they are also the maximal nilpotent orbits whose associated degenerate Whittaker models are nonzero. However, in the literature there are two versions commonly used, one defining maximality using analytic closure and the other using Zariski closure. We show that these two definitions are inequivalent for $G = \mathrm{Sp}_4$.

1. Introduction

Let F be a finite extension of $\mathbb{Q}_p$ and G be a connected reductive group over F. Write $\mathfrak{g} := \operatorname{Lie} G$. The local character expansion of Howe and of Harish-Chandra [1999, Theorem 16.2] asserts that, for any irreducible admissible $\mathbb{C}$-representation π of $G(F)$, there exist constants $c_{\mathcal{O}}(\pi) \in \mathbb{C}$ indexed by nilpotent $\mathrm{Ad}(G(F))$-orbits $\mathcal{O} \subset \mathfrak{g}(F)$, together with a neighborhood $U = U_\pi$ of $0 \in \mathfrak{g}(F)$, such that the character Θ_π of π satisfies the following identity of distributions on U:

$$(\Theta_\pi \circ \log^*)|_U \equiv \sum_{\mathcal{O}} c_{\mathcal{O}}(\pi) \hat{I}_{\mathcal{O}}|_U. \tag{1}$$

Here $I_{\mathcal{O}}$ is the distribution of integrating a function on $\mathcal{O}$ with any $G(F)$-invariant positive measure, and $\hat{I}_{\mathcal{O}}$ its Fourier transform, namely $\hat{I}_{\mathcal{O}}(f) := I_{\mathcal{O}}(\hat{f})$.

Mœglin and Waldspurger [1987] generalized a result of Rodier [1975] and showed that, for $\mathcal{O} \in \max\{\mathcal{O} : c_{\mathcal{O}}(\pi) \neq 0\}$, the quantity $c_{\mathcal{O}}(\pi)$ with suitable normalization is equal to the dimension of the degenerate Whittaker model for π relative to $\mathcal{O}$. Degenerate Whittaker models are local analogues and necessary conditions for existence of Fourier coefficients for automorphic forms. The set $\max\{\mathcal{O} : c_{\mathcal{O}}(\pi) \neq 0\}$ is therefore of particular interest, and is typically called the *wave-front set*. However, there are two partial orders commonly used in the literature: for two nilpotent $\mathrm{Ad}(G(F))$-orbits $\mathcal{O}_1$ and $\mathcal{O}_2$ the partial order $\mathcal{O}_1 < \mathcal{O}_2$ is defined either (i) if the analytic closure (using the Hausdorff p-adic topology on $\mathfrak{g}(F)$) of $\mathcal{O}_1$ is strictly contained in the analytic closure of $\mathcal{O}_2$, or alternatively (ii) if the Zariski closure of $\mathcal{O}_1$ is strictly contained in the Zariski closure of $\mathcal{O}_2$.

Let us denote by $\mathrm{WF}^{\mathrm{rat}}(\pi) := \max\{\mathcal{O} : c_{\mathcal{O}}(\pi) \neq 0\}$ the set given by the first definition, and by $\mathrm{WF}^{\mathrm{Zar}}(\pi)$ the analogous set given by the second definition. Since the Zariski closure is larger than the analytic closure, we have an obvious inclusion $\mathrm{WF}^{\mathrm{rat}}(\pi) \supseteq \mathrm{WF}^{\mathrm{Zar}}(\pi)$. At the same time, there is the notion

---

The author is supported by NSTC grants 110-2115-M-001-002-MY3, 113-2115-M-001-002 and 113-2628-M-001-012.
MSC2020: primary 22E35; secondary 11F30, 22E50.
Keywords: wave-front sets, p-adic groups.

of geometric wave-front sets: Fix an algebraic closure $\bar{F}$ of F and let $\overline{\mathrm{WF}}^{\mathrm{rat}}(\pi)$ (resp. $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$) be the set of $\mathrm{Ad}(G(\bar{F}))$-orbits in $\mathfrak{g}(\bar{F})$ that meet those in $\mathrm{WF}^{\mathrm{rat}}(\pi)$ (resp. $\mathrm{WF}^{\mathrm{Zar}}(\pi)$). Again we have $\overline{\mathrm{WF}}^{\mathrm{rat}}(\pi) \supseteq \overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$. We note that by [Poonen 2017, Proposition 3.5.75], any $G(F)$-orbit $\mathcal{O} \subset \mathfrak{g}(F)$ is Zariski dense in the $G(\bar{F})$-orbit it sits in. Hence $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$ is equal to the set of maximal geometric orbits that appear in (1). We thank Emile Okada for clarifying this.

The set $\mathrm{WF}^{\mathrm{rat}}(\pi)$ was used in [Mœglin and Waldspurger 1987; Mœglin 1996; Gomez et al. 2021] and many others. On the other hand, $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$ was used in, for example, [Waldspurger 2018]. Both $\mathrm{WF}^{\mathrm{rat}}(\pi)$ and $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$ were discussed in [Ciubotaru et al. 2025], while their main results determine $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$ but not $\overline{\mathrm{WF}}^{\mathrm{rat}}(\pi)$. Nevertheless, in [Jiang et al. 2022] the main conjecture, Conjecture 1.3, is stated for $\mathrm{WF}^{\mathrm{Zar}}(\pi)$ but it seems that the spirit might work for $\mathrm{WF}^{\mathrm{rat}}(\pi)$ as well. Given the abundance of results on the topic, it is desirable to know how/whether $\mathrm{WF}^{\mathrm{rat}}(\pi)$ and $\mathrm{WF}^{\mathrm{Zar}}(\pi)$ (resp. $\overline{\mathrm{WF}}^{\mathrm{rat}}(\pi)$ and $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$) could be different. In fact, the longstanding conjecture about geometric wave-front sets, proposed and proved for GL_n in [Mœglin and Waldspurger 1987], asserted that:

Conjecture 1.1. *For any irreducible admissible representation π of $G(F)$, the set $\overline{\mathrm{WF}}^{\mathrm{rat}}(\pi)$ is a singleton.*

Since $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$ is obviously nonempty, the validity of Conjecture 1.1 for any π is equivalent to the validities of the following two statements:

Conjecture 1.2 (counterexample in [Tsai 2024, Theorem 1.1]). $\overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$ *is a singleton.*

Conjecture 1.3. *We have* $\overline{\mathrm{WF}}^{\mathrm{rat}}(\pi) = \overline{\mathrm{WF}}^{\mathrm{Zar}}(\pi)$ *or equivalently* $\mathrm{WF}^{\mathrm{rat}}(\pi) = \mathrm{WF}^{\mathrm{Zar}}(\pi)$.

As indicated above, the first counterexample for Conjecture 1.1 is a counterexample to Conjecture 1.2. The purpose of this paper is to show that Conjecture 1.3 also has a counterexample, in fact, in the case of split rank 2, which is the smallest absolute rank where Conjecture 1.3 becomes nontrivial.

Let $p \geq 11$ be any prime number, q a power of p with $q \equiv 1 \pmod 4$, and F any nonarchimedean local field with residue field $\mathbb{F}_q$ and fixed uniformizer $\varpi \in F$. Let $G = \mathrm{Sp}_4/_F$ be the group of linear operators on F^4 that preserve the symplectic form

$$\langle \vec{x}, \vec{y} \rangle = x_1 y_4 + x_2 y_3 - x_3 y_2 - x_4 y_1. \tag{2}$$

Denote by $\mathfrak{m}$ the maximal ideal and $\mathfrak{m}^0 = \mathcal{O}_F$ the ring of integers in F. Consider the Moy–Prasad filtration $(G(F)_r)_{r \in (1/2)\mathbb{Z}_{\geq 0}}$ "associated to the Siegel parahoric." It is given by

$$G(F)_n := \left\{ g \in G(F) : g - \mathrm{Id}_4 \in \begin{bmatrix} \mathfrak{m}^n & \mathfrak{m}^n & \mathfrak{m}^n & \mathfrak{m}^n \\ \mathfrak{m}^n & \mathfrak{m}^n & \mathfrak{m}^n & \mathfrak{m}^n \\ \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^n & \mathfrak{m}^n \\ \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^n & \mathfrak{m}^n \end{bmatrix} \right\},$$

$$G(F)_{n+1/2} := \left\{ g \in G(F) : g - \mathrm{Id}_4 \in \begin{bmatrix} \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^n & \mathfrak{m}^n \\ \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^n & \mathfrak{m}^n \\ \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} \\ \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} & \mathfrak{m}^{n+1} \end{bmatrix} \right\} \tag{3}$$

for $n \in \mathbb{Z}_{\geq 0}$. The group $G(F)_1$ is a normal subgroup of $G(F)_{1/2}$, and the quotient may be identified as

$$V := G(F)_{1/2}/G(F)_1 \cong \left\{ \begin{bmatrix} 0 & 0 & b & a \\ 0 & 0 & c & b \\ e & d & 0 & 0 \\ f & e & 0 & 0 \end{bmatrix} : a, b, c \in \mathcal{O}_F/\mathfrak{m},\ d, e, f \in \mathfrak{m}/\mathfrak{m}^2 \right\}.$$

Fix an additive character $\psi : F \to \mathbb{C}^\times$ that is trivial on $\mathfrak{m}$ but nontrivial on $\mathcal{O}_F$. Consider

$$A := \begin{bmatrix} 0 & 0 & 0 & \varpi^{-1} \\ 0 & 0 & \varpi^{-1} & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \in \mathfrak{g}(F). \tag{4}$$

Denote by $\psi_A : V \to \mathbb{C}^\times$ the character $B \mapsto \psi(\mathrm{Tr}(AB))$, and by $\tilde{\psi}_A$ its pullback to $G(F)_{1/2}$. Conjecture 1.3 is disproved by:

Theorem 1.4. *For any irreducible component π of the compact induction*

$$\text{c-ind}_{G(F)_{1/2}}^{G(F)} \tilde{\psi}_A,$$

we have that $\mathrm{WF}^{\mathrm{rat}}(\pi)$ contains two regular nilpotent orbits and also a subregular nilpotent orbit. Consequently $\mathrm{WF}^{\mathrm{Zar}}(\pi)$ contains only the two regular nilpotent orbits.

In fact, the subregular orbit is the unique one not contained in the analytic closure of the previous two regular nilpotent orbits. The representation π is one of the so-called epipelagic representations in [Reeder and Yu 2014]. Prior to this work, similar representations for much higher-rank groups had already been studied in a joint work in progress of Chi-Heng Lo and the author to produce a counterexample to Conjecture 1.2 for split groups (rather than for ramified groups as in [Tsai 2024]). We also remark that in the language of the newer paper [Tsai 2023], we have $\mathrm{WF}^{\mathrm{rat}}(\pi) = \mathrm{WF}^{\mathrm{rat}}(A)$ and the result may well be interpreted as for the wave-front set of $A \in \mathfrak{g}(F)$.

2. Nilpotent orbits

For our $G = \mathrm{Sp}_4$, the subregular nilpotent $\mathrm{Ad}(G(F))$-orbits correspond to partition $[2^2]$ and any such orbit has a representative of the form

$$e_{a,b,c} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ b & a & 0 & 0 \\ c & b & 0 & 0 \end{bmatrix}, \quad a, b, c \in F.$$

Denote by v_i $(1 \leq i \leq 4)$ the i-th coordinate vector of our 4-dimensional symplectic space. The operator $e_{a,b,c}$ defines a nondegenerate quadratic form on $\mathrm{span}(v_1, v_2)$ by

$$(X, Y)_{a,b,c} := \langle X, e_{a,b,c} Y \rangle, \tag{5}$$

where $\langle\cdot,\cdot\rangle$ is as in (2). The $\mathrm{Ad}(G(F))$-orbit of $e_{a,b,c}$ is uniquely determined [Nevins 2011, Proposition 5] by the isomorphism class of the quadratic form $(\cdot,\cdot)_{a,b,c}$. Similarly, a regular nilpotent $\mathrm{Ad}(G(F))$-orbit has a representative of the form

$$n_d=\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&d&0&0\\0&0&-1&0\end{bmatrix},\quad d\in F^\times. \tag{6}$$

The orbit is again uniquely determined by the image of d in $F^\times/(F^\times)^2$. We show that:

Lemma 2.1. *The element $e_{a,b,c}$ lies in the analytic closure of $\mathrm{Ad}(G(F))n_d$ if and only if the quadratic form $(\cdot,\cdot)_{(a,b,c)}$ represents d, namely $(v,v)_{a,b,c}=d$ for some $v\in\mathrm{span}(v_1,v_2)$.*

Proof. Suppose $(v,v)_{a,b,c}=d$ for some $v\in\mathrm{span}(v_1,v_2)$. Then with a change of basis we may assume $(v_2,v_2)_{a,b,c}=d$, i.e., $a=d$. We have (with all hidden entries being 0's) for $e,f\in F,\ h\in F^\times$ that

$$\begin{bmatrix}1&&&\\-e&1&&\\f&&1&\\&f&e&1\end{bmatrix}\begin{bmatrix}0&&&\\1&0&&\\&d&0&\\&&1&0\end{bmatrix}\begin{bmatrix}1&&&\\e&1&&\\-f&&1&\\&-f&-e&1\end{bmatrix}=\begin{bmatrix}0&&&\\1&0&&\\de&d&0&\\2f+de^2&de&1&0\end{bmatrix},$$

$$\begin{bmatrix}h^{-1}&&&\\&1&&\\&&1&\\&&&h\end{bmatrix}\begin{bmatrix}0&&&\\1&0&&\\de&d&0&\\2f+de^2&de&1&0\end{bmatrix}\begin{bmatrix}h&&&\\&1&&\\&&1&\\&&&h^{-1}\end{bmatrix}=\begin{bmatrix}0&&&\\h&0&&\\deh&d&0&\\2fh^2+de^2h^2&deh&h&0\end{bmatrix}.$$

For arbitrarily small h we can choose $e,f\in F$ so that $deh=b,\ 2fh^2+de^2h^2=c$. Hence the above converges to $e_{a,b,c}$ as desired.

Now suppose $e_{a,b,c}$ is in the analytic closure of $\mathrm{Ad}(G(F))n_d$, i.e., there is a sequence $g_i\in G(F)$ such that $\mathrm{Ad}(g_i)^{-1}n_d$ converges to $e_{a,b,c}$. To show that $(\cdot,\cdot)_{a,b,c}$ represents d we follow the method of [Djoković 1981, Theorem 6] for real groups. The quadratic form

$$(X,Y)_{d,g_i}:=\langle X,\mathrm{Ad}(g_i)^{-1}(n_d)Y\rangle=\langle g_iX,n_dg_iY\rangle$$

has to converge to $(X,Y)_{a,b,c}$ on $\mathrm{span}(v_1,v_2)$. Since being isomorphic to a nondegenerate quadratic form over F is an open condition in the space of (not necessarily nondegenerate) quadratic forms, for $i\gg 0$ we have that $(\cdot,\cdot)_{d,g_i}|_{\mathrm{span}(v_1,v_2)}\cong(\cdot,\cdot)_{a,b,c}$. In particular, there exists a 2-dimensional subspace W in F^4 such that the restriction of the form $(X,Y)_d:=\langle X,n_dY\rangle$ is isomorphic to $(X,Y)_{a,b,c}$.

Observe the form $(X,Y)_d$ restricts to a rank-2 hyperbolic form on $\mathrm{span}(v_1,v_3)$. The orthogonal complement of $\mathrm{span}(v_1,v_3)$ under it is $\mathrm{span}(v_2)\oplus\mathrm{span}(v_4)$, where the form has discriminant d on $\mathrm{span}(v_2)$ and has $\mathrm{span}(v_4)$ in its kernel. The subspace W must not intersect $\mathrm{span}(v_4)$; hence its image to $F^4/\mathrm{span}(v_4)\cong\mathrm{span}(v_1,v_2,v_3)$ is again 2-dimensional. Denote by $W^\perp$ the orthogonal complement of W in $\mathrm{span}(v_1,v_2,v_3)$. Since $(\cdot,\cdot)_d|_W\cong(\cdot,\cdot)_{a,b,c}$, we have that

$$(\cdot,\cdot)_d|_{\mathrm{span}(v_1,v_3)}\oplus(\cdot,\cdot)_d|_{\mathrm{span}(v_2)}\cong(\cdot,\cdot)_d|_{\mathrm{span}(v_1,v_2,v_3)}\cong(\cdot,\cdot)_{a,b,c}\oplus(\cdot,\cdot)_d|_{W^\perp}$$

are isomorphic as quadratic spaces. Since $(\cdot,\cdot)_d|_{\mathrm{span}(v_1,v_3)}$ is hyperbolic, it is isomorphic to the direct sum of $(\cdot,\cdot)_d|_{W^\perp}$ and some other 1-dimensional quadratic space. By the cancellation theorem of quadratic spaces [Serre 1973, p. 34, Theorem 4], we then have $(\cdot,\cdot)_{a,b,c}$ is isomorphic to the direct sum of $(\cdot,\cdot)_d|_{\mathrm{span}(v_2)}$ and this 1-dimensional space, i.e., $(\cdot,\cdot)_{a,b,c}$, represents d, as asserted. □

3. Shalika germs and the proof of Theorem 1.4

We normalize our Fourier transforms as

$$\hat{f}(B) := \int_{\mathfrak{g}(F)} \psi(\mathrm{Tr}(AB)) f(A)\, dA,$$

where elements in $\mathfrak{g}(F) = \mathfrak{sp}_4(F)$ are identified as 4×4 matrices as usual, i.e., as in (2). It is known (see the main result of [Kim and Murnaghan 2003], or [Kaletha 2015, (6.1)] for a more direct exhibition) that, for any π in Theorem 1.4, on some sufficiently small neighborhood U of $0 \in \mathfrak{g}(F)$ we have

$$(\Theta_\pi \circ \log^*)|_U \equiv c \cdot \hat{I}_A|_U \tag{7}$$

for some $c \in \mathbb{Q}_{>0}$.

Since $p \geq 11$, the hypotheses needed for [DeBacker 2002, Theorem 2.1.5] are satisfied and it gives the following analogue of (1), the *Shalika germ expansion*:

$$I_A(f) = \sum c_{\mathcal{O}}(A) I_{\mathcal{O}}(f). \tag{8}$$

Here $\mathcal{O}$ runs over nilpotent $\mathrm{Ad}(G(F))$-orbits in $\mathfrak{g}(F)$ as in (1), and f has to be a function of depth $-\frac{1}{2}$; a condition that will be automatically met if $\hat{f}$ is supported in a small enough neighborhood. Comparing (1), (7) and (8), we see that the coefficients in (1) satisfy $c_{\mathcal{O}}(\pi) = c \cdot c_{\mathcal{O}}(A)$. In particular, $c_{\mathcal{O}}(\pi) \neq 0$ if and only if $c_{\mathcal{O}}(A) \neq 0$, and we have $\mathrm{WF}^{\mathrm{rat}}(\pi) = \max\{\mathcal{O} : c_{\mathcal{O}}(A) \neq 0\}$, where the partial order is given by the (analytic) closure relation. Fix $\epsilon \in \mathcal{O}_F^\times$ any nonsquare and

$$e = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -\varpi^{-1}\epsilon & 0 & 0 \\ \epsilon & 0 & 0 & 0 \end{bmatrix}. \tag{9}$$

Theorem 1.4 now follows from:

Proposition 3.1. *The Shalika germ $c_{\mathcal{O}}(A)$ is zero for a regular nilpotent orbit $\mathcal{O}$ if and only if the closure of $\mathcal{O}$ contains e.*

Proposition 3.2. *The Shalika germ $c_{\mathcal{O}}(A)$ is nonzero for the subregular nilpotent orbit $\mathcal{O} = \mathrm{Ad}(G(F))e$.*

Remark 3.3. It might look like there are smart choices behind e and A. In fact, a random choice of A has about $\frac{1}{2}$ probability to work; it secretly needs a certain invariant in $\mathcal{O}_F^\times$ to be a square. Once that is met, Proposition 3.2 will work for any such A and some e it picks out. Our choice merely gives a nicer matrix calculation. The assumption $q \equiv 1 \pmod 4$ is also taken to simplify the exposition and is not essentially needed.

The rest of the section is devoted to the proofs of Propositions 3.1 and 3.2.

Proof of Proposition 3.1. A result of Shelstad [1989], combined with another by Kottwitz [1999, Theorem 5.1] (we thank Alexander Bertoloni Meli for clarifying this), showed that, for a regular nilpotent orbit $\mathcal{O}$, $c_{\mathcal{O}}(A)=0$ if and only if $\mathrm{Ad}(G(F))A$ does not meet the Kostant section associated to any element in $\mathcal{O}$. The theory of the Kostant section also gives that, for any fixed regular $\mathcal{O}$, among the stable orbit of A there is exactly one rational orbit that meets the Kostant section. We have

$$\mathrm{Ad}\left(\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&-1&0&0\\0&0&0&1\end{bmatrix}\right)A=\begin{bmatrix}0&0&0&\varpi^{-1}\\1&0&0&0\\0&-\varpi^{-1}&0&0\\0&0&-1&0\end{bmatrix}$$

is in the Kostant section for $n_{-\varpi^{-1}}$. Since $q\equiv 1 \pmod 4$, we may fix $i:=\sqrt{-1}$ a square root of -1 in $\mathcal{O}_F$. We have

$$\mathrm{Ad}\left(\begin{bmatrix}1&i&0&0\\1&-i&0&0\\0&0&\frac{1}{2}i&\frac{1}{2}\\0&0&-\frac{1}{2}i&\frac{1}{2}\end{bmatrix}\right)A=\begin{bmatrix}0&0&2\varpi^{-1}&0\\0&0&0&2\varpi^{-1}\\0&\frac{1}{2}i&0&0\\-\frac{1}{2}i&0&0&0\end{bmatrix}.$$

Hence

$$\mathrm{Ad}\left(\begin{bmatrix}2\varpi^{-1}&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&\frac{1}{2}\varpi\end{bmatrix}\begin{bmatrix}0&0&0&1\\0&1&0&0\\0&0&1&0\\-1&0&0&0\end{bmatrix}\cdot\begin{bmatrix}1&i&0&0\\1&-i&0&0\\0&0&\frac{1}{2}i&\frac{1}{2}\\0&0&-\frac{1}{2}i&\frac{1}{2}\end{bmatrix}\right)A=\begin{bmatrix}0&0&0&2\varpi^{-2}i\\1&0&0&0\\0&\frac{1}{2}i&0&0\\0&0&1&0\end{bmatrix}$$

is in the Kostant section for $n_{i/2}$. We note that both -1 and $\frac{1}{2}i$ are squares in $\mathcal{O}_F^{\times}$, and thus Lemma 2.1 shows that $n_{-\varpi^{-1}}$ and $n_{i/2}$ are exactly the two regular nilpotent orbits whose closure does not contain e. This shows that if $c_{\mathcal{O}}(A)=0$ for a regular nilpotent $\mathcal{O}$, then the closure of $\mathcal{O}$ must contain e. It remains to show that for any regular nilpotent orbit $\mathcal{O}$ different from that of $n_{-\varpi^{-1}}$ and $n_{i/2}$, we have $c_{\mathcal{O}}(A)=0$.

Let $\sqrt{\epsilon}$ be a square root of ϵ in an unramified quadratic extension of F. The element

$$d:=\begin{bmatrix}\sqrt{\epsilon}^{-1}&0&0&0\\0&\sqrt{\epsilon}^{-1}&0&0\\0&0&\sqrt{\epsilon}&0\\0&0&0&\sqrt{\epsilon}\end{bmatrix}$$

has image in $G_{ad}(F)=\mathrm{PSp}_4(F)$. Since the orbit of A meets the Kostant section for $n_{-\varpi^{-1}}$ and $n_{i/2}$, the orbit of $\mathrm{Ad}(d)A$ meets the Kostant section of $\mathrm{Ad}(d)n_{-\varpi^{-1}}$ and $\mathrm{Ad}(d)n_{i/2}$. As $\mathrm{Ad}(d)n_{-\varpi^{-1}}=n_{-\epsilon\varpi^{-1}}$ and $\mathrm{Ad}(d)n_{i/2}=n_{\epsilon i/2}$ are the other two regular nilpotent orbits, using results of Shelstad and Kottwitz and the classical result that a Kostant section meets an $\mathrm{Ad}(G(\bar{F}))$-orbit at one point, it remains to prove that $\mathrm{Ad}(d)A$ and A live in different $\mathrm{Ad}(G(F))$-orbits. The element d defines a class $\alpha_d\in Z^1(F,Z(G))$ and the assertion that $\mathrm{Ad}(d)A$ and A live in different $\mathrm{Ad}(G(F))$-orbits is equivalent to the fact that the image of α_d in $H^1(F,Z_G(A))$ is nontrivial. Observe that α_d is trivial on inertia and sends Frob to

$-1 \in \mu_2 = Z(G)$. Since $Z_G(A)$ is anisotropic over the maximal unramified extension of F, the image of α_d is nontrivial in $H^1(F, Z_G(A)) = H^1(\mathrm{Frob}, X_*(Z_G(A))_{I_F}) = H^1(\mathrm{Frob}, \mu_2^2)$, as claimed. □

Proof of Proposition 3.2. Consider the characteristic function of the set

$$\left\{ X \in \mathfrak{g}(F) : X \text{ is of the form } \begin{bmatrix} \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^{-1} & \mathfrak{m}^{-1} \\ \mathfrak{m}^0 & \mathfrak{m}^0 & \varpi^{-1}\epsilon+\mathfrak{m}^0 & \mathfrak{m}^{-1} \\ \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 \\ \epsilon+\mathfrak{m} & \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 \end{bmatrix} \right\}.$$

Call this function f. It has the property that $f(X+Y) = f(X)$ whenever Y is of the form

$$\begin{bmatrix} \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^{-1} & \mathfrak{m}^{-1} \\ \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^{-1} \\ \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 \\ \mathfrak{m} & \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 \end{bmatrix}.$$

The set of elements of the above form is a Moy–Prasad lattice of depth $-\frac{1}{2}$. Since A is of depth $-\frac{1}{2}$, [DeBacker 2002, Theorem 2.1.5] (or its application to Conjecture 2 of that work) shows that (8) holds for f. Let e be as in (9). We claim that:

Lemma 3.4. *Suppose $I_{\mathcal{O}}(f) \neq 0$ for a nilpotent $\mathrm{Ad}(G(F))$-orbit $\mathcal{O}$. Then e lies in the closure of $\mathcal{O}$.*

Lemma 3.5. $I_A(f) \neq 0$.

With both lemmas, (8) gives $\sum_{\mathcal{O}} c_{\mathcal{O}}(A) I_{\mathcal{O}}(f) = I_A(f) \neq 0$. By Proposition 3.1 and Lemma 3.4, the only nilpotent orbit $\mathcal{O}$ that can contribute to the sum is $\mathcal{O} = \mathrm{Ad}(G(F))e$, which proves Proposition 3.2. □

Proof of Lemma 3.4. We have,

$$\text{for } w = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathrm{Ad}(w)\,\mathrm{supp}(f) = \begin{bmatrix} \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^{-1} & \mathfrak{m}^{-1} \\ \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^{-1} \\ \mathfrak{m}^0 & -\varpi^{-1}\epsilon+\mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 \\ \epsilon+\mathfrak{m} & \mathfrak{m}^0 & \mathfrak{m}^0 & \mathfrak{m}^0 \end{bmatrix}.$$

For any $X \in \mathcal{O} \cap \mathrm{Ad}(w)\,\mathrm{supp}(f)$, we observe that

$$\varpi^{2n}\,\mathrm{Ad}\left(\begin{bmatrix} \varpi^n & 0 & 0 & 0 \\ 0 & \varpi^n & 0 & 0 \\ 0 & 0 & \varpi^{-n} & 0 \\ 0 & 0 & 0 & \varpi^{-n} \end{bmatrix}\right) X \in \mathcal{O} \cap \begin{bmatrix} \mathfrak{m}^{2n} & \mathfrak{m}^{2n} & \mathfrak{m}^{4n-1} & \mathfrak{m}^{4n-1} \\ \mathfrak{m}^{2n} & \mathfrak{m}^{2n} & \mathfrak{m}^{4n} & \mathfrak{m}^{4n-1} \\ \mathfrak{m}^0 & -\varpi^{-1}\epsilon+\mathfrak{m}^0 & \mathfrak{m}^{2n} & \mathfrak{m}^{2n} \\ \epsilon+\mathfrak{m} & \mathfrak{m}^0 & \mathfrak{m}^{2n} & \mathfrak{m}^{2n} \end{bmatrix}.$$

As n goes to $+\infty$, such elements converge (or a subsequence does) to an element

$$e' \in \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \mathfrak{m}^0 & -\varpi^{-1}\epsilon+\mathfrak{m}^0 & 0 & 0 \\ \epsilon+\mathfrak{m} & \mathfrak{m}^0 & 0 & 0 \end{bmatrix}.$$

This element e' lives in the same $\mathrm{Ad}(G(F))$-orbit as e because the bottom-left 2×2 matrix defines an isomorphic quadratic form. □

Proof of Lemma 3.5. We look at

$$\mathrm{Ad}\left(\begin{bmatrix}1&0&0&0\\z&1&0&0\\0&0&1&0\\0&0&-z&1\end{bmatrix}\right)A=\begin{bmatrix}0&0&\varpi^{-1}z&\varpi^{-1}\\0&0&\varpi^{-1}(1+z^2)&\varpi^{-1}z\\1&0&0&0\\-2z&1&0&0\end{bmatrix},$$

$$\mathrm{Ad}\left(\begin{bmatrix}x&0&0&0\\0&y^{-1}&0&0\\0&0&y&0\\0&0&0&x^{-1}\end{bmatrix}\cdot\begin{bmatrix}1&0&0&0\\z&1&0&0\\0&0&1&0\\0&0&-z&1\end{bmatrix}\right)A=\begin{bmatrix}0&0&\varpi^{-1}xy^{-1}z&\varpi^{-1}x^2\\0&0&\varpi^{-1}y^{-2}(1+z^2)&\varpi^{-1}xy^{-1}z\\x^{-1}y&0&0&0\\-2x^{-2}z&x^{-1}y&0&0\end{bmatrix}.$$

Suppose $x, y\in\mathcal{O}_F^\times$ and $z\in\mathcal{O}_F$. Denote by $\bar{x},\bar{y},\bar{z},\bar{\epsilon}\in\mathbb{F}_q$ the respective reductions. The right-hand side of the last equation lies in the support of f if and only if

$$\begin{cases}-2\bar{x}^{-2}\bar{z}=\bar{\epsilon},\\ \bar{y}^{-2}(1+\bar{z}^2)=\bar{\epsilon}\end{cases}\iff\begin{cases}\bar{z}=-\frac{1}{2}\bar{\epsilon}\bar{x}^2,\\ \frac{1}{4}\bar{\epsilon}^2\bar{x}^4+1=\bar{\epsilon}\bar{y}^2.\end{cases}$$

That is, as long as the curve $E=\left(\bar{\epsilon}\bar{y}^2=\frac{1}{4}\bar{\epsilon}^2\bar{x}^4+1:(\bar{x},\bar{y})\in(\mathbb{G}_m)^2/_{\mathbb{F}_q}\right)$ has an $\mathbb{F}_q$-point, there exists $g\in G(F)$ such that $\mathrm{Ad}(g)A\in\mathrm{supp}(f)$, i.e., $I_A(f)\neq 0$. Such an $\mathbb{F}_q$-point always exists. Indeed, E differs from its smooth completion E^c by eight $\overline{\mathbb{F}}_q$-points (two for $\bar{x}=0$, four for $\bar{y}=0$ and two at infinity), and none of them is defined over $\mathbb{F}_q$ because $\bar{\epsilon}\in\mathbb{F}_q$ is a nonsquare. Hence $E(\mathbb{F}_q)=E^c(\mathbb{F}_q)$, while E^c is a geometrically connected projective smooth genus-1 curve and always has an $\mathbb{F}_q$-point. □

Remark 3.6. This "none of the boundary points is defined over the residue field" phenomenon seems to be related to the vanishing of $c_{\mathcal{O}}(A)$ for those $\mathcal{O}>\mathrm{Ad}(G(F))e$.

Remark 3.7. Using a special case of [Kim and Murnaghan 2003, Theorem 2.3.1] that $\mathrm{Ad}(g)A\in\mathfrak{g}(F)_{-1/2}\implies g\in G(F)_0$, one may reduce the computation of orbital integrals and thus $c_{\mathcal{O}}(A)$ and $c_{\mathcal{O}}(\pi)$ (for $\mathcal{O}=\mathrm{Ad}(G(F))e$) to $\#E(\mathbb{F}_q)$. We predict the dimension of the associated degenerate Whittaker model to be $\frac{1}{4}\#E(\mathbb{F}_q)$, analogous to [Tsai 2017, Theorem 4.10 and Corollary 6.2].

Remark 3.8. We may also work with representations of depth $n+\frac{1}{2}$ by replacing A by $\varpi^{-n}A$ and replacing $G(F)_{1/2}$ by $G(F)_{n+1/2}$ in Theorem 1.4. The same proof works, except that e needs to be replaced by $\varpi^{-n}e$, resulting in every $\mathcal{O}\in\mathrm{WF}^{\mathrm{rat}}(\pi)$ being replaced by $\varpi^{-n}\mathcal{O}$.

4. Langlands parameters

The determination of the Langlands parameter corresponding to an individual π in Theorem 1.4 is part of the difficult problem solved in [Kaletha 2015] with deep insight into the rectifying characters and their relation with transfer factors. The *collection* of all Langlands parameters corresponding to components π in Theorem 1.4 is nevertheless simpler, because it happens in this case that the rectifying characters can be absorbed into the choice of an irreducible component in $\text{c-ind}_{G(F)_{1/2}}^{G(F)}\tilde{\psi}_A$. We describe the collection of such Langlands parameters, in the hope that it may be useful to interested readers.

Consider the ramified quadratic extension $E = F(\sqrt{\varpi})$. Write $\varpi_E = \sqrt{\varpi}$. A Langlands parameter we seek for is a homomorphism $\rho : W_F \to \mathrm{SO}_5(\mathbb{C})$. It has image in $\mathrm{O}_2(\mathbb{C}) \times \mathrm{O}_2(\mathbb{C}) \times \mathrm{SO}_1(\mathbb{C})$, i.e., ρ can be viewed as the sum of two orthogonal self-dual representations and a trivial representation. We have $\rho = \rho_1 \oplus \rho_2 \oplus \mathrm{triv}$, where $\rho_j = \mathrm{Ind}_{W_E}^{W_F} \chi_j$ for $j = 1, 2$. Write $\alpha_1 = 1$ and $\alpha_2 = \sqrt{-1}$ for any choice of square root of -1 in $\mathcal{O}_F^\times$. Then χ_j is a character on $E^\times$ satisfying:

(a) $\chi_j|_{F^\times} \equiv 1$.

(b) $\chi_j(1 + x\varpi) = 1$ for all $x \in \mathcal{O}_E$.

(c) $\chi_j(1 + x\varpi_E) = \psi(2x\alpha_j)$ for all $x \in \mathcal{O}_E$.

Here ψ is as chosen before (4). We note that each χ_j is determined up to a freedom of $\chi_j(\varpi_E) \in \{\pm 1\}$, and consequently there are $2^2 = 4$ candidates for such ρ. Relatedly, there are also 2^2 components of π in Theorem 1.4.

Acknowledgments

I wish to thank Chi-Heng Lo, Lei Zhang, Emile Okada and Alexander Bertoloni Meli for very helpful conversations. I am grateful to the National University of Singapore for a wonderful environment and their hospitality during a visit. I am thankful to the referee for helpful suggestions. Lastly I thank ChatGPT for polishing some English writing.

References

[Ciubotaru et al. 2025] D. Ciubotaru, L. Mason-Brown, and E. Okada, "Wavefront sets of unipotent representations of reductive p-adic groups, II", *J. Reine Angew. Math.* **823** (2025), 191–253. MR Zbl

[DeBacker 2002] S. DeBacker, "Homogeneity results for invariant distributions of a reductive p-adic group", *Ann. Sci. École Norm. Sup.* (4) **35**:3 (2002), 391–422. MR Zbl

[Djoković 1981] D. Ž. Djoković, "Closures of conjugacy classes in classical real linear Lie groups", pp. 63–83 in *Algebra: Lie algebras, group theory, and partially ordered algebraic structures* (Carbondale, IL, 1980), edited by R. K. Amayo, Lecture Notes in Math. **848**, Springer, 1981. MR Zbl

[Gomez et al. 2021] R. Gomez, D. Gourevitch, and S. Sahi, "Whittaker supports for representations of reductive groups", *Ann. Inst. Fourier (Grenoble)* **71**:1 (2021), 239–286. MR Zbl

[Harish-Chandra 1999] Harish-Chandra, *Admissible invariant distributions on reductive p-adic groups*, Univ. Lect. Ser. **16**, Amer. Math. Soc., Providence, RI, 1999. MR Zbl

[Jiang et al. 2022] D. Jiang, D. Liu, and L. Zhang, "Arithmetic wavefront sets and generic L-packets", preprint, 2022. arXiv 2207.04700

[Kaletha 2015] T. Kaletha, "Epipelagic L-packets and rectifying characters", *Invent. Math.* **202**:1 (2015), 1–89. MR Zbl

[Kim and Murnaghan 2003] J.-L. Kim and F. Murnaghan, "Character expansions and unrefined minimal K-types", *Amer. J. Math.* **125**:6 (2003), 1199–1234. MR Zbl

[Kottwitz 1999] R. E. Kottwitz, "Transfer factors for Lie algebras", *Represent. Theory* **3** (1999), 127–138. MR Zbl

[Mœglin 1996] C. Mœglin, "Front d'onde des représentations des groupes classiques p-adiques", *Amer. J. Math.* **118**:6 (1996), 1313–1346. MR Zbl

[Mœglin and Waldspurger 1987] C. Mœglin and J.-L. Waldspurger, "Modèles de Whittaker dégénérés pour des groupes p-adiques", *Math. Z.* **196**:3 (1987), 427–452. MR Zbl

[Nevins 2011] M. Nevins, "On nilpotent orbits of SL_n and Sp_{2n} over a local non-Archimedean field", *Algebr. Represent. Theory* **14**:1 (2011), 161–190. MR Zbl

[Poonen 2017] B. Poonen, *Rational points on varieties*, Grad. Stud. in Math. **186**, Amer. Math. Soc., Providence, RI, 2017. MR Zbl

[Reeder and Yu 2014] M. Reeder and J.-K. Yu, "Epipelagic representations and invariant theory", *J. Amer. Math. Soc.* **27**:2 (2014), 437–477. MR Zbl

[Rodier 1975] F. Rodier, "Modèle de Whittaker et caractères de représentations", pp. 151–171 in *Non-commutative harmonic analysis* (Marseille/Luminy, 1974), edited by J. Carmona et al., Lecture Notes in Math. **466**, Springer, 1975. MR Zbl

[Serre 1973] J.-P. Serre, *A course in arithmetic*, Grad. Texts in Math. **7**, Springer, 1973. MR Zbl

[Shelstad 1989] D. Shelstad, "A formula for regular unipotent germs", pp. 275–277 in *Orbites unipotentes et représentations*, *II*, Astérisque **171-172**, Soc. Math. France, Paris, 1989. MR Zbl

[Tsai 2017] C.-C. Tsai, "A formula for certain Shalika germs of ramified unitary groups", *Compos. Math.* **153**:1 (2017), 175–213. MR Zbl

[Tsai 2023] C.-C. Tsai, "Wave-front sets for p-adic Lie algebras", preprint, 2023. arXiv 2311.08078

[Tsai 2024] C.-C. Tsai, "Geometric wave-front set may not be a singleton", *J. Amer. Math. Soc.* **37**:1 (2024), 281–304. MR Zbl

[Waldspurger 2018] J.-L. Waldspurger, "Représentations de réduction unipotente pour $\mathrm{SO}(2n+1)$, III: Exemples de fronts d'onde", *Algebra Number Theory* **12**:5 (2018), 1107–1171. MR Zbl

Communicated by Wee Teck Gan
Received 2024-04-15 Revised 2024-08-23 Accepted 2024-10-21

chchtsai@gate.sinica.edu.tw

Institute of Mathematics, Academia Sinica, Taipei, Taiwan

Department of Applied Mathematics, National Sun Yat-Sen University, Kaohsiung, Taiwan

Department of Mathematics, National Taiwan University, Taipei, Taiwan

ALGEBRA AND NUMBER THEORY 19:12 (2025)
https://doi.org/10.2140/ant.2025.19.2481

Irregular Hodge filtration of hypergeometric differential equations

Yichen Qin and Daxin Xu

Fedorov and Sabbah–Yu calculated the (irregular) Hodge numbers of hypergeometric connections. In this paper, we study the irregular Hodge filtrations on hypergeometric connections defined by rational parameters and provide a new proof of the aforementioned results. Our approach is based on a geometric interpretation of hypergeometric connections, which enables us to show that certain hypergeometric sums are everywhere ordinary on $|\mathbb{G}_{m,\mathbb{F}_p}|$; i.e., "Frobenius Newton polygon equals the irregular Hodge polygon".

1. Introduction

Our primary focus is to investigate the Hodge theoretic properties of confluent hypergeometric differential equations. These differential equations have irregular singularities and are equipped with *irregular Hodge filtrations*, which are defined in [Sabbah 2018]. The irregular Hodge theory, initiated by Deligne [2007a; 2007b], extends the classical Hodge theory and has been developed in a series of works; see [Sabbah 2010; Kontsevich and Soibelman 2011; Yu 2014; Esnault et al. 2017; Sabbah and Yu 2015; Sabbah 2018].

Let $n \geq m$ be two nonnegative integers, λ a real number, and $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_m)$ two nondecreasing sequences of real numbers in $[0, 1)$. Let S be the scheme $\mathbb{G}_m \backslash \{1\}$ (resp. $\mathbb{G}_m$) if $n = m$ (resp. $n > m$) with coordinate z. The *hypergeometric equation* is the linear differential equation defined by the differential operator

$$\mathrm{Hyp}_\lambda(\alpha; \beta) := \lambda \prod_{i=1}^{n} (z\partial_z - \alpha_i) - z \prod_{j=1}^{m} (z\partial_z - \beta_j). \tag{1.0.0.1}$$

The *hypergeometric connection* $\mathcal{H}yp_\lambda(\alpha; \beta)$ is the associated connection on the complex algebraic variety $S_{\mathbb{C}}$; see (2.1.1.1). We say that the pair (α, β) is *nonresonant* if $\alpha_i \neq \beta_j$ for any i and j. In this case, the hypergeometric connection $\mathcal{H}yp_\lambda(\alpha; \beta)$ is irreducible and rigid, as seen by combining the works [Beukers and Heckman 1989] and [Katz 1990].

When $n = m$, hypergeometric connections have regular singularities at 0, 1, and ∞. Simpson [1990, Corollary 8.1] demonstrated that rigid irreducible connections on curves with regular singularities whose eigenvalues of monodromy actions at singularities have norm 1 underlie complex variations of Hodge structure. In this case, Fedorov [2018] computed the Hodge numbers associated with the Hodge filtrations of irreducible hypergeometric connections, and Martin [2021] gave an alternative proof.

MSC2020: primary 14D07; secondary 11T23, 14F30, 14F40, 33C15.
Keywords: hypergeometric connections, irregular Hodge filtration, p-adic slopes, exponential sums.

When $n > m$, hypergeometric connections are called *confluent*, indicating the merging of singularities, and have a regular singularity at 0 and an irregular singularity at ∞. Sabbah [2018, Theorem 0.7] showed that a rigid irreducible connection on $\mathbb{P}^1$ with real formal exponents at each singular point admits a variation of irregular Hodge structure away from singularities. For confluent hypergeometric connections, Sabbah and Yu [2019] computed the corresponding irregular Hodge numbers. In addition, Castaño Domínguez and Sevenheck [2021, Theorem 4.7] and Castaño Domínguez, Reichelt and Sevenheck [Castaño Domínguez et al. 2019, Theorem 5.8] explicitly calculated the irregular Hodge filtration for $m = 0$ or 1, respectively.

This article focuses on cases where λ, α, and β are rational numbers. We explicitly construct the irregular Hodge filtration $F^{\bullet}_{\mathrm{irr}}$ on hypergeometric connections in Theorem 3.3.1 and provide a uniform method for reproving the results of Fedorov and Sabbah–Yu.

Theorem 1.0.1 (3.3.1). *Suppose (α, β) is nonresonant. We define a map $\theta : \{1, \dots, n\} \to \mathbb{R}$ by*

$$\theta(k) = (n-m)\alpha_k + \#\{i \mid \beta_i < \alpha_k\} + (n-k) - \sum_{i=1}^{n} \alpha_i + \sum_{j=1}^{m} \beta_j. \tag{1.0.1.1}$$

Then, up to an $\mathbb{R}$-shift,[1] *the jumps of the irregular Hodge filtration on $\mathcal{H}yp_\lambda(\alpha, \beta)$ occur at $\theta(k)$ and, for any $p \in \mathbb{R}$, we have*

$$\operatorname{rk} \operatorname{gr}^p_{F_{\mathrm{irr}}} \mathcal{H}yp_\lambda(\alpha; \beta) = \#\theta^{-1}(p).$$

1.1. ***Application to Frobenius slopes of hypergeometric sums.*** Our method has an arithmetic application to the Frobenius slopes of hypergeometric sums: the arithmetic incarnation of hypergeometric functions [Katz 1990].

Let K be a p-adic field with residue field $\mathbb{F}_p$ containing an element π satisfying $\pi^{p-1} = -p$. Such an element π corresponds to an additive character $\psi : \mathbb{F}_p \to K^\times$ by Dwork's theory [1974]. Suppose that (α, β) is nonresonant and that

$$\alpha_i = \frac{a_i}{p-1},\ \beta_j = \frac{b_j}{p-1} \in \frac{\mathbb{Z}}{p-1}.$$

Miyatani [2020] showed that there exists a unique Frobenius structure φ (up to a scalar) on the analytification of the hypergeometric connection $\mathcal{H}yp_{(-1)^{m+np}/\pi^{n-m}}(\alpha; \beta)$ on S_K, which underlies an overconvergent F-isocrystal on the special fiber of S (called the *hypergeometric F-isocrystal*). The Frobenius trace of φ at an $\mathbb{F}_q$-point a of S is given by the *hypergeometric sum* $\mathrm{Hyp}(\alpha; \beta)(a)$, defined by

$$\sum_{\substack{x_i, y_j \in \mathbb{F}_q^\times \\ x_1 \cdots x_n = a y_1 \cdots y_m}} \psi\left(\mathrm{Tr}\left(\sum_{i=1}^{n} x_i - \sum_{j=1}^{m} y_j\right)\right) \cdot \prod_{i=1}^{n} \omega^{a_i}(\mathrm{Nm}(x_i)) \prod_{j=1}^{m} \omega^{-b_j}(\mathrm{Nm}(y_j)),$$

where $\omega : \mathbb{F}_p^\times \to K^\times$ denotes the Teichmüller lift, $\mathrm{Tr} = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}$, and $\mathrm{Nm} = \mathrm{Nm}_{\mathbb{F}_q/\mathbb{F}_p}$.

[1] Our Hodge numbers $\theta(k)$ are normalized according to the geometric interpretation in Proposition 2.4.1, which is different from those of Fedorov and Sabbah–Yu by a shift.

Frobenius eigenvalues of φ at a are Weil numbers and have complex absolute valuations $q^{(n+m-1)/2}$ via an isomorphism $\overline{K} \simeq \mathbb{C}$. When (α, β) is resonant, the above hypergeometric sum can also be written as a sum of n Weil numbers. It is expected that the p-adic valuations of these Frobenius eigenvalues (called *Frobenius slopes*) are related to the (irregular) Hodge filtration. Our geometric construction of hypergeometric connections allows us to show the following result.

Theorem 1.1.1 (4.0.2). *Suppose $n > m$ and that α_i and β_j lie in $\frac{1}{p-1}\mathbb{Z} \cap [0, 1)$. For every p-power q and $a \in \mathbb{G}_m(\mathbb{F}_q)$, the multiset of Frobenius eigenvalues of* $\mathrm{Hyp}(\alpha; \beta)(a)$ (*normalized by* ord_q) *coincides with the multiset of irregular Hodge numbers $\{\theta(1), \dots, \theta(n)\}$ defined in (1.0.1.1).*

Following [Mazur 1972], we encode the information of the p-adic valuations of Frobenius eigenvalues and (irregular) Hodge numbers into the Newton polygon and the (irregular) Hodge polygon, respectively, as defined in Definition 4.0.1.

For crystalline cohomology groups of a smooth proper variety over k, Mazur and Ogus showed that the associated (Frobenius) Newton polygon lies above the Hodge polygon defined by Hodge numbers [Mazur 1972; Berthelot and Ogus 1978]. For F-isocrystals associated with exponential sums, "Newton above Hodge" type results were studied by Dwork's school. For example, Dwork [1974], Sperber [1977], and Wan [1993] proved that Kloosterman sums (hypergeometric sums of type $(n, 0)$ with $\alpha = (0, \dots, 0)$) are everywhere ordinary on $|\mathbb{G}_{m,\mathbb{F}_p}|$; i.e., two polygons coincide for every closed point $a \in |\mathbb{G}_m|$. We use a "Newton above Hodge" result of Adolphson and Sperber [1989; 1993] and identify their (combinatorial) Hodge polygon for the above hypergeometric sums with the irregular Hodge polygon of hypergeometric connections. Finally, we deduce "Newton equals Hodge" by a criterion for ordinariness due to Wan [1993].

Remark 1.1.2. (i) One may also consider the Frobenius Newton polygon of hypergeometric sums defined by multiplicative characters of orders dividing $p^s - 1$ for a positive integer s. In this case, Adolphson and Sperber showed that the associated Frobenius Newton polygon lies above their (combinatorial) Hodge polygon, which can be viewed as an average of irregular Hodge polygons. However, the associated hypergeometric sums may not be ordinary in the case $s > 1$. There is an example of hypergeometric sums (of type $(n, m) = (2, 0)$) for which the Frobenius Newton polygon lies strictly above Adolphson and Sperber's Hodge polygon [1987] for every $a \in |\mathbb{G}_{m,\mathbb{F}_p}|$.

(ii) The ordinariness of hypergeometric sums also fails in the nonconfluent case (i.e., $n = m$). For $p = 31$ and the hypergeometric sum defined by $\alpha = (0, 0, 0, 0)$, $\beta = \left(\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}\right)$ at $a = 4$ or 17, its Newton polygon $\left(\text{with slope } \left(\frac{5}{2}, \frac{5}{2}, \frac{9}{2}, \frac{9}{2}\right)\right)$ [Drinfeld and Kedlaya 2017, Appendix A.5][2] strictly lies above the irregular Hodge polygon (with slope $(2, 3, 4, 5)$).

1.2. *Strategy of proof.* The proof of Theorem 1.0.1 can be reduced to calculating the irregular Hodge filtration on each fiber of $\mathcal{H}yp_\lambda(\alpha, \beta)$. We adopt an approach similar to those used in [Fresán et al. 2022; Sabbah and Yu 2023; Qin 2024], where the authors calculated the Hodge numbers of motives attached to Kloosterman and Airy moments. The key ingredient of this argument is an (exponentially) geometric

[2] In [loc. cit.], the Frobenius slopes are normalized and are different from our convention by a shift of 2.

interpretation of hypergeometric connections in Corollary 2.3.3. More precisely, there exists a smooth quasiprojective variety X with a regular function $g : X \times S \to \mathbb{A}^1$ such that the hypergeometric connections are subquotients of the $\mathscr{D}_S$-module $\mathcal{H}^N \mathrm{pr}_+(\mathcal{O}_{X\times S}, \mathrm{d} + \mathrm{d}g)$, where $N = \dim X$ and pr is the projection $\mathrm{pr} : X \times S \to S$. Our construction is motivated by Katz's hypergeometric sums and the function-sheaf dictionary. A related construction can be found in [Kamgarpour and Yi 2021].

Through this geometric interpretation, each fiber $\mathcal{H}yp_\lambda(\alpha, \beta)_a$ at a closed point a of S is identified with a subquotient of the twisted de Rham cohomology of the pair $(X, g_a := g|_{\mathrm{pr}_z^{-1}(a)})$, i.e., the hypercohomology of the twisted de Rham complex $(\Omega^\bullet_X, \mathrm{d}+\mathrm{d}g_a)$. Then, we reduce to calculate the irregular Hodge filtration on the twisted de Rham cohomology of the pair (X, g_a) (up to a shift).

The irregular Hodge filtration on the twisted de Rham cohomology of the pairs (X, g_a) has been studied in [Yu 2014]. In the context of our case, we can select $X = \mathbb{G}_m^{n+m-1}$ and g_a as a Laurent polynomial with good properties; see Corollary 2.3.3. Under these assumptions, Yu showed that the irregular Hodge filtration on $\mathrm{H}^{n+m-1}_{\mathrm{dR}}(X, g_a)$ can be calculated by the Newton polyhedron filtration on the Newton polytope $\Delta(g_a)$ (3.1.1.1). This identification enables us to prove, via a combinatorial approach, a fiberwise version of Theorem 1.0.1 as follows.

Theorem 1.2.1 (3.3.3). *Up to an $\mathbb{R}$-shift, the jumps of the irregular Hodge filtration $F^\bullet_{\mathrm{irr}}$ on the fiber $\mathcal{H}yp(\alpha; \beta)_a$ occur at $\theta(k)$ from (1.0.1.1) for $1 \le k \le n$. Moreover, we have*

$$\dim \mathrm{gr}^p_{F_{\mathrm{irr}}} \mathcal{H}yp(\alpha; \beta)_a = \#\theta^{-1}(p) \quad \textit{for any } p \in \mathbb{R}.$$

In addition, our geometric construction allows us to answer a question of Katz [1990, 6.3.8] on the comparison between modified hypergeometric $\mathscr{D}$-modules and hypergeometric connections in the resonant case (see Proposition 2.4.7) when the parameters are rational.

1.3. ***Organization of this article.*** We present a geometric interpretation of hypergeometric connections in Section 2. Section 3 is devoted to the proofs of Theorems 1.2.1 and 1.0.1. In Section 4, we study hypergeometric sums defined by multiplicative characters of orders dividing $p-1$ and prove that they are ordinary (Theorem 1.1.1).

2. Hypergeometric connections

In this section, we give an (exponentially) geometric interpretation of the hypergeometric connections in Proposition 2.3.1, Corollary 2.3.3, and Proposition 2.4.1. We work with varieties over $\mathbb{C}$ in Sections 2 and 3.

2.1. *Review of hypergeometric connections following* [Katz 1990].

2.1.1. *Hypergeometric connections.* Let $n \ge m \ge 0$ be two integers, $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_j)$ two sequences of nondecreasing rational numbers (and we don't require that they lie in [0, 1) as in the introduction), and $\lambda \in \mathbb{Q}$. Let $\mathscr{D}_S$ be the sheaf of differential operators on the scheme S, which is $\mathbb{G}_m \backslash \{1\}$

(resp. $\mathbb{G}_m$) if $n = m$ (resp. $n > m$) with coordinate z. Then, the hypergeometric connection $\mathcal{H}yp_\lambda(\alpha;\beta)$ on S is defined by the differential operator in (1.0.0.1) as

$$\mathscr{D}_S/\operatorname{Hyp}_\lambda(\alpha;\beta). \tag{2.1.1.1}$$

By [Katz 1990, (3.1)], one has for $\gamma \in \mathbb{Q}$ that

$$\mathcal{H}yp_\lambda(\alpha;\beta)\otimes\Big(\mathcal{O}, \mathrm{d}+\gamma\frac{\mathrm{d}z}{z}\Big) \simeq \mathcal{H}yp_\lambda(\alpha+\gamma;\beta+\gamma), \tag{2.1.1.2}$$

where $\alpha+\gamma$ (resp. $\beta+\gamma$) is the sequence consisting of $\alpha_i+\gamma$ (resp. $\beta_j+\gamma$). Furthermore, one has for $\mu \in \mathbb{Q}^\times$ that

$$[x \mapsto \mu\cdot x]^+\mathcal{H}yp_\lambda(\alpha;\beta) \simeq \mathcal{H}yp_{\lambda/\mu}(\alpha;\beta). \tag{2.1.1.3}$$

Thanks to the above relations, we can often assume that $\lambda = 1$ and $\alpha_1 = 0$. For simplicity, we denote by $\mathcal{H}yp(\alpha;\beta)$ the connection $\mathcal{H}yp_1(\alpha;\beta)$.

When the pair (α, β) is nonresonant, i.e., $\alpha_i - \beta_j \notin \mathbb{Z}$ for any i and j, Katz [1990, Proposition 3.2] showed that $\mathcal{H}yp(\alpha;\beta)$ is irreducible and only depends on $\alpha \bmod \mathbb{Z}$ and $\beta \bmod \mathbb{Z}$. In this case, we may assume that α and β are two nondecreasing sequences of rational numbers in $[0, 1)$.

2.1.2. *Modified hypergeometric $\mathscr{D}$-modules.* Given a morphism g between smooth varieties, for a bounded complex of holonomic algebraic $\mathscr{D}$-modules, following [Fresán et al. 2022, Appendix A.1], we denote by g^+, g_+, and $g_\dagger$ the derived pullback functor, the pushforward functor, and the pushforward with compact support functor, respectively. The k-th cohomology of a complex K is denoted by $\mathcal{H}^k(K)$.

Let $\mathrm{mult}: \mathbb{G}_m \times \mathbb{G}_m \to \mathbb{G}_m$ be the product map. The convolution functors $\star_*$ and $\star_!$ on $\mathbb{G}_m$ are defined, for two objects M and N of $\mathrm{D}^b(\mathscr{D}_{\mathbb{G}_m})$, by

$$M \star_* N := \mathrm{mult}_+(M \boxtimes N) \quad \text{and} \quad M \star_! N := \mathrm{mult}_\dagger(M \boxtimes N),$$

respectively. These convolution functors are associative and commutative. Moreover, the duality functor $\mathbb{D}$ interchanges $\star_!$ and $\star_*$.

Definition 2.1.3. Let α and β be two sequences of rational numbers. For $? \in \{!, *\}$, the convolution

$$\mathcal{H}yp(\alpha_1;\varnothing)\star_? \cdots \star_? \mathcal{H}yp(\alpha_n;\varnothing)\star_? \mathcal{H}yp(\varnothing;\beta_1)\star_? \cdots \star_? \mathcal{H}yp(\varnothing;\beta_m)$$

is a holonomic $\mathscr{D}_{\mathbb{G}_m}$-module [Katz 1990, (6.3.6)]. We denote it by $\mathcal{H}yp(?;\alpha;\beta)$ and call it a *modified hypergeometric $\mathscr{D}$-module*.

The restrictions of the above two modified hypergeometric $\mathscr{D}$-modules to S are generally not isomorphic to the hypergeometric connections. When (α, β) is nonresonant, the natural map

$$\mathcal{H}yp(!;\alpha;\beta) \to \mathcal{H}yp(*;\alpha;\beta) \tag{2.1.3.1}$$

is an isomorphism, as seen by using an argument similar to those in [Katz 1990, Theorem 8.4.2 (5)] and [Miyatani 2020, Proposition 3.3.3]. In this case, both modified hypergeometric $\mathscr{D}_{\mathbb{G}_m}$-modules, restricted to S, are isomorphic to the hypergeometric connection $\mathcal{H}yp(\alpha;\beta)$ by [Katz 1990, (5.3.1)].

2.2. ***The Newton polytope of a Laurent polynomial.*** We study the Newton polytope of a Laurent polynomial appearing in the geometric interpretation of hypergeometric connections in Proposition 2.4.1.

Definition 2.2.1. Let N be a positive integer and

$$g(z_1, \dots, z_N) = \sum_{\tau \in \mathbb{Z}^N} c(\tau) z^\tau$$

be a Laurent polynomial in variables $z_1, \dots, z_N$, with $z^\tau = \prod_{i=1}^N z_i^{\tau_i}$ for $\tau = (\tau_1, \dots, \tau_N)$.

(1) The support of g is the subset $\operatorname{Supp}(g) = \{\tau \mid c(\tau) \neq 0\}$ of $\mathbb{Z}^N$.

(2) The *Newton polytope* $\Delta(g)$ is the convex hull of the set $\operatorname{Supp}(g) \cup \{0\}$ in $\mathbb{R}^N$.

(3) The Laurent polynomial g is called *nondegenerate* with respect to $\Delta(g)$ (or simply nondegenerate) if, for each face $\sigma \subset \Delta(g)$ not passing through 0, the Laurent polynomial $g_\sigma := \sum_{\tau \in \sigma \cap \mathbb{Z}^N} c(\tau) z^\tau$ has no critical point in $(\mathbb{C}^\times)^N$.

Let $n \geq m \geq 0$ and $d \geq 1$ be three integers, $f : \mathbb{G}_m^{n+m} \to \mathbb{A}^1$ the Laurent polynomial

$$f : (x_2, \dots, x_n, y_1, \dots, y_m, z) \mapsto \sum_{i=2}^n x_i^d - \sum_{j=1}^m y_j^d + z \cdot \frac{\prod_{j=1}^m y_j^d}{\prod_{i=2}^n x_i^d}, \tag{2.2.1.1}$$

and $\mathrm{pr}_z : \mathbb{G}_m^{n+m} \to \mathbb{G}_m$ the projection onto the z-coordinate. For $a \in \mathbb{C}^\times$, we set $f_a = f|_{\mathrm{pr}_z^{-1}(a)}$.

We denote by $\{u_i, v_j\}_{2 \leq i \leq n, 1 \leq j \leq m}$ the coordinates in $\mathbb{R}^{n+m-1}$, and identify a monomial $\prod_i x_i^{a_i} \cdot \prod_j y_j^{b_j}$ with a lattice point $(a_i, b_j) \in \mathbb{Z}^{n+m-1} \subset \mathbb{R}^{n+m-1}$.

Lemma 2.2.2. *Assume that $n > m = 0$ and $a \in \mathbb{C}^\times$.*

(1) *The Laurent polynomial f_a is convenient*; *i.e.*, *the origin is in the interior of $\Delta(f_a)$.*

(2) *The Newton polytope $\Delta(f_a)$ is defined by*

$$h_{n+1} := \sum_{i=2}^n u_i \leq d \qquad \textit{and} \qquad h_{i_0} := \sum_{i=2}^n u_i - (n-m) u_{i_0} \leq d, \quad 2 \leq i_0 \leq n. \tag{2.2.2.1}$$

(3) *The Laurent polynomial f_a is nondegenerate with respect to $\Delta(f_a)$.*

Proof. (1) Let P_i for $2 \leq i \leq n$ and R be the points in $\mathbb{Z}^{n-1}$ corresponding to x_i^d and $1/\prod_{i=2}^n x_i^d$, respectively. Observe that 0 is an interior point of the Newton polytope, as $0 = \big(\sum_{i=2}^n P_i + R\big)/n$.

(2) A face $\sigma \subset \Delta(f_a)$ of dimension $n-2$ must pass through $n-1$ points among $\{P_i, R\}$. So either $R \notin \sigma$ or there exists a $P_{i_0} \notin \sigma$. In the first case, the face lies on the hyperplane defined by the equation $h_{n+1} = d$. In the latter case, the face lies on the hyperplane defined by the equations $h_{i_0} = d$.

(3) Let σ be a face which does not pass through 0. Since the support of f_a has n points, it must pass through at most $n-1$ points in $\operatorname{Supp}(f_a)$. Let $I \subset \{2, \dots, n\}$ be a subset of the indices. Then $f_{a,\sigma}$ is either

$$f_{a,\sigma} = \sum_{i \in I} x_i^d \qquad \text{or} \qquad f_{a,\sigma} = \sum_{i \in I} x_i^d + \frac{a}{\prod_{i=2}^n x_i^d} \quad \text{for } |I| \leq n-2.$$

We can check that they are smooth on $\mathbb{G}_m^{n-1}$. Therefore f_a is nondegenerate. □

Lemma 2.2.3. *Assume that $n > m \neq 0$ and $a \in \mathbb{C}^\times$.*

(1) *The cone $\mathbb{R}_{\geq 0} \cdot \Delta(f_a)$ is defined by*

$$u_i + v_j \geq 0, \quad v_j \geq 0$$

for $i = 2, \dots, n$ and $j = 1, \dots, m$,

(2) *The Newton polytope $\Delta(f_a)$ is defined by*

$$u_i + v_j \geq 0, \quad v_j \geq 0, \quad h_{n+1} := \sum u_i + \sum v_j \leq d$$

and

$$h_{i_0} := \sum_i u_i + \sum_j v_j - (n-m)u_{i_0} \leq d, \quad 2 \leq i_0 \leq n. \tag{2.2.3.1}$$

(3) *The Laurent polynomial f_a is nondegenerate with respect to $\Delta(f_a)$.*

Proof. Let P_i and Q_j be the points in $\mathbb{Z}^{n+m-1}$ corresponding to monomials x_i^d and y_j^d for $2 \leq i \leq n$ and $1 \leq j \leq m$, respectively, and R the lattice point corresponding to

$$\prod_{j=1}^m y_j^d \Big/ \prod_{i=2}^n x_i^d.$$

In this case, the origin 0 is not an interior point of the Newton polytope. So $\Delta(f_a)$ has $(n+m+1)$-many vertices. To determine a face of dimension $n+m-2$, we need to choose $(n+m-1)$-many points among $\{P_i, Q_j, R\}$.

(1) For the first part, it suffices to determine faces $\sigma \subset \Delta(f_a)$ with dimensions $n+m-2$ containing 0.

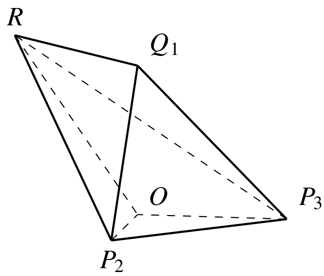
- If σ does not pass through R, it contains $(n+m-2)$ distinct points in $\{P_i, Q_j\}$. In this case, σ misses one point Q_{j_0} and lies on the hyperplane $v_{j_0} = 0$. Otherwise, σ misses one point P_{i_0}. Hence the hyperplane is given by the equation $u_{i_0} = 0$. Therefore, R and P_{i_0} lie on the two sides of the hyperplane, respectively, which is absurd.
- If σ passes through R, it contains $(n+m-3)$ distinct points in $\{P_i, Q_j\}$. In this case, σ has to miss one P_{i_0} and one Q_{j_0} and lies on the hyperplane $u_{i_0} + v_{j_0} = 0$. Otherwise, σ misses two $P_{i_0}, P_{i'_0}$ or $Q_{j_0}, Q_{j'_0}$. So σ lies on the hyperplane $u_{i_0} - u_{i'_0} = 0$ or $v_{j_0} - v_{j'_0} = 0$. However, the points $P_{i_0}, P_{i'_0}$ or $Q_{j_0}, Q_{j'_0}$ lie on different sides of the hyperplane $u_{i_0} - u_{i'_0} = 0$ or $v_{j_0} - v_{j'_0} = 0$, which contradicts the definition of σ.

(2) For the second part, it suffices to determine faces of dimension $n+m-2$ that do not pass through the origin.

- If $R \notin \sigma$, then σ contains all points P_i and Q_j. In this case, σ lies on the hyperplane $\sum u_i + \sum v_j = d$.
- If $R \in \sigma$, then σ contains $n+m-2$ points among $\{P_i, Q_j\}$. In this case, σ misses one P_{i_0} and lies on the hyperplane $h_{i_0} = d$. Otherwise, it misses one Q_{j_0} and lies on the hyperplane

$$\sum_{i=2}^n u_i + \sum_{j=1}^m v_j + (n-m)v_{j_0} = d.$$

However, the points 0 and Q_{j_0} are on different sides of the hyperplane.



(3) Let σ be a face which does not pass through 0. Since the support of f_a has $n+m$ points, it must pass through at most $n+m-1$ points in $\mathrm{Supp}(f_a)$. Let $I \subset \{2, \dots, n\}$ and $J \subset \{1, \dots, m\}$ be two subsets of the indices. Then $f_{a,\sigma}$ is either

$$f_{a,\sigma} = \sum_{i\in I} x_i^d - \sum_{j\in J} y_j^d \qquad \text{or} \qquad f_{a,\sigma} = \sum_{i\in I} x_i^d - \sum_{j\in J} y_j^d + a\cdot \frac{\prod_{j=1}^m y_j^d}{\prod_{i=2}^m x_i^d} \quad \text{for } |I|+|J| \le n+m-2.$$

To see that the partial Laurent polynomials $f_{a,\sigma}$ are all smooth on $(\mathbb{G}_m)^{n+m-1}$, it suffices to show that the system of equations

$$\{f_{a,\sigma} = \partial_{x_i} f_{a,\sigma} = \partial_{y_j} f_{a,\sigma} = 0 \mid 2 \le i \le n,\, 1 \le j \le m\}$$

has no solutions in $(\mathbb{G}_m)^{n+m-1}$. In fact, in the first case above, taking any $i_0 \in I$ or $j_0 \in J$, we have the equation $0 = \partial_{x_{i_0}} f_{a,\sigma} = dx_{i_0}^{d-1}$ or $0 = \partial_{y_{j_0}} f_{a,\sigma} = dj_{j_0}^{d-1}$, which is impossible if $d = 1$. If $d \ge 2$, then x_{i_0} or y_{j_0} is forced to be 0. In the second case, for any $i_0 \notin I$ or $j_0 \notin J$, we have

$$0 = \partial_{x_{i_0}} f_{a,\sigma} = -\frac{d}{x_{i_0}}\cdot a\cdot \frac{\prod_{j=1}^m y_j^d}{\prod_{i=2}^n x_i^d} \quad \text{or} \quad 0 = \partial_{y_j} f_{a,\sigma} = \frac{d}{y_{j_0}}\cdot a\cdot \frac{\prod_{j=1}^m y_j^d}{\prod_{i=2}^n x_i^d},$$

which again forces some $y_j = 0$.

Consequently, all the $f_{a,\sigma}$ have no critical points in $\mathbb{G}_m^{n+m-1}$, and therefore f_a is nondegenerate. $\square$

Lemma 2.2.4. *Assume that $n = m$ and $a \in \mathbb{C}^\times$.*

(1) *The cone $\mathbb{R}_{\ge 0}\cdot \Delta(f_a)$ is defined by*

$$u_i + v_j \ge 0, \quad v_j \ge 0$$

for $i = 2, \dots, n$ and $j = 1, \dots, m$.

(2) *The Newton polytope $\Delta(f_a)$ is defined by*

$$u_i + v_j \ge 0, \quad v_j \ge 0, \quad \textit{and} \quad h_{n+1} := \sum u_i + \sum v_j \le d. \tag{2.2.4.1}$$

(3) *The Laurent polynomial f_a is nondegenerate with respect to $\Delta(f_a)$ if $a \ne 1$.*

Proof. We use the same notation as in Lemma 2.2.3. The proof of the first assertion is the same as that in Lemma 2.2.3. The second assertion follows from the observation that the points $\{P_i, Q_j, R\}$ all lie on the hyperplane

$$\sum u_i + \sum v_j - d = 0.$$

Let σ be the face passing through $\{P_i, Q_j, R\}$. If a face τ of $\Delta(f_a)$ does not contain 0, it is a face of σ. Similar to the proof of Lemma 2.2.3, one can check that, if τ is a proper face of σ, there is no solution for the system of equations

$$\{f_{a,\tau} = \partial_{x_i} f_{a,\tau} = \partial_{y_j} f_{a,\tau} = 0 \mid 2 \le i \le n, 1 \le j \le m\}.$$

If $\tau = \sigma$, the system of equations

$$\{f_a = \partial_{x_i} f_a = \partial_{y_j} f_a = 0 \mid 2 \le i \le n, 1 \le j \le m\}$$

has solutions in $\mathbb{G}_m^{n+m-1}$ if and only if $a = 1$ (in such cases $x_i = y_j = c \in \mathbb{R}$ are solutions). So f_a is nondegenerate with respect to $\Delta(f_a)$ if $a \neq 1$. $\square$

Remark 2.2.5. The volume of $\Delta(f_a)$ is $d^{n+m-1}n/(n+m-1)!$. In fact, the Newton polytope can be decomposed into n-copies $(n+m-1)$-simplexes, and each of them has volume $d^{n+m-1}/(n+m-1)!$.

2.3. ***Geometric interpretations.*** We present some geometric interpretations of hypergeometric connections here. Let d be a common denominator of α_i and β_j, and set $a_i = d \cdot \alpha_i$ and $b_j = d \cdot \beta_j$. To α_i (resp. β_j), we associate the character $\chi_i : \mu_d \to \mathbb{C}^\times$ (resp. ρ_j) which sends ζ_d to $\zeta_d^{a_i}$ (resp. $\zeta_d^{b_j}$). Set

$$\begin{aligned} \chi \times \rho &= \chi_1 \times \cdots \times \chi_n \times \rho_1^{-1} \times \cdots \times \rho_m^{-1}, \\ \tilde{\chi} \times \rho &= \chi_2 \times \cdots \times \chi_n \times \rho_1^{-1} \times \cdots \times \rho_m^{-1} \end{aligned} \tag{2.3.0.1}$$

as products of these characters.

Now we introduce two diagrams as follows:

- Let $\mathbb{G}_m^{n+m}$ be the torus with coordinates x_i and y_j for $1 \le i \le n$ and $1 \le j \le m$. We consider the diagram

$$\begin{tikzcd} & \mathbb{G}_m^{n+m} \arrow[dl, "\sigma"'] \arrow[dr, "\varpi"] & \\ \mathbb{A}_t^1 & & \mathbb{G}_m \end{tikzcd} \tag{2.3.0.2}$$

where

$$\sigma(x_i, y_j) = \sum_{i=1}^n x_i^d - \sum_{j=1}^m y_j^d \quad \text{and} \quad \varpi(x_i, y_j) = \prod_{i=1}^n x_i^d \Big/ \prod_{j=1}^m y_j^d.$$

Let the group μ_d^{n+m} act on $\mathbb{G}_m^{n+m}$ by multiplication and on $\mathbb{A}_t^1$ and $\mathbb{G}_m$ trivially. Then, it can be verified that σ and ϖ are μ_d^{n+m}-equivariant.

• Let $\mathbb{G}_m^{n+m}$ be the torus with coordinates z, x_i, and y_j for $2 \le i \le n$ and $1 \le j \le m$, and S be $\mathbb{G}_{m,z}$ (resp. $\mathbb{G}_{m,z} \backslash \{1\}$) if $n \neq m$ (resp. $n = m$). We consider the diagram

$$\begin{tikzcd} & \mathbb{G}_m^{n+m} \arrow[dl, "f"'] \arrow[dr, "\mathrm{pr}_z"] & U := S \times \mathbb{G}_m^{n+m-1} \arrow[l, hook'] \arrow[dr, "\mathrm{pr}_z"] & \\ \mathbb{A}^1_t & & \mathbb{G}_{m,z} & S \arrow[l, hook'] \end{tikzcd} \tag{2.3.0.3}$$

where pr_z is the projection to the z-coordinate and f is the Laurent polynomial

$$\sum_{i=2}^{n} x_i^d - \sum_{j=1}^{m} y_j^d + z \cdot \frac{\prod_{j=1}^m y_j^d}{\prod_{i=2}^n x_i^d}$$

defined in (2.2.1.1). Let the group $G = \mu_d^{n+m-1}$ act on $\mathbb{G}_m^{n+m}$ (resp. $S \times \mathbb{G}_m^{n+m-1}$) by multiplication on the coordinates x_i and y_j and trivially on z, and on $\mathbb{A}^1_t$, $\mathbb{G}_{m,z}$, and S trivially. Then f and pr_z are μ_d^{n+m-1}-equivariant.

Let $\mathcal{E}^t = (\mathcal{O}, \mathrm{d} + \mathrm{d}t)$ be the exponential $\mathscr{D}$-module on $\mathbb{A}^1_t$. For a regular function $h : X \to \mathbb{A}^1_t$, we denote by $\mathcal{E}^h := h^+\mathcal{E}^t$ the connection $(\mathcal{O}_X, \mathrm{d} + \mathrm{d}h)$ on X.

Proposition 2.3.1. *Let α and β be as above. The complexes $\varpi_?\mathcal{E}^\sigma$ are μ_d^{n+m}-equivariant and concentrated in degree 0 for $? \in \{\dagger, +\}$. Moreover, we have isomorphisms of $\mathscr{D}_{\mathbb{G}_m}$-modules*

$$\mathcal{H}yp(*; \alpha; \beta) \simeq (\varpi_+\mathcal{E}^\sigma)^{(\mu_d^{n+m}, \chi\times\rho)},$$
$$\mathcal{H}yp(!; \alpha; \beta) \simeq (\varpi_\dagger\mathcal{E}^\sigma)^{(\mu_d^{n+m}, \chi\times\rho)},$$

where the exponent $(\mu_d^{n+m}, \chi \times \rho)$ means taking the $(\chi \times \rho)$-isotypic component with respect to the action of μ_d^{n+m}.

Proof. The case of $\mathcal{H}yp(!; \alpha; \beta)$ can be deduced from the case of $\mathcal{H}yp(*; \alpha; \beta)$ by applying the duality functor. So, we only prove the latter case. Recall that the action of μ_d^{n+m} on $\mathbb{A}^1_t$ is trivial in diagram (2.3.0.2). So the $\mathscr{D}_{\mathbb{A}^1_t}$-module $\mathcal{E}^t$ is μ_d^{n+m}-equivariant. Since σ and ϖ are both μ_d^{n+m}-equivariant morphisms, φ_+ and σ^+ preserve μ_d^{n+m}-equivariant objects. Hence the complex

$$\varphi_+\mathcal{E}^\sigma = \varphi_+\sigma^+\mathcal{E}^t$$

is μ_d^{n+m}-equivariant.

Assume that $(n, m) = (1, 0)$. Then $\sigma : \mathbb{G}_{m,x_1} \to \mathbb{A}^1$ is the map $x_1 \mapsto x_1^d$ and $\varpi : \mathbb{G}_{m,x_1} \to \mathbb{G}_{m,z}$ is the d-th power map. So by the identity

$$\varpi_+\mathcal{O}_{\mathbb{G}_m} = \bigoplus_{i=0}^{d-1}\Big(\mathcal{O}_{\mathbb{G}_m}, \mathrm{d} + \frac{i}{d}\frac{\mathrm{d}z}{z}\Big)$$

and the projection formula, we have

$$(\varpi_+\mathcal{E}^\sigma) = \mathcal{E}^z \otimes (\varpi_+\mathcal{O}_{\mathbb{G}_m}) = \bigoplus_{i=0}^{d-1} \mathcal{E}^z \otimes \Big(\mathcal{O}_{\mathbb{G}_m}, \mathrm{d} + \frac{i}{d}\frac{\mathrm{d}z}{z}\Big),$$

which is concentrated in degree 0. Taking the isotypic component, we have

$$(\varpi_+\mathcal{E}^\sigma)^{(\mu_d^{n+m},\chi\times\rho)} = (\varpi_+\mathcal{E}^{x_1^d})^{(\mu_d,\chi_1)} = \mathcal{E}^z\otimes(\varpi_+\mathcal{O}_{\mathbb{G}_m})^{(\mu_d,\chi_1)}$$
$$= \Big(\mathcal{O}_{\mathbb{G}_m}, \mathrm{d}+\mathrm{d}z+\alpha_1\frac{\mathrm{d}z}{z}\Big) = \mathcal{H}yp(*;\alpha_1;\varnothing)$$

in the case where $(n,m)=(1,0)$. The proof of the case where $(n,m)=(0,1)$ is similar. In general, we use induction on $n+m$. The proof follows from the following lemma. □

Lemma 2.3.2. *Let α, α', β and β' be four sequences of rational numbers with common denominator d, whose lengths are n, n', m and m', respectively. We denote by χ_i, χ'_i, ρ_j, and ρ'_j characters of μ_d corresponding to α_i, α'_i, β_j, and β'_j, respectively. Let σ and ϖ (resp. σ' and ϖ') be the maps for (n,m) (resp. (n',m')) in the diagram (2.3.0.2).*

Suppose that $(\varpi_+\mathcal{E}^\sigma)$ and $(\varpi'_+\mathcal{E}^{\sigma'})$ are concentrated in degree 0, and there are isomorphisms of $\mathscr{D}$-modules

$$\mathcal{H}yp(*;\alpha;\beta)\simeq(\varpi_+\mathcal{E}^\sigma)^{(\mu_d^{n+m},\chi\times\rho)},$$
$$\mathcal{H}yp(*;\alpha';\beta')\simeq(\varpi_+\mathcal{E}^{\sigma'})^{(\mu_d^{n+m},\chi'\times\rho')}.$$

Then $((\varpi\cdot\varpi')_+\mathcal{E}^{\sigma\boxplus\sigma'})$ is also concentrated in degree 0, and we have an isomorphism of $\mathscr{D}$-modules

$$\mathcal{H}yp(*;\alpha,\alpha';\beta,\beta')\simeq((\varpi\cdot\varpi')_+\mathcal{E}^{\sigma\boxplus\sigma})^{(\mu_d^{n+n'+m+m'},\chi\times\chi'\times\rho\times\rho')},$$

where $\varpi\cdot\varpi' = \mathrm{mult}\circ(\varpi\times\varpi')$, pr and pr' are the projections from $\mathbb{G}_m^{n+n'+m+m'}$ to $\mathbb{G}_m^{n+m}$ and $\mathbb{G}_m^{n'+m'}$, respectively, and $\sigma\boxplus\sigma' = \sigma\circ\mathrm{pr}+\sigma'\circ\mathrm{pr}'$ is the Thom–Sebastiani sum.

Proof. The proof of this lemma is essentially that of [Katz 1990, Lemma 5.4.3]. Notice that the exterior product $\mathcal{E}^\sigma\boxtimes\mathcal{E}^{\sigma'}$ is $\mathcal{E}^{\sigma\boxplus\sigma'}$. Then

$$(\varpi_+\mathcal{E}^\sigma)\star_*(\varpi'_+\mathcal{E}^{\sigma'}) = \mathrm{mult}_+((\varpi_+\mathcal{E}^\sigma)\boxtimes(\varpi'_+\mathcal{E}^{\sigma'})) = \mathrm{mult}_+(\varpi\times\varpi')_+(\mathcal{E}^\sigma\boxtimes\mathcal{E}^{\sigma'}) = (\varpi\cdot\varpi')_+\mathcal{E}^{\sigma\boxplus\sigma'}.$$

By the Künneth formula [Hotta et al. 2008, Proposition 1.5.28 (i) and Proposition 1.5.30], we conclude that $(\varpi\cdot\varpi')_+\mathcal{E}^{\sigma\boxplus\sigma'}$ is again concentrated in degree 0.

Viewing μ_d^{n+m}-equivariant and $\mu_d^{n'+m'}$-equivariant objects as $\mu_d^{n+m+n'+m'}$-equivariant via the identifications

$$\mu_d^{n+m}\simeq\mu_d^{n+m}\times 1\quad\text{and}\quad\mu_d^{n'+m'}\simeq 1\times\mu_d^{n'+m'},$$

we can verify that both $\boxtimes$ and mult are $\mu_d^{n+m+n'+m'}$-equivariant. Hence the convolution product $\star_*$ is also $\mu_d^{n+m+n'+m'}$-equivariant. Therefore, we conclude the lemma by taking the corresponding isotypic components of the above formula. □

Corollary 2.3.3. *Let α and β be as above and $\alpha_1=0$. The complexes of $\mathscr{D}_{\mathbb{G}_m}$-modules $\mathrm{pr}_{z?}\mathcal{E}^f$ are μ_d^{n+m-1}-equivariant and concentrated in degree 0 for $?\in\{\dagger,+\}$. Moreover, we have*

$$\mathcal{H}yp(*;\alpha;\beta)\simeq(\mathcal{H}^0\mathrm{pr}_{z+}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)},$$
$$\mathcal{H}yp(!;\alpha;\beta)\simeq(\mathcal{H}^0\mathrm{pr}_{z\dagger}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}.$$

Proof. Similar to Proposition 2.3.1, we only consider the case of $\mathcal{H}yp(*;\alpha;\beta)$. By the construction of the diagram in (2.3.0.3), the morphisms pr_z and f are μ_d^{n+m-1}-equivariant. Hence the complex of $\mathscr{D}_{\mathbb{G}_m}$-modules $\mathrm{pr}_{z+}\mathcal{E}^f=\mathrm{pr}_{z+}f^+\mathcal{E}^t$ is μ_d^{n+m-1}-equivariant.

By assumption, we set $\alpha_1=0$, which implies that the character χ_1 is trivial. By Proposition 2.3.1, we have

$$\begin{aligned}(\varpi_+\mathcal{E}^\sigma)^{(\mu_d^{n+m},\chi\times\rho)}&=\biggl(\biggl(x_1\cdot\prod_{i=2}^n x_i^d\Big/\prod_{j=1}^m y_j^d\biggr)_+\mathcal{E}^{x_1+\sum_{i=2}^m x_i^d-\sum_j y_j^d}\biggr)^{(1\times G,1\times\tilde{\chi}\times\rho)}\\&=(\mathrm{pr}_{z+}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)},\end{aligned}\tag{2.3.3.1}$$

where we performed a change of variable $z=x_1\cdot\prod_{i=2}^n x_i^d/\prod_{j=1}^m y_j^d$ to get rid of the variable x_1 in the last isomorphism. Because $(\varpi_+\mathcal{E}^\sigma)$ is concentrated in degree 0 and isomorphic to $\mathcal{H}yp(*;\alpha;\beta)$, so is $(\mathrm{pr}_{z+}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}$. □

Corollary 2.3.4. *Assume that (α,β) is nonresonant and $\alpha_1=0$. Then, the natural map*

$$(\mathcal{H}^0\mathrm{pr}_{z\dagger}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}\to(\mathcal{H}^0\mathrm{pr}_{z+}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}$$

is an isomorphism of $\mathscr{D}_{\mathbb{G}_m}$-modules. In particular, for a closed point[3] *a of S, the forget-support map*

$$\mathrm{H}^{n+m-1}_{\mathrm{dR},c}(\mathbb{G}_m^{n+m-1},f_a)^{(G,\tilde{\chi}\times\rho)}\to\mathrm{H}^{n+m-1}_{\mathrm{dR}}(\mathbb{G}_m^{n+m-1},f_a)^{(G,\tilde{\chi}\times\rho)}$$

is an isomorphism.

Proof. Using induction on the size of α and β, one can verify that the diagram

$$\begin{tikzcd}
\mathcal{H}yp(!;\alpha;\beta) \arrow[r, "\simeq"] \arrow[d, "\simeq"] & \mathcal{H}yp(*;\alpha;\beta) \arrow[d, "\simeq"] \\
(\mathcal{H}^0\varpi_\dagger\mathcal{E}^\sigma)^{(\mu_d^{n+m},\chi\times\rho)} \arrow[r] \arrow[d, "\simeq"] & (\mathcal{H}^0\varpi_+\mathcal{E}^\sigma)^{(\mu_d^{n+m},\chi\times\rho)} \arrow[d, "\simeq"] \\
(\mathcal{H}^0\mathrm{pr}_{z\dagger}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)} \arrow[r] & (\mathcal{H}^0\mathrm{pr}_{z+}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}
\end{tikzcd}$$

is commutative, where the horizontal morphisms are the canonical forget-support morphisms with the top one being (2.1.3.1), the two upper vertical morphisms are those from Proposition 2.3.1, and the two lower vertical morphisms are (2.3.3.1). So, we deduce the isomorphism

$$(\mathcal{H}^0\mathrm{pr}_{z\dagger}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}\to(\mathcal{H}^0\mathrm{pr}_{z+}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}.$$

At last, we take the noncharacteristic inverse image along $a:\mathrm{Spec}(\mathbb{C})\to\mathbb{G}_m$ and the base change theorem [Hotta et al. 2008, Theorem 1.7.3 & Proposition 1.5.28] to conclude the isomorphism of twisted de Rham cohomology groups. □

[3]The modified hypergeometric $\mathscr{D}_{\mathbb{G}_m}$-modules $\mathcal{H}yp(?,\alpha;\beta)$ are smooth on S, on which the hypergeometric connections $\mathcal{H}yp(\alpha;\beta)$ are defined.

Remark 2.3.5. When (α, β) is nonresonant and $\alpha_1 = 0$, we deduce from Corollary 2.3.3 the isomorphism

$$[z \mapsto (-1)^{n-m}z]^+\mathcal{H}yp(\alpha;\beta) \simeq (\mathcal{H}^0\mathrm{pr}_{z+}\mathcal{E}^{-f})^{(G,\tilde{\chi}\times\rho)}$$

by performing a change of variable by sending x_i and y_j to $-x_i$ and $-y_j$, respectively, in the diagram (2.3.0.3). According to (2.1.1.3), the first term in the above is $\mathcal{H}yp_{(-1)^{n-m}}(\alpha;\beta)$. In particular, the results in Corollary 2.3.4 remain valid if we replace f with $-f$.

2.4. ***Explicit cyclic vectors for hypergeometric connections.*** We present explicit cyclic vectors for $\mathcal{H}yp(\alpha;\beta)$ in terms of sections of some subquotients of some relative de Rham cohomologies equipped with their Gauss–Manin connections. This point of view will be used in the computation of Hodge filtrations in Section 3.

Recall that d is an integer such that $a_i = d\alpha_i$ and $b_j = d\beta_j$ are integers for all i and j, and we take notation from (2.3.0.3). When (α, β) is nonresonant and $\alpha_1 = 0$, there exists an isomorphism between the hypergeometric connection $\mathcal{H}yp(\alpha;\beta)$ and $(\mathcal{H}^0\mathrm{pr}_{z\dagger}\mathcal{E}^f)^{(G,\tilde{\chi}\times\rho)}|_S$ by (2.1.3.1) and Corollary 2.3.3. From now on, we will identify the latter with the relative de Rham cohomology $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$ on S equipped with the Gauss–Manin connection, where $U = S \times \mathbb{G}_m^{n+m-1}$.

Proposition 2.4.1. *Suppose that $\alpha_1 = 0$ and (α, β) is nonresonant. The relative de Rham cohomology $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$ admits a cyclic vector, defined by the cohomology class of the differential form*

$$\omega = \prod_{i=2}^{n} x_i^{a_i} \cdot \prod_{j=1}^{m} y_j^{-b_j} \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m}. \tag{2.4.1.1}$$

Remark 2.4.2. Under the above assumption, the isomorphism class of $\mathcal{H}yp(\alpha;\beta)$ depends only on the congruence classes of α and β modulo $\mathbb{Z}$. Then, any differential form

$$\prod_{i=2}^{n} x_i^{u_i} \cdot \prod_{j=1}^{m} y_j^{-v_j} \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m},$$

satisfying $u_i \equiv a_i$, $v_j \equiv b_j$ modulo d, is a cyclic vector of $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$.

Proof. The morphism $\mathrm{pr}_z : U = S \times \mathbb{G}_m^{n+m-1} \to S$ in (2.3.0.3) is smooth. It follows that the relative de Rham cohomologies $\mathcal{H}^i_{\mathrm{dR}}(U/S, f)$ are equipped with the Gauss–Manin connections $D := \nabla_{z\partial_z}$, given by

$$\nabla_{z\partial_z}\omega = z\partial_z\omega + z\partial_z(f)\omega \tag{2.4.2.1}$$

for $0 \le i \le n+m-1$. By Lemmas 2.2.2, 2.2.3, and 2.2.4, the Laurent polynomial $f_a := f|_{\mathrm{pr}_z^{-1}(a)}$ is nondegenerate for each closed point a of S. By [Adolphson and Sperber 1997, Theorems 1.4 and 4.1], the cohomology group $\mathcal{H}^i_{\mathrm{dR}}(U/S, f_a)$ vanishes if $i \neq n+m-1$.

Now we consider the $(G, \tilde{\chi}\times\rho)$-isotypic component of the connection $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)$. It remains to prove that the cohomology class defined by the differential form ω (2.4.1.1) is a cyclic vector for $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$.

Lemma 2.4.3. *Let $t_2, \dots, t_n$ and $s_1, \dots, s_m$ be integers, and set*

$$\tilde{\omega} := \prod_{i=2}^{n} x_i^{t_i} \cdot \prod_{j=1}^{m} y_j^{s_j} \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m}$$

as a class in $\mathcal{H}_{\mathrm{dR}}^{n+m-1}(U/S,\, f)$. For each u and v such that $2 \le u \le n$ and $1 \le v \le m$, respectively, we have

$$\left(D - \frac{t_u}{d}\right)\tilde{\omega} = x_u^d \cdot \tilde{\omega} \quad \textit{and} \quad \left(D + \frac{s_v}{d}\right)\tilde{\omega} = y_v^d \cdot \tilde{\omega}.$$

Proof. We prove the identity $(D - t_2/d)\tilde{\omega} = x_2^d \cdot \tilde{\omega}$, and the proof for the rest is identical. By (2.4.2.1), we have

$$D\tilde{\omega} = z \cdot \frac{\prod_{j=1}^{m} y_j^d}{\prod_{i=2}^{n} x_i^d} \tilde{\omega}.$$

Since U and S are affine, the image of any $(n+m-2)$-form under the relative differential

$$\nabla_{U/S} : \Omega_{U/S}^{n+m-2} \to \Omega_{U/S}^{n+m-1}$$

in $\mathcal{H}_{\mathrm{dR}}^{n+m-1}(U/S,\, f)$ is zero. Then, we have

$$\begin{aligned} 0 = \nabla_{U/S}\left(\prod_{i=2}^{n} x_i^{t_i} \cdot \prod_{j=1}^{m} y_j^{s_j} \frac{\mathrm{d}x_3}{x_3} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m}\right) &= t_2 \cdot \tilde{\omega} + x_2 \cdot \partial_{x_2} f \cdot \tilde{\omega} \\ = t_2 \cdot \tilde{\omega} + x_2 \cdot \left(d x_2^{d-1} - d x_2^{-1} \cdot z \cdot \frac{\prod_{j=1}^{m} y_j^d}{\prod_{i=2}^{n} x_i^d}\right)\tilde{\omega} &= d\left(x_2^d - \left(D - \frac{t_2}{d}\right)\right)\tilde{\omega}. \end{aligned}$$

This is exactly what we want to prove. □

We show that ω (2.4.1.1) satisfies the hypergeometric differential equation $\mathrm{Hyp}(\alpha;\, \beta)$. By Lemma 2.4.3, we have

$$\prod_{i=2}^{n}(D - \alpha_i)\omega = \prod_{i=2}^{n} x_i^d \cdot \omega \quad \text{and} \quad \prod_{j=1}^{m}(D - \beta_j)\omega = \prod_{j=1}^{m} y_j^d \cdot \omega.$$

Then, we deduce from (2.4.2.1) that

$$\prod_{i=1}^{n}(D - \alpha_i)\omega = D\left(\prod_{i=2}^{n} x_i^d \cdot \omega\right) = z \prod_{j=1}^{m} y_j^d \cdot \omega = z \prod_{j=1}^{m}(D - \beta_j)\omega.$$

Lemma 2.4.4. *The cohomology class of ω in $\mathcal{H}_{\mathrm{dR}}^{n+m-1}(U/S,\, f)^{(G, \tilde{\chi} \times \rho)}$ is nonzero.*

Proof. The lemma is obviously true if $n = 1$ and $m = 0$. In general, assume that $\omega = 0$. Each point $(A,\, B) \in (a_i,\, -b_j) + d \cdot \mathbb{Z}^{n+m-1}$ corresponds to a differential form

$$\prod_{i=2}^{n} x_i^{A_i} \cdot \prod_{j=1}^{m} y_j^{B_j} \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m}. \tag{2.4.4.1}$$

Notice that we can also take the isotypic components on the level of complexes of differential forms, and the relative differential $\nabla_{U/S}$ respects the corresponding isotopic components. Thus, the differential form in (2.4.4.1) defines a cohomology class in $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$. On the other hand, any cohomology class has a representative that is a linear combination of such differential forms.

Since $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$ is nonzero, we can select a differential form $\omega^{(0)}$ that defines a nonzero cohomology class. Given that $\omega^{(0)}$ is a linear combination of differential forms of the form in (2.4.4.1), at least one of such forms defines a nonzero cohomology class. We may assume, without loss of generality, that $\omega^{(0)}$ itself is of the form in (2.4.4.1). Using Lemma 2.4.3, we obtain a sequence of differential forms $\{\omega^{(i)}\}_{i=0}^N$ corresponding to points $(A^{(i)}, B^{(i)})$ such that $\omega^{(i+1)} = (D-\gamma_i)\omega^{(i)}$ for some rational number γ_i, and

$$(A^{(N)}, B^{(N)}) \in (a_i, -b_j) + d\cdot\mathbb{N}^{n+m-1}.$$

Applying Lemma 2.4.3 again, we observe that $\omega^{(N)}$ can be expressed as a linear combination of $\{D^k\omega\}_{k\in\mathbb{N}}$, and is thus equal to 0. Hence there exists $M < N$ such that $\omega^{(M)}$ has a nonzero cohomology class and $(D-\gamma_M)\omega^{(M)} = 0$. Thus, $\mathcal{O}_S\cdot\omega^{(M)}$ is the hypergeometric connection $\mathcal{H}yp(\gamma_M; \varnothing)$. Since it is a subconnection of the irreducible connection $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$, it must be isomorphic to $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$, leading to a contradiction. $\square$

In summary, we obtain a nonzero morphism

$$\mathscr{D}_S/\operatorname{Hyp}(\alpha;\beta) \to \bigoplus_{i=0}^{n-1}\mathcal{O}_S\cdot D^i\omega \subset \mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)} \tag{2.4.4.2}$$

defined by sending 1 to ω. Since the left-hand side of the morphism is irreducible, it must be a subconnection of the irreducible connection $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$ on the right-hand side. Since both sides have the same rank, the above morphism is an isomorphism, implying that ω is a cyclic vector of $\operatorname{Hyp}(\alpha;\beta) \simeq \mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G,\tilde{\chi}\times\rho)}$. $\square$

Remark 2.4.5. If we replace $\mathcal{E}^f$ by

$$(\mathbb{G}_m^{n+m-1}, \mathrm{d}-\mathrm{d}f) = (\mathbb{G}_m^{n+m-1}, \mathrm{d}+\mathrm{d}f)^\vee,$$

the direct sum $\bigoplus_{i=0}^{n-1}\mathcal{O}_S\cdot D^i\omega$ is the $(G, \tilde{\chi}\times\rho)$-isotypic component of $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, -f)$, isomorphic to the connection $\mathcal{H}yp_{(-1)^{n-m}}(\alpha;\beta)$. To see this, it suffices to notice that the corresponding identities in Lemma 2.4.3 become

$$\left(D-\frac{t_u}{d}\right)\omega_{t,s} = -x_u^d\omega_{t,s} \quad\text{and}\quad \left(D+\frac{s_v}{d}\right)\omega_{t,s} = -y_v^d\omega_{t,s}$$

in this case. The rest of the proof relies on the above calculation and Remark 2.3.5.

2.4.6. *Resonant case.* When (α, β) is resonant, the modified hypergeometric $\mathscr{D}$-module $\mathcal{H}yp(*;\alpha;\beta)$ depends only on the classes of α and β modulo $\mathbb{Z}$. Katz [1990, 6.3.8] asked whether $\mathcal{H}yp(*;\alpha;\beta)|_S$ is isomorphic to the connection $\mathcal{H}yp((\alpha_i+r_i);(\beta_j+s_j))$ (2.1.1.1) for suitable integers $r_i, s_j\in\mathbb{Z}$. We provide a positive answer to this question in the following proposition.

Proposition 2.4.7. *When (α, β) is resonant, there exists a positive integer h depending on $\alpha \bmod \mathbb{Z}$ and $\beta \bmod \mathbb{Z}$, such that, for any integers $r, s > h$, the modified hypergeometric $\mathscr{D}$-module $\mathcal{H}yp(*; \alpha; \beta)|_S$ is isomorphic to the hypergeometric connection $\mathcal{H}yp((\alpha_1, \alpha_2 - r, \dots, \alpha_n - r); \beta + s)$.*

Proof. We may assume that $\alpha_1 = 0$. Let $\tilde{\omega}_1, \dots, \tilde{\omega}_n$ be a representative of a basis of the connection $\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G, \tilde{\chi}\times\rho)}$. More precisely, we can write

$$\tilde{\omega}_k = \sum_{e\in\mathbb{Z}^{n-1}, f\in\mathbb{Z}^m} \epsilon_{k,e,f} \prod_{i=2}^{n} x_i^{a_i + d\cdot e_i} \prod_{j=1}^{m} y_j^{-b_j + d\cdot f_j} \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m},$$

where only finitely many $\epsilon_{k,e,f}$ are nonzero. We equip $\mathbb{Z}^{n+m-1}$ with the partial order defined by the relation $a \geq b$ if $a - b \in \mathbb{N}^{n+m-1}$. Let (e_0, f_0) be a maximal element in the set

$$\{(e', f') \mid (e', f') \leq (e, f) \text{ if } \epsilon_{k,e,f} \neq 0\}.$$

Then we take h to be the maximal value among $\{|(e_0)|_i, |(f_0)|_j\}$.

For any $r, s > h$, as in Proposition 2.4.1, we define a morphism of $\mathscr{D}$-modules:

$$\mathscr{D}_S / \operatorname{Hyp}(0, \alpha_2 - r, \dots, \alpha_n - r; \beta + s) \to \bigoplus_{i=0}^{n-1} \mathcal{O}_{\mathbb{G}_m} \cdot D^i \omega \subset \mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S, f)^{(G, \tilde{\chi}\times\rho)} \tag{2.4.7.1}$$

by sending 1 to

$$\omega = \prod_{i=2}^{n} x_i^{a_i - d\cdot r} \cdot \prod_{j=1}^{m} y_j^{-b_j - d\cdot s} \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m}.$$

Since, for all (e, f) with $\epsilon_{k,e,f} \neq 0$, we have $a_i + d\cdot e_i \geq a_i - d\cdot r$ and $b_j + d\cdot f_j \geq b_j - d\cdot s$ for any i and j, we deduce that the class defined by

$$\prod_{i=2}^{n} x_i^{a_i + d\cdot e_i} \prod_{j=1}^{m} y_j^{-b_j + d\cdot f_j} \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m}$$

lies in the image of (2.4.7.1) by Lemma 2.4.3. This morphism is a surjection between two connections of rank n and is, hence, an isomorphism. □

3. Irregular Hodge filtration of hypergeometric connections

This section aims to calculate the (irregular) Hodge filtrations of hypergeometric connections (see Theorems 3.3.1 and 3.3.3). Throughout this section, let $n \geq m \geq 0$ be two integers, and let $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_j)$ be two sequences of nondecreasing rational numbers in $[0, 1)$.

3.1. ***Exponential mixed Hodge structures.*** To explain certain duality on the irregular Hodge filtration of hypergeometric connections, we use the language of exponential mixed Hodge structures introduced by Kontsevich and Soibelman [2011]. We recall the basic definitions of exponential mixed Hodge structures from [Fresán et al. 2022, Appendix].

Let X be a smooth algebraic variety and K a number field. We denote by $\mathrm{MHM}(X, K)$ the abelian category of *mixed Hodge modules* on X with coefficients in K. In particular, when $X = \mathrm{Spec}(\mathbb{C})$, the category $\mathrm{MHM}(X, K)$ is equivalent to the category of mixed K-Hodge structures. Moreover, the bounded derived categories $\mathrm{D}^b(\mathrm{MHM}(X, K))$ admit the six functor formalism. For more details about mixed Hodge modules, see [Saito 1990].

Let $\pi : \mathbb{A}^1 \to \mathrm{Spec}(\mathbb{C})$ be the structure morphism. The category $\mathrm{EMHS}(K)$ of *exponential mixed Hodge structures* with coefficients in K is defined as the full subcategory of $\mathrm{MHM}(\mathbb{A}^1, K)$, whose objects N^{H} have vanishing cohomology on $\mathbb{A}^1$, i.e., those satisfying $\pi_* N^{\mathrm{H}} = 0$.

There is an exact functor $\Pi : \mathrm{MHM}(\mathbb{A}^1, K) \to \mathrm{MHM}(\mathbb{A}^1, K)$ defined by

$$N^{\mathrm{H}} \mapsto s_*(N^{\mathrm{H}} \boxtimes j_! \mathcal{O}^{\mathrm{H}}_{\mathbb{G}_m}), \tag{3.1.0.1}$$

where $j : \mathbb{G}_{m,\mathbb{C}} \to \mathbb{A}^1$ is the inclusion and $s : \mathbb{A}^1 \times \mathbb{A}^1 \to \mathbb{A}^1$ is the summation map. The functor Π is a projector onto $\mathrm{EMHS}(K)$; i.e., it factors through $\mathrm{EMHS}(K)$ with essential image $\mathrm{EMHS}(K)$. Using this functor, the dual of an object M in $\mathrm{EMHS}(K)$ is defined by $\Pi([t \mapsto -t]^* \mathbb{D}(M))$, where t is the coordinate of $\mathbb{A}^1$.

For each object $\Pi(N^{\mathrm{H}})$ of the category $\mathrm{EMHS}(K)$, there exists a weight filtration $W^{\mathrm{EMHS}}_{\bullet}$ on $\Pi(N^{\mathrm{H}})$, defined by the weight filtration on N^{H}: $W^{\mathrm{EMHS}}_n \Pi(N^{\mathrm{H}}) := \Pi(W_n N^{\mathrm{H}})$. We will drop the superscript for simplicity.

The *de Rham fiber* functor from $\mathrm{EMHS}(K)$ to $\mathrm{Vect}_{\mathbb{C}}$ is defined by

$$\Pi(N^{\mathrm{H}}) \mapsto \mathrm{H}^1_{\mathrm{dR}}(\mathbb{A}^1, \Pi(N) \otimes \mathcal{E}^t), \tag{3.1.0.2}$$

where $\Pi(N)$ denotes the underlying $\mathscr{D}$-module of $\Pi(N^{\mathrm{H}})$ and $\mathcal{E}^t$ denotes the exponential $\mathscr{D}$-module $(\mathcal{O}_{\mathbb{A}^1}, \mathrm{d} + \mathrm{d}t)$.

The de Rham fiber functor is faithful, and one can associate an *irregular Hodge filtration* $F^{\bullet}_{\mathrm{irr}}$ on the de Rham fibers of objects in $\mathrm{EMHS}(K)$ by [Fresán et al. 2022, Proposition A.10], constructed using a generalization of Deligne's filtration [Sabbah 2010, §6.b]; see also [Esnault et al. 2017, §1.6].

3.1.1. *Objects of* EMHS *attached to regular functions.* Let X be a smooth affine variety of dimension n and K a number field. We denote by K^{H}_X the trivial Hodge module on X with coefficients in K. For a regular function $g : X \to \mathbb{A}^1$ and an integer r, we consider the exponential mixed Hodge structures

$$\mathrm{H}^r(X, g) := \Pi(\mathcal{H}^{r-n} g_* K^{\mathrm{H}}_X), \quad \mathrm{H}^r_c(X, g) := \Pi(\mathcal{H}^{r-n} g_! K^{\mathrm{H}}_X).$$

The exponential mixed Hodge structures $\mathrm{H}^r(X, g)$ and $\mathrm{H}^r_c(X, g)$ are mixed of weights at least r and mixed of weights at most r, respectively, by [Fresán et al. 2022, A.19].

The de Rham fiber of $\mathrm{H}^r_?(X, g)$ is isomorphic to $\mathrm{H}^r_{\mathrm{dR},?}(X, g)$ for $? \in \{\varnothing, c\}$. In this case, Esnault, Sabbah, and Yu showed [Esnault et al. 2017, Proposition 1.7.4] that the irregular Hodge filtration on the de Rham fiber coincides with the Yu filtration [2014] on the twisted de Rham cohomologies, where the two filtrations are denoted by $F^{\bullet}_{\mathrm{Del}} = F^{\mathrm{Del}}_{-\bullet}$ and $F^{\bullet}_{\mathrm{Yu}} = F^{\mathrm{Yu}}_{-\bullet}$, respectively, in [loc. cit.].

3.1.2. *Irregular Hodge filtration and Newton monomial filtration.* We briefly recall the definition of the irregular Hodge filtration on the twisted de Rham cohomology following [Yu 2014]. Let X and g be as above, $j : X \to \overline{X}$ a smooth compactification of X, and $D := \overline{X} \backslash X$ the boundary divisor. The pair $(\overline{X}, D)$ is called a *good compactification* of the pair (X, g) if D is normal crossing and g extends to a morphism $\bar{g} : \overline{X} \to \mathbb{P}^1$.

Let P be the pole divisor of g. The twisted de Rham complex $(\Omega^{\bullet}_{\overline{X}}(*D), \nabla = \mathrm{d} + \mathrm{d}g)$ admits a decreasing filtration $F^{\lambda}(\nabla) := F^0(\lambda)^{\geq \lceil \lambda \rceil}$, indexed by nonnegative real numbers λ, where $F^0(\lambda)$ is the Yu complex

$$\mathcal{O}_{\overline{X}}(\lfloor -\lambda P \rfloor) \xrightarrow{\nabla} \Omega^1_{\overline{X}}(\log D)(\lfloor (1-\lambda)P \rfloor) \to \cdots \to \Omega^p_{\overline{X}}(\log D)(\lfloor (p-\lambda)P \rfloor) \to \cdots .$$

The *irregular Hodge filtration* on the de Rham cohomology $\mathrm{H}^i_{\mathrm{dR}}(X, g)$ is defined by

$$F^{\lambda}_{\mathrm{irr}} \mathrm{H}^i_{\mathrm{dR}}(X, g) := \mathrm{im}(\mathbb{H}^i(\overline{X}, F^{\lambda}(\nabla)) \to \mathrm{H}^i_{\mathrm{dR}}(X, g)), \tag{3.1.0.3}$$

which is independent of the choice of the good compactification $(\overline{X}, D)$ [Yu 2014, Theorem 1.7].

When X is isomorphic to a torus $\mathbb{G}^n_m$, the regular function g on X is a Laurent polynomial of the form $\sum_{P=(p_1,\ldots,p_n)} c(P)x^P$. We refine the normal fan of the Newton polytope $\Delta(g)$ to make the associated toric variety X_{tor} smooth proper. Although $(X_{\mathrm{tor}}, D_{\mathrm{tor}} = X_{\mathrm{tor}} \backslash X)$ is not a good compactification for the pair (X, g) in general, we can still define $F^{\lambda}_{\mathrm{NP}}(\nabla)$ and the *Newton polyhedron filtration* $F^{\lambda}_{\mathrm{NP}} \mathrm{H}^i_{\mathrm{dR}}(U, \nabla)$ similarly to that in (3.1.0.3) by replacing the good compactification $(\overline{X}, D)$ with $(X_{\mathrm{tor}}, D_{\mathrm{tor}})$,

When g is nondegenerate with respect to $\Delta(g)$, the only nonvanishing twisted de Rham cohomology group of the pair (X, g) is the middle cohomology group $\mathrm{H}^n_{\mathrm{dR}}(X, g)$ by [Adolphson and Sperber 1997, Theorem 1.4]. Moreover, we have the following theorem.

Theorem 3.1.1 [Yu 2014, Theorem 4.6]. *When g is nondegenerate with respect to $\Delta(g)$, the irregular Hodge filtration $F^{\bullet}_{\mathrm{irr}}$ agrees with the Newton polyhedron filtration $F^{\bullet}_{\mathrm{NP}}$ on $\mathrm{H}^n_{\mathrm{dR}}(X, g)$.*

In particular, when g is nondegenerate, we have

$$\mathbb{H}^i(X_{\mathrm{tor}}, F^{\lambda}_{\mathrm{NP}}(\nabla)) = \mathrm{H}^i(\Gamma(X_{\mathrm{tor}}, F^{\lambda}_{\mathrm{NP}}(\nabla))),$$

which allows us to compute the irregular Hodge filtration using the knowledge of $\Delta(g)$.

Now, we present an explicit way to calculate the Newton polyhedron filtration. For a cohomology class

$$\omega = x^Q \frac{\mathrm{d}x_1}{x_1} \wedge \cdots \wedge \frac{\mathrm{d}x_n}{x_n}$$

such that the lattice point $Q = (q_1, \ldots, q_n)$ lies in $\mathbb{R}_{\geq 0}\Delta(g)$, we define $w(Q)$ to be the weight of Q in the sense of [Adolphson and Sperber 1997], i.e., the minimal positive real number w such that $Q \in w \cdot \Delta(g)$. The associated cohomology class of ω lies in $F^{\lambda}_{\mathrm{NP}} \mathrm{H}^n_{\mathrm{dR}}(X, g)$ if

$$\omega \in \Gamma(X_{\mathrm{tor}}, \Omega^n_{X_{\mathrm{tor}}}(\log D_{\mathrm{tor}})(\lfloor (n-\lambda)P \rfloor)).$$

Notice that each ray ρ in the normal fan of $\Delta(g)$ corresponds to an irreducible component P_ρ of P. Let v_ρ be a primitive vector of the ray ρ. Then, the multiplicity of ω along P_ρ is given by $\langle Q, v_\rho\rangle$ [Fulton 1993, p. 61]. Taking the multiplicities of P_ρ in P into account, we have

$$x^Q \frac{\mathrm{d}x_1}{x_1} \wedge \cdots \wedge \frac{\mathrm{d}x_n}{x_n} \in F_{\mathrm{NP}}^{n-w(Q)} \mathrm{H}^n_{\mathrm{dR}}(X, g), \tag{3.1.1.1}$$

as remarked in [Yu 2014, p. 126 footnote].

3.1.3. *The* EMHS *associated with hypergeometric connections.* In this subsection, we assume $\alpha_1 = 0$ and let $\tilde{\chi} \times \rho$ be the product of characters associated with α_i and β_j in (2.3.0.1).

Definition 3.1.2. Let K be the number field $\mathbb{Q}(\zeta_d^{a_i}, \zeta_d^{b_j})$ and a a closed point of S. For $? \in \{\varnothing, c\}$, we define

$$E_?(a; \alpha; \beta) := \mathrm{H}^{n+m-1}(\mathbb{G}_m^{n+m-1}, f_a)^{(G, \tilde{\chi}\times\rho)}$$

as exponential mixed Hodge structures with coefficients in K in the sense of Section 3.1.1.

By Corollary 2.3.3 and the base change theorem, the de Rham fiber of $E(a; \alpha; \beta)$ is isomorphic to the fiber of $\mathcal{H}yp_\lambda(\alpha; \beta)$ at the closed point $a \cdot \lambda$ of S for $\lambda \in \mathbb{Q}^\times$. In other words, the fiber $\mathcal{H}yp_\lambda(\alpha; \beta)_{a\lambda}$ underlies the above exponential mixed Hodge structure and is equipped with an irregular Hodge filtration F_{irr}, which coincides with the Yu filtration on $\mathrm{H}^{n+m-1}_{\mathrm{dR}}(\mathbb{G}_m^{n+m-1}, f_a)^{(G, \tilde{\chi}\times\rho)}$ as explained in Section 3.1.1.

Remark 3.1.3. The geometric interpretations of the hypergeometric connection are not unique, and their associated Yu filtrations on $\mathcal{H}yp(\alpha; \beta)_a$ coincide only up to certain shifts. For example, we can alternatively identify $\mathcal{H}yp(\alpha; \beta)_a$ with

$$\mathrm{H}^{n+m-1}_{\mathrm{dR}}(\mathbb{G}_m^{n+m-1}, f_a)^{(G, \tilde{\chi}\times\rho)} \otimes \mathrm{H}^1_{\mathrm{dR}}(\mathbb{G}_m, x),$$

where the Yu filtration on the one-dimensional vector space $\mathrm{H}^1_{\mathrm{dR}}(\mathbb{G}_m, x)$ jumps at 1. Consequently, we deduce a new irregular Hodge filtration on $\mathcal{H}yp(\alpha; \beta)_a$ which differs from our current one by a shift of 1. For this reason, we made a choice of a uniform shift for irregular Hodge filtrations on fibers of $\mathcal{H}yp(\alpha; \beta)$ at closed points of S, using the exponential mixed Hodge structures in Definition 3.1.2. Moreover, this specifically chosen shift determines the shift of the function θ in (1.0.1.1).

Let t be the largest natural number such that $\alpha_t = 0$. We let $\bar{\alpha}$ and $\bar{\beta}$ be the sequences of rational numbers defined by

$$\bar{\alpha}_i = \begin{cases} 0, & 1 \le k \le t, \\ 1 - \alpha_{n+t+1-k}, & t+1 \le k \le n, \end{cases} \quad \text{and} \quad \bar{\beta}_k = 1 - \beta_k. \tag{3.1.3.1}$$

Proposition 3.1.4. (1) *The dual of the exponential mixed Hodge structure $E_c(a; \alpha; \beta)$ is isomorphic to $E((-1)^{n-m}a; \bar{\alpha}; \bar{\beta})(n+m-1)$.*

(2) *When (α, β) is nonresonant, the exponential mixed Hodge structures $E_?(a; \alpha; \beta)$ for $? \in \{\varnothing, c\}$ are isomorphic. In particular, they are pure of weight $n+m-1$.*

Proof. (1) The EMHS $\mathrm{H}_c^{n+m-1}(\mathbb{G}_m^{n+m-1}, f_a)$ is dual to

$$\mathrm{H}^{n+m-1}(\mathbb{G}_m^{n+m-1}, -f_a) \otimes \mathrm{H}_c^{2n+2m-2}(\mathbb{G}_m^{n+m-1})^\vee,$$

which is also isomorphic to $\mathrm{H}^{n+m-1}(\mathbb{G}_m^{n+m-1}, f_{(-1)^{n-m}a}) \otimes \mathrm{H}_c^{2n+2m-2}(\mathbb{G}_m^{n+m-1})^\vee$. We deduce the first assertion by taking their corresponding isotypic components.

(2) Since the de Rham fiber functor is faithful, the forget-support morphism

$$E_c(a; \alpha; \beta) \to E(a; \alpha; \beta)$$

is an isomorphism by Corollary 2.3.4. Hence the exponential mixed Hodge structures $E_c(a; \alpha; \beta)$ and $E(a; \alpha; \beta)$ are isomorphic and are pure of weight $n+m-1$. □

3.2. *A basis in relative twisted de Rham cohomology.* In this subsection, we assume $\alpha_1 = 0$. We define positive integers $s_1, \dots, s_{m+1}$ by

$$s_r = \begin{cases} 1, & r = 0, \\ \#\{i : \alpha_i < \beta_r\}, & 1 \le r \le m, \\ n+1, & r = m+1, \end{cases} \tag{3.2.0.1}$$

and, for r and ℓ such that $0 \le r \le m$ and $1 \le \ell \le s_{r+1} - s_r$, we set

$$g_{r,\ell} = x_2^{a_2} \cdots x_{s_r+\ell-1}^{a_{s_r+\ell-1}} \cdot x_{s_r+\ell}^{a_{s_r+\ell}-d} \cdots x_n^{a_n-d} \cdot y_1^{d-b_1} \cdots y_r^{d-b_r} \cdot y_{r+1}^{2d-b_{r+1}} \cdots y_m^{2d-b_m}.$$

Let

$$\eta = \frac{\mathrm{d}x_2}{x_2} \cdots \frac{\mathrm{d}x_n}{x_n} \frac{\mathrm{d}y_1}{y_1} \cdots \frac{\mathrm{d}y_m}{y_m}$$

and $\omega_{r,\ell} = g_{r,\ell} \cdot \eta$ be the corresponding differential forms in $\mathcal{H}_{\mathrm{dR}}^{n+m-1}(U/S, \pm f)^{(G,\tilde{\chi}\times\rho)}$, where U and S are defined in (2.3.0.3).

Proposition 3.2.1. *If (α, β) is nonresonant, then the cohomology classes defined by*

$$\omega_{r,\ell}, \quad 0 \le r \le m, \quad 1 \le \ell \le s_{r+1} - s_r$$

in $\mathcal{H}_{\mathrm{dR}}^{n+m-1}(U/S, \pm f)^{(G,\tilde{\chi}\times\rho)}$ form a basis over $\mathcal{O}_S$.

Proof. It suffices to show that $\mathrm{Span}(\omega_{r,\ell}) = \mathrm{Span}(D^i\omega \mid 0 \le i \le n-1)$ for a cyclic vector ω.

To a Laurent monomial $g = \prod_{i=2}^n x_i^{u_i} \prod_{j=1}^m y_j^{v_j}$ in variables $\{x_i\}_{i=2}^n$ and $\{y_j\}_{j=1}^m$ we associate a lattice point $\mathcal{P}(g) = (u_2, \dots, u_n, v_1, \dots, v_m) \in \mathbb{Z}^{n+m-1} \subset \mathbb{R}^{n+m-1}$. If $\omega = g \cdot \eta$ is the product of a monomial g with the differential form η, we set $\mathcal{P}(\omega) := \mathcal{P}(g)$ for the corresponding point.

Let π_1 and π_2 be the projections from $\mathbb{R}^{n+m-1}$ to $\mathbb{R}_{u_i}^{n-1}$ and $\mathbb{R}_{v_j}^m$, respectively. Then, for the differential forms $\omega_{r,\ell}$, we have

$$\pi_1(\mathcal{P}(\omega_{r,\ell})) = (a_2, \dots, a_{s_r+\ell-1}, a_{s_r+\ell} - d, \dots, a_n - d)$$

and

$$\pi_2(\mathcal{P}(\omega_{r,\ell})) = (d - b_1, \dots, d - b_r, 2d - b_{r+1}, \dots, 2d - b_m).$$

In Lemmas 2.2.2, 2.2.3, and 2.2.4, we have written down the defining inequalities of the cone $\mathbb{R}_{\geq 0}\cdot\Delta(f_a)$ explicitly as $u_i+v_j\geq 0$ and $v_j\geq 0$. For the points $\mathcal{P}(\omega_{r,\ell})$ corresponding to the values of $\omega_{r,\ell}$ at closed points a of S, we can verify that they satisfy $u_i+v_j\geq 0$ and $v_j\geq 0$ using the fact that $a_i, b_j\leq d$ for any i and j, and $a_i\geq b_j$ when $i\geq s_j+1$. Thus, all the points $\mathcal{P}(\omega_{r,\ell})$ lie within the cone $\mathbb{R}_{\geq 0}\cdot\Delta(f_a)$.

Let P_i and Q_j be the points corresponding to monomials x_i^d and y_j^d, respectively, for $2\leq i\leq n$ and $1\leq j\leq m$.

Lemma 3.2.2. *For a point $P\in\mathbb{Z}^{n+m-1}$ and two integers $2\leq i_0\leq n$ and $1\leq j_0\leq m$, let ω_0, ω_1, and ω_2 be the corresponding differential forms of the points P, $P+Q_{j_0}$, and $P+P_{i_0}$ in $\mathbb{Z}^{n+m-1}$. If the i_0-th coordinate of P is different from the negative of the j_0-th coordinate of P, then we have*

$$\mathrm{Span}(\omega_0,\omega_2)=\mathrm{Span}(\omega_1,\omega_2)\quad \textit{in } \mathrm{H}^{n+m-1}_{\mathrm{dR}}(U/S,\ f)^{(G,\tilde{\chi}\times\rho)}.$$

Proof. Let P be the point $(t_i, s_j)\in\mathbb{Z}^{n+m-1}$ and ω_0 be the associated differential form. By assumption, we have $t_{i_0}\neq -s_{j_0}$. Therefore, we can express

$$\omega_0=\frac{-d}{t_{i_0}+s_{j_0}}\left(\left(D-\frac{t_{i_0}}{d}\right)\omega_0-\left(D+\frac{s_{j_0}}{d}\right)\omega_0\right).$$

In particular, we have

$$\mathrm{Span}\left(\omega_0,\left(D-\frac{t_{i_0}}{d}\right)\omega_0\right)=\mathrm{Span}\left(\left(D-\frac{t_{i_0}}{d}\right)\omega_0,\left(D+\frac{s_{j_0}}{d}\right)\omega_0\right).$$

At last, notice that we have

$$\omega_1=\left(D+\frac{s_{j_0}}{d}\right)\omega_0\quad\text{and}\quad\omega_2=\left(D-\frac{t_{i_0}}{d}\right)\omega_0$$

by Lemma 2.4.3 and Remark 2.4.5. □

Step 1: If $s_1-s_0=0$, we skip this step and put $\omega^{(1)}_{r,\ell}=\omega_{r,\ell}$ for any r and ℓ. Otherwise, for $r=0$ and $1\leq\ell\leq s_1-s_0$, we replace the differential forms $\omega_{0,\ell}$ by differential forms $\omega^{(1)}_{0,\ell}$ of the form $g\cdot\eta$ for some monomials g such that

$$\mathcal{P}(\omega^{(1)}_{0,\ell})=\mathcal{P}(\omega_{0,\ell})-Q_1.$$

More precisely, we keep the first $n-1$ coordinates of $\mathcal{P}(\omega_{0,\ell})$ unchanged and replace the last m coordinates of $\mathcal{P}(\omega_{0,\ell})$ by that of $\mathcal{P}(\omega^{(1)}_{0,\ell})$:

$$(d-b_1, 2d-b_2,\ldots,2d-b_m).$$

In particular, by Lemma 2.4.3, one has

$$(D+1-\beta_1)\omega^{(1)}_{0,\ell}=\omega_{0,\ell},\qquad (D+1-\alpha_{\ell+1})\omega^{(1)}_{0,\ell}=\omega^{(1)}_{0,\ell+1},$$

and

$$(D+1-\alpha_{s_1-s_0})\omega^{(1)}_{0,s_1-s_0}=\omega_{e,1},$$

where e is the least integer such that $s_e>s_0=1$.

We also put $\omega^{(1)}_{r,\ell} = \omega_{r,\ell}$ for $r \geq 1$. Then, using Lemma 3.2.2 for $\omega_0 = \omega^{(1)}_{0,s_1-s_0}$, $\omega_1 = \omega_{0,s_1-s_0}$, and $\omega_2 = \omega_{e,1}$, we have

$$\begin{aligned}\mathrm{Span}\{\omega_{r,\ell} \mid r, \ell\} &= \mathrm{Span}\big(\ldots, \omega_{0,s_1-s_0}\big(= (D+1-\beta_1)\omega^{(1)}_{0,s_1-s_0}\big), \omega_{e,1}\big(= (D+1-\alpha_{s_1-s_0})\omega^{(1)}_{0,s_1-s_0}\big), \ldots\big)\\ &= \mathrm{Span}(\{\omega_{0,1}, \ldots, \omega_{0,s_1-s_0-1}, \omega^{(1)}_{0,s_1-s_0}\} \cup \{\omega^{(1)}_{r,\ell} \mid r \geq 1, \ell\}),\end{aligned}$$

where $0 \leq r \leq m$ and $1 \leq \ell \leq s_{r+1} - s_r$. Continuing to use Lemma 3.2.2 for $\omega_0 = \omega^{(1)}_{0,\ell}$, $\omega_1 = \omega_{0,\ell}$, and $\omega_2 = \omega^{(1)}_{0,\ell+1}$ for $\ell = s_1 - s_0 - 1$ and $s_1 - s_0 - 2, \ldots, 1$, we have

$$\begin{aligned}\mathrm{Span}\{\omega_{r,\ell} \mid r, \ell\} &= \mathrm{Span}(\{\omega_{0,1}, \ldots, \omega_{0,s_1-s_0-1}, \omega^{(1)}_{0,s_1-s_0}\} \cup \{\omega^{(1)}_{r,\ell} \mid r \geq 1, \ell\})\\ &= \mathrm{Span}(\{\omega_{0,1}, \omega^{(1)}_{0,2}, \ldots, \omega^{(1)}_{0,s_1-s_0}\} \cup \{\omega^{(1)}_{r,\ell} \mid r \geq 1, \ell\})\\ &= \mathrm{Span}(\omega^{(1)}_{r,\ell} \mid r, \ell).\end{aligned}$$

Step $i \geq 2$: Assume that we have already obtained elements $\omega^{(i-1)}_{r,\ell}$ for $i \geq 2$. If $s_i = s_{i-1}$, we skip this step and put $\omega^{(i)}_{r,\ell} = \omega^{(i-1)}_{r,\ell}$ for any r and ℓ. Otherwise, let $\omega^{(i)}_{r,\ell}$ be differential forms of the form $g \cdot \eta$ for some monomials g such that

$$\mathcal{P}(\omega^{(i)}_{r,\ell}) = \begin{cases} \mathcal{P}(\omega^{(i-1)}_{r,\ell}) - Q_i & \text{if } r \leq i-1,\\ \mathcal{P}(\omega^{(i-1)}_{r,\ell}) & \text{if } i \leq r \leq m.\end{cases}$$

More precisely, when $r \leq i-1$, we keep the first $n-1$ coordinates of $\mathcal{P}(\omega^{(i-1)}_{r,\ell})$ unchanged and replace the last m coordinates of $\mathcal{P}(\omega^{(i-1)}_{r,\ell})$ by that of $\mathcal{P}(\omega^{(i)}_{r,\ell})$:

$$(d - b_1, \ldots, d - b_i, 2d - b_{i+1}, \ldots, 2d - b_m).$$

Similar to Step 1, we use Lemmas 2.4.3 and 3.2.2 $(s_{r+1} - s_r)$-many times to deduce

$$\mathrm{Span}(\omega^{(i)}_{r,\ell} \mid r, \ell) = \mathrm{Span}(\omega^{(i-1)}_{r,\ell} \mid r, \ell) = \mathrm{Span}(\omega_{r,\ell} \mid r, \ell),$$

where $0 \leq r \leq m$ and $1 \leq \ell \leq s_{r+1} - s_r$.

After Step m: After m steps, we get $\omega^{(m)}_{r,\ell}$ such that

$$\mathcal{P}(\omega^{(m)}_{r,\ell}) = (a_2, \ldots, a_{s_r+\ell-1}, a_{s_r+\ell} - d, \ldots, a_n - d, d - b_1, \ldots, d - b_m).$$

Note that there is a bijection between $\{(r, \ell)\}_{0 \leq r \leq m, 1 \leq \ell \leq s_{r+1}-s_r}$ and $\{1, \ldots, n\}$ by sending (r, ℓ) to $s_r + \ell - 1$. We set $\tilde{\omega}_{s_r+\ell-1} = \omega^{(m)}_{r,\ell}$ via this map. Then, by Lemma 2.4.3, we have

$$\tilde{\omega}_{i+1} = (D + 1 - \alpha_{i+1})\tilde{\omega}_i \quad \text{for } 1 \leq i \leq n-1.$$

It follows that

$$\begin{aligned}\mathrm{Span}(D^i \tilde{\omega}_1 \mid 0 \leq i \leq n-1) &= \mathrm{Span}(\tilde{\omega}_i \mid 1 \leq i \leq n)\\ &= \mathrm{Span}(\omega^{(m)}_{r,\ell} \mid r, \ell) = \mathrm{Span}(\omega_{r,\ell} \mid r, \ell).\end{aligned}$$

By Proposition 2.4.1 and Remark 2.4.2, $\tilde{\omega}_1$ is a cyclic vector, from which we showed that $\{\omega_{r,\ell}\}_{r,\ell}$ form a basis. This finishes the proof. $\square$

3.3. ***Calculation of the irregular Hodge filtration.*** Using the fact that a nonresonant hypergeometric connection is rigid or its geometric interpretation Corollary 2.3.3, it underlies an irregular Hodge module on $\mathbb{P}^1$ of weight $n+m-1$ by [Sabbah 2018, Theorem 0.7 & p. 78] and, therefore, admits a unique irregular Hodge filtration $F^{\bullet}_{\mathrm{irr}}$. When $n=m$, this irregular Hodge module coincides with the variation of Hodge structures on $\mathcal{H}yp(\alpha;\beta)$.

Recall that, for (α,β), we defined in (1.0.1.1) the numbers

$$\theta(k)=(n-m)\alpha_k+\#\{i\mid\beta_i<\alpha_k\}+(n-k)-\sum_{i=1}^{n}\alpha_i+\sum_{j=1}^{m}\beta_j.$$

Theorem 3.3.1. *Assume (α,β) is nonresonant.*

(1) *When $\alpha_1=0$, via the isomorphism $\mathcal{H}yp(\alpha;\beta)\simeq\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S,f)^{(G,\tilde\chi\times\rho)}$, the irregular Hodge filtration on $\mathcal{H}yp(\alpha;\beta)$ can be identified with the following filtration of subbundles*:

$$F^{p}_{\mathrm{irr}}\mathcal{H}^{n+m-1}_{\mathrm{dR}}(U/S,f)^{(G,\tilde\chi\times\rho)}=\bigoplus_{n+m-1-w(\omega_{r,s})\geq p}\omega_{r,s}\mathcal{O}_S.$$

(2) *Up to an $\mathbb{R}$-shift, the jumps of the irregular Hodge filtration on $\mathcal{H}yp(\alpha;\beta)$ occur at $\theta(k)$ and, for any $p\in\mathbb{R}$, we have*

$$\operatorname{rk}\operatorname{gr}^{p}_{F_{\mathrm{irr}}}\mathcal{H}yp(\alpha;\beta)=\#\theta^{-1}(p).$$

Remark 3.3.2. (i) By [Sabbah and Yu 2015, Remark 6.3], the irregular Hodge filtration satisfies the Griffiths' transversality; that is, $\nabla(F^{p}_{\mathrm{irr}}\mathcal{H}yp(\alpha;\beta))\subset\Omega^1_S\otimes F^{p-1}_{\mathrm{irr}}\mathcal{H}yp(\alpha;\beta)$ for all $p\in\mathbb{R}$.

(ii) Inspired by the Griffiths' transversality, we expect that there exists an oper structure on the hypergeometric connections which refines the irregular Hodge filtration. An oper structure is essential in the geometric Langlands correspondence [Beilinson and Drinfeld 1997; Zhu 2017; Kamgarpour et al. 2023].

To prove the above theorem, we study the Hodge numbers of the irregular Hodge filtration on fibers as explained in Section 1.2.

Theorem 3.3.3. *Up to an $\mathbb{R}$-shift, the jumps of the irregular Hodge filtration $F^{\bullet}_{\mathrm{irr}}$ on the fiber $\mathcal{H}yp(\alpha;\beta)_a$ occur at $\theta(k)$ for $1\leq k\leq n$. Moreover, we have* $\dim\operatorname{gr}^{p}_{F_{\mathrm{irr}}}\mathcal{H}yp(\alpha;\beta)_a=\#\theta^{-1}(p)$ *for any $p\in\mathbb{R}$.*

3.3.1. *Proof of Theorem 3.3.3.* We may assume $\alpha_1=0$ by (2.1.1.2). By Corollary 2.3.3 and Definition 3.1.2, we have

$$\begin{aligned}F^{\bullet}_{\mathrm{irr}}\mathcal{H}yp(\alpha;\beta)_a&\simeq F^{\bullet}_{\mathrm{irr}}\mathrm{H}^{n+m-1}_{\mathrm{dR}}(\mathbb{G}_m^{n+m-1},f_a)^{(G,\tilde\chi\times\rho)}\\&\simeq F^{\bullet}_{\mathrm{irr}}\mathrm{H}^{n+m-1}_{\mathrm{dR}}(\mathbb{G}_m^{n+m-1},-f_{(-1)^{n-m}a})^{(G,\tilde\chi\times\rho)},\end{aligned}\tag{3.3.3.1}$$

where $\tilde\chi$ and ρ are products of characters corresponding to α_i and β_j from (2.3.0.1). So it suffices to compute the irregular Hodge filtration on the twisted de Rham cohomologies $\mathrm{H}^{n+m-1}_{\mathrm{dR}}(\mathbb{G}_m^{n+m},\pm f_a)^{(G,\tilde\chi\times\rho)}$. Since f_a is nondegenerate with respect to $\Delta(f_a)$, we can compute the filtration in terms of Newton polyhedron filtration.

Let $\omega_{r,\ell}$ be the basis of $\mathcal{H}yp(\alpha;\beta)_a$ from Proposition 3.2.1. Recall that $w(\omega_{r,\ell})$ is the minimal positive real number w such that $\mathcal{P}(g_{r,\ell}) \in w\cdot\Delta(f_a)$. It follows from (3.1.1.1) that

$$\omega_{r,\ell} \in F_{\mathrm{irr}}^{n+m-1-w(\omega_{r,\ell})}\mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)}.$$

We consider an auxiliary filtration $G^\bullet$ on $\mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)}$ defined by

$$G^p := \mathrm{Span}\{\omega_{r,\ell} \mid n+m-1-w(\omega_{r,\ell}) \geq p\}. \tag{3.3.3.2}$$

By the following Lemmas 3.3.4, 3.3.5, and 3.3.6, the filtration $F^\bullet$ coincides with $G^\bullet$, which finishes the proof of the theorem. □

Lemma 3.3.4. *We set $\theta(n+1)=\theta(1)$. For $0\leq r\leq m$ and $1\leq \ell\leq s_{r+1}-s_r$, we have*

$$n+m-1-w(\omega_{r,\ell}) = \theta(s_r+\ell).$$

Lemma 3.3.5. *For $0\leq p\leq n+m-1$, we have*

$$\dim \mathrm{gr}_G^p\, \mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)} = \dim \mathrm{gr}_G^{n+m-1-p}\, \mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \mp f_a)^{(G,\tilde{\chi}^{-1}\times\rho^{-1})}.$$

Lemma 3.3.6. *The two filtrations $F_{\mathrm{irr}}^\bullet$ and $G^\bullet$ coincide.*

Proof of Lemma 3.3.4. By Lemmas 2.2.2, 2.2.3, and 2.2.4, the weight $w(\omega_{r,\ell})$ equals the number $\max_k\{h_k(g_{r,\ell})/d\}$, where the h_k are defined in (2.2.2.1), (2.2.3.1), and (2.2.4.1). We can check that

$$w(\omega_{r,\ell}) = \frac{h_{s_r+\ell}(g_{r,\ell})}{d},$$

where we put $h_1=\cdots=h_n=h_{n+1}$ when $n=m$. Now, it suffices to check that $n+m-1-w(\omega_{r,\ell})$ agrees with one of the jumps of the irregular Hodge numbers of $\mathcal{H}yp(\alpha;\beta)_a$.

If $s_r+\ell=n+1$, the monomial $g_{m,n+1-s_m}$ corresponds to the point

$$(a_2,\ldots,a_n,d-b_1,\ldots,d-b_m).$$

Then we have

$$n+m-1-\frac{h_{n+1}(g_{m,n+1-s_m})}{d} = n-1-\sum_{i=1}^n\alpha_i+\sum_{j=1}^m\beta_j = \theta(1).$$

If $s_r+\ell<n+1$, we have

$$\begin{aligned}
n+m-1-&\frac{h_{s_r+\ell}(g_{r,\ell})}{d}\\
&= n+m-1-\Big(\sum_{i=1}^n\alpha_i-(n+1-s_r-\ell)-\sum_{j=1}^m\beta_j+(2m-r)-(n-m)(\alpha_{s_r+\ell}-1)\Big)\\
&= (n-m)\alpha_{s_r+\ell}+r+(n-s_r-\ell)-\sum_{i=1}^n\alpha_i+\sum_{j=1}^m\beta_j,
\end{aligned}$$

which is exactly $\theta(s_r+\ell)$. □

Proof of Lemma 3.3.5. For simplicity, we write

$$\delta_p^{\pm}(\alpha,\beta) := \dim \mathrm{gr}_G^p \mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)}. \tag{3.3.6.1}$$

Recall that, in (3.1.3.1), we let t be the biggest natural number such that $\alpha_t = 0$. For $1 \le k \le t$, the numbers α_k and $\bar{\alpha}_{t+1-k}$ are 0. And, for $t+1 \le k \le n$, we have $\bar{\alpha}_{n-k+t+1} = 1-\alpha_k$. Then

$$\sum_{i=1}^n \alpha_i + \sum_{i=1}^n \bar{\alpha}_i = n-t \quad \text{and} \quad \sum_{j=1}^m \beta_j + \sum_{j=1}^m \bar{\beta}_j = m.$$

Similar to the number $\theta(k)$, we let $\bar{\theta}(k)$ be the numbers

$$(n-m)\bar{\alpha}_k + \#\{i \mid \bar{\beta}_i < \bar{\alpha}_k\} + (n-k) - \sum_{i=1}^n \bar{\alpha}_i + \sum_{j=1}^m \bar{\beta}_j, \quad 1 \le k \le n,$$

for the sequences $\bar{\alpha}$ and $\bar{\beta}$. Then, for $1 \le k \le t$, we have

$$\begin{aligned}\theta(k) + \bar{\theta}(t+1-k) &= \left(n-k-\sum_{i=1}^n \alpha_i + \sum_{j=1}^m \beta_j\right) + \left(n-(t+1-k) - \sum_{i=1}^n \bar{\alpha}_i + \sum_{j=1}^m \bar{\beta}_j\right)\\ &= (2n-t-1) - (n-t) + m = n+m-1.\end{aligned}$$

For $t+1 \le k \le n$, we have

$$\begin{aligned}&\theta(k) + \bar{\theta}(n-k+t+1)\\ &\quad= \left((n-m)\alpha_k + \#\{j \mid \beta_j < \alpha_k\} + n-k-\sum_{i=1}^n \alpha_i + \sum_{j=1}^m \beta_j\right)\\ &\qquad\qquad + \left((n-m)\bar{\alpha}_{n-k+t+1} + \#\{j \mid \bar{\beta}_j < \bar{\alpha}_{n-k+t+1}\} + n-(n-k+t+1) - \sum_{i=1}^n \bar{\alpha}_i + \sum_{j=1}^m \bar{\beta}_j\right)\\ &\quad= (n-m) + m + (n-t-1) - (n-t) + m = n+m-1.\end{aligned}$$

So there exists a permutation $\sigma \in S_n$ such that $\theta(k) + \bar{\theta}(\sigma(k)) = n+m-1$. It follows that

$$\begin{aligned}\delta_p^{\pm}(\alpha,\beta) &= \#\{k \mid \theta(k) = p\} = \#\{k \mid n+m-1-p = n+m-1-\theta(k)\}\\ &= \#\{k \mid \bar{\theta}(k) = n+m-1-p\} = \delta^{\mp}_{n+m-1-p}(\bar{\alpha},\bar{\beta}).\end{aligned}$$

□

Proof of Lemma 3.3.6. For simplicity, we write

$$h_p^{\pm}(\alpha,\beta) := \dim \mathrm{gr}_{F_{\mathrm{irr}}}^p \mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)}. \tag{3.3.6.2}$$

By (3.1.1.1) and the construction of the auxiliary filtration G (3.3.3.2), for every $p \in \mathbb{Q}$, we have

$$G^p \mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)} \subset F_{\mathrm{irr}}^p \mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)}, \tag{3.3.6.3}$$

which implies that $\sum_{q\le p} \delta_q^{\pm}(\alpha,\beta) \le \sum_{q \le p} h_q^{\pm}(\alpha,\beta)$.

To prove the reverse inclusion, we consider the duality between the two filtered vector spaces

$$(\mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)}, F_{\mathrm{irr}}^{\bullet}) \quad \text{and} \quad (\mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \mp f_a)^{(G,\tilde{\chi}^{-1}\times\rho^{-1})}, F_{\mathrm{irr}}^{\bullet}),$$

induced by Proposition 3.1.4. In particular, we have

$$h_p^{\pm}(\alpha, \beta) = h_{n+m-1-p}^{\mp}(\bar{\alpha}, \bar{\beta}). \tag{3.3.6.4}$$

Combining Lemma 3.3.5 and equations (3.3.6.3) and (3.3.6.4), we see for any $p \in \mathbb{R}$ that

$$\begin{aligned}\dim G^p\mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)} &= \sum_{q\le p}\delta_q^{\pm}(\alpha,\beta) \le \sum_{q\le p} h_q^{\pm}(\alpha,\beta) = \sum_{q\ge n+m-1-p} h_q^{\mp}(\bar{\alpha},\bar{\beta})\\ &\le \sum_{q\ge n+m-1-p}\delta_q^{\mp}(\bar{\alpha},\bar{\beta}) = \sum_{q\le p}\delta_q^{\pm}(\alpha,\beta)\\ &= \dim G^p\mathrm{H}_{\mathrm{dR}}^{n+m-1}(\mathbb{G}_m^{n+m-1}, \pm f_a)^{(G,\tilde{\chi}\times\rho)}.\end{aligned}$$

Hence both sides in (3.3.6.3) have the same dimension for every p. Then Lemma 3.3.6 follows. □

3.3.2. *Proof of Theorem 3.3.1.* We may assume $\alpha_1 = 0$ by (2.1.1.2). By [Sabbah 2018, Proposition 3.54] and [Mochizuki 2025, Proposition 11.22], the irregular Hodge filtration on $\mathcal{H}yp(\alpha;\beta)$ induces those on fibers $\mathcal{H}yp(\alpha;\beta)_a$ at closed points of S; i.e.,

$$(F_{\mathrm{irr}}^{\bullet}\mathcal{H}yp(\alpha;\beta))_a = F_{\mathrm{irr}}^{\bullet}(\mathcal{H}yp(\alpha;\beta))_a.$$

We have shown in Theorem 3.3.3 that the irregular Hodge filtration on the fibers $\mathcal{H}yp(\alpha;\beta)_a$ are given in terms of the cohomology classes $\omega_{r,s}$ in (3.3.3.2). Hence we deduce that the irregular Hodge filtration on $\mathcal{H}yp(\alpha;\beta)$ is the one in assertion (1).

From (1), we deduce that the irregular Hodge numbers $\operatorname{rk}\operatorname{gr}_{F_{\mathrm{irr}}}^p \mathcal{H}yp(\alpha;\beta)$ are given by

$$\#\{(r,s) \mid n+m-1-w(\omega_{r,s}) = p\}.$$

Recall that we have the bijection between the sets

$$\{1,\dots,n\} \quad \text{and} \quad \{(r,\ell) \mid 0\le r\le m,\ 1\le \ell\le s_{r+1}-s_r\},$$

where s_r are numbers defined in (3.2.0.1). Using Lemma 3.3.4, we deduce assertion (2); i.e., the irregular Hodge numbers $\operatorname{rk}\operatorname{gr}_{F_{\mathrm{irr}}}^p \mathcal{H}yp(\alpha;\beta)$ coincide with the numbers $\#\theta^{-1}(p)$. □

4. Frobenius structures on hypergeometric connections and p-adic estimates

In this section, let p be a prime number and $k = \mathbb{F}_q$ the finite field with $q = p^s$ elements for an integer $s \ge 1$. Let K be a finite extension of $\mathbb{Q}_p$ with residue field k containing an element π satisfying $\pi^{p-1} = -p$. We fix such an element π and denote the associated additive character by $\psi : \mathbb{F}_p \to K^\times$ [Berthelot 1984, (1.3)]. The q-th power Frobenius on k admits a lift $\sigma = \mathrm{id}$ on $\mathcal{O}_K$.

Let $n > m$ be two integers,

$$\alpha = \left(\alpha_i = \frac{a_i}{q-1}\right)_{i=1}^{n} \quad \text{and} \quad \beta = \left(\beta_j = \frac{b_j}{q-1}\right)_{j=1}^{m}$$

be two sequences of nondecreasing rational numbers in $[0, 1)$ with denominator $q-1$. Let $\omega : k^\times \to K^\times$ be the Teichmüller character, and set $\chi_i = \omega^{a_i}$ and $\rho_j = \omega^{b_j}$. The hypergeometric sum associated to ψ, $\chi = (\chi_1, \dots, \chi_n)$ and $\rho = (\rho_1, \dots, \rho_m)$ is defined, for $a \in k^\times$, by

$$\mathrm{Hyp}_{(n,m)}(\chi; \rho)(a) = \sum_{\substack{x_i, y_j \in k^\times \\ x_1 \cdots x_n = a y_1 \cdots y_m}} \psi\left(\mathrm{Tr}_{k/\mathbb{F}_p}\left(\sum_{i=1}^{n} x_i - \sum_{j=1}^{m} y_j\right)\right) \cdot \prod_{i=1}^{n} \chi_i(x_i) \prod_{j=1}^{m} \rho_j^{-1}(y_j). \tag{4.0.0.1}$$

When (χ, ρ) is nonresonant, the above sum equals (up to a sign) the Frobenius trace of the hypergeometric overconvergent F-isocrystal $\mathscr{H}yp(\chi, \rho)$ at $a \in \mathbb{G}_m(k)$ [Miyatani 2020] and therefore can be written as a sum of n Frobenius eigenvalues. Its underlying connection is the hypergeometric connection $\mathcal{H}yp_{(-1)^{m+np}/\pi^{n-m}}(\alpha; \beta)$ [Miyatani 2020, Theorem 4.1.3]. When (χ, ρ) is resonant, the above sum can also be written as a sum of n Frobenius eigenvalues (see Section 4.2.1 for a direct proof by induction on n).

We are interested in the p-adic valuation of Frobenius eigenvalues (normalized by ord_q) of the above sum (called *Frobenius slopes*), encoded in the Frobenius Newton polygon [Mazur 1972, §2].

Recall that the irregular Hodge numbers of the hypergeometric connection $\mathcal{H}yp(\alpha; \beta)$ are given by the function $\theta : \{1, \dots, n\} \to \mathbb{Q}$ (1.0.1.1), defined by

$$\theta(k) = (n-m)\alpha_k + \#\{i \mid \beta_i < \alpha_k\} + (n-k) - \sum_{i=1}^{n} \alpha_i + \sum_{j=1}^{m} \beta_j. \tag{4.0.0.2}$$

Definition 4.0.1. Let $\delta_1 < \cdots < \delta_k$ be the Frobenius slopes of $\mathrm{Hyp}_{(n,m)}(\chi; \rho)(a)$, normalized by $\mathrm{ord}_q(q) = 1$, (resp. irregular Hodge numbers of $\mathcal{H}yp(\alpha, \beta)$) with multiplicities $\lambda_1, \dots, \lambda_k$. The Newton polygon (resp. irregular Hodge polygon) is defined as the union of segments in $\mathbb{R}^2$ joining points P_i and P_{i+1} for $0 \le i \le k-1$, where the P_i are given by

$$P_0 = (0, 0), \qquad P_i = \left(\sum_{j=1}^{i} \lambda_j, \sum_{j=1}^{i} \lambda_j \delta_j\right) \quad \text{for } 1 \le i \le k.$$

Theorem 4.0.2. *Suppose $n > m$ and the orders of χ_i, ρ_j divide $p-1$. Then, for each $a \in \mathbb{G}_m(k)$, the Frobenius Newton polygon of* $\mathrm{Hyp}_{(n,m)}(\chi; \rho)(a)$ *coincides with the irregular Hodge polygon defined by* (4.0.0.2).

A "Newton above Hodge" type result for twisted exponential sums was obtained in [Adolphson and Sperber 1993]. In our case, we show that the (combinatorial) Hodge polygon in [loc. cit.] for hypergeometric sums coincides with the irregular Hodge polygon of hypergeometric connections. Then, we apply a result of Wan [1993] to conclude "Newton equals Hodge".

4.1. ***Frobenius Newton polygon above Hodge polygon.*** In this subsection, we revise Adolphson and Sperber's definition [1993] of (combinatorial) Hodge polygons and their result on "Newton above Hodge" for certain twisted exponential sums. Finally, we can identify their Hodge polygon with the irregular Hodge polygon of hypergeometric connections (Proposition 4.1.7).

4.1.1. Let N be a positive integer,

$$\chi = (\chi_1, \dots, \chi_N) : (k^\times)^N \to K^\times$$

a multiplicative character, and $g : \mathbb{G}_m^N \to \mathbb{A}^1$ a morphism defined by a Laurent polynomial

$$g(x_1, \dots, x_N) = \sum_{j=1}^{M} a_j x^{u_j} \in k[x_1^{\pm}, \dots, x_N^{\pm}],$$

where $\{u_j\}_{j=1}^M$ is a finite subset of $\mathbb{Z}^N$ and $a_j \in k^\times$. For $m \in \mathbb{N}$, we consider the twisted exponential sum

$$S_m(\chi, g) = \sum_{x \in \mathbb{G}_m^N(\mathbb{F}_{q^m})} \chi^{(m)}(x) \psi^{(m)}(g(x)), \tag{4.1.1.1}$$

where $\chi^{(m)} = \chi \circ \mathrm{Nm}_{\mathbb{F}_{q^m}/k} and\ \psi^{(m)} = \psi \circ \mathrm{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_p}$. The associated L-function

$$L(\chi, g; T) = \exp\left(\sum_{m \geq 1} S_m(\chi, g) \frac{T^m}{m}\right) \tag{4.1.1.2}$$

is a rational function in T by the Grothendieck–Lefschetz trace formula (or the Dwork trace formula).

Recall that we denote by $\Delta = \Delta(g)$ the convex closure in $\mathbb{R}^N$ generated by the origin and lattices defined by the exponents $\{u_j\}$ of g in Definition 2.2.1. Let $C(g)$ be the cone over Δ, i.e., the union of all rays in $\mathbb{R}^N$ emanating from the origin and passing through Δ.

We set $M(g) = C(g) \cap \mathbb{Z}^N$. Adolphson and Sperber [1989, (1.7)] considered a subring $R(g)$ of $k[x_1^{\pm}, \dots, x_N^{\pm}]$ defined by monomials with exponents in $M(g)$:

$$R(g) = k[x^{M(g)}].$$

We take $d_i \in [0, q-2]$ such that[4] $\chi_i = \omega^{-d_i}$. We set

$$\bar{d}_i = \begin{cases} q - 1 - d_i, & d_i \neq 0, \\ d_i, & d_i = 0, \end{cases}$$

and

$$\boldsymbol{d} = (d_1, \dots, d_N), \quad \bar{\boldsymbol{d}} = \{\bar{d}_1, \dots, \bar{d}_N\}, \quad N_{\boldsymbol{d}} = (q-1)^{-1}\boldsymbol{d} + \mathbb{Z}^N.$$

We define an $R(g)$-module $R_{\boldsymbol{d}}(g)$ [Adolphson and Sperber 1989, (1.12)] by

$$R_{\boldsymbol{d}}(g) = \left\{ \sum_{\text{finite}} b_u x^u \;\middle|\; u \in N_{\boldsymbol{d}} \cap C(g), b_u \in k \right\}.$$

[4] Adolphson–Sperber's convention $\chi_i = \omega^{-d_i}$ is different from our convention in the beginning of Section 4 by a minus sign.

There exists a (minimal) positive integer M such that, for all $u \in \left(\frac{1}{q-1}\mathbb{Z}\right)^{\mathbb{N}} \cap C(g)$, the weight function $w(u)$, defined as the minimal positive real number w such that $u \in w\Delta(g)$, is a rational number with denominator dividing M. Then w defines an increasing filtration on $R(g)$ by

$$R(g)_{i/M} = \left\{ \sum_{u \in M(g)} b_u x^u : w(u) \leq \frac{i}{M} \text{ for all } u \text{ with } b_u \neq 0 \right\}.$$

We denote the associated graded module by

$$\bar{R}(g) = \bigoplus_{i \geq 0} \bar{R}(g)_{i/M},$$
$$\bar{R}(g)_{i/M} = R(g)_{i/M} / R(g)_{(i-1)/M}.$$

Similarly, we equip $R_{\boldsymbol{d}}(g)$ with a filtration compatible with that of $R(g)$ and let $\bar{R}_{\boldsymbol{d}}(g)$ be the associated graded $\bar{R}(g)$-module.

4.1.2. In the following, we assume that g is *nondegenerate* and that $\dim \Delta(g) = N$.

For $1 \leq i \leq N$, let $\bar{g}_i$ be the image of $x_i\, \partial/\partial x_i\, g$ in $\bar{R}(g)_1$, and set

$$\bar{I}_{\boldsymbol{d}} = \bar{g}_1 \bar{R}(g)_{\boldsymbol{d}} + \cdots + \bar{g}_N \bar{R}(g)_{\boldsymbol{d}},$$

a graded submodule of $\bar{R}(g)_{\boldsymbol{d}}$. For each $i \geq 0$, we define a finite set

$$S_{\boldsymbol{d}}^{i/M} \subset N_{\boldsymbol{d}} \cap C(g)$$

of exponents as follows [Adolphson and Sperber 1991, §3]: Take a k-linearly independent set of monomials $\{x^\mu \mid \mu \in S_{\boldsymbol{d}}^{i/M}\}$ of weight i/M which spans a k-subspace $\bar{V}_{\boldsymbol{d},i/M}$ complement to $\bar{R}(g)_{\boldsymbol{d},i/M} \cap \bar{I}_{\boldsymbol{d}}$; i.e.,

$$\bar{R}(g)_{\boldsymbol{d},i/M} = \bar{V}_{\boldsymbol{d},i/M} \oplus (\bar{R}(g)_{\boldsymbol{d},i/M} \cap \bar{I}_{\boldsymbol{d},i/M}).$$

Set

$$S_{\boldsymbol{d}} = \bigcup_{i \geq 0} S_{\boldsymbol{d}}^{i/M},$$

which we also denote by $S_{\boldsymbol{d}}(g)$, and let $V(g)$ be the volume of $\Delta(g)$. The quotient $\bar{R}(g)_{\boldsymbol{d}}/\bar{I}_{\boldsymbol{d}}$ admits a basis of monomials in $S_{\boldsymbol{d}}$ and has dimension [loc. cit., Lemma 2.8]

$$\dim \bar{R}(g)_{\boldsymbol{d}}/\bar{I}_{\boldsymbol{d}} = N!V(g).$$

In this case, the L-function $L(\chi, g; T)^{(-1)^{N-1}}$ (4.1.1.2) is a polynomial of degree $N!V(g)$ [loc. cit., Corollary 2.12]. The q-order of roots of this polynomial are called *Frobenius slopes* of the twisted exponential sums $S_m(\chi, g)$.

Adolphson and Sperber studied the *Frobenius Newton polygon* defined by Frobenius slopes of this L-function (Definition 4.0.1) and compared it with a Hodge polygon defined as below.

For an integer $0 \leq d \leq q-2$, let d' be the nonnegative residue of pd modulo $q-1$. Recall that $q = p^s$ for an integer $s \geq 1$. For $\boldsymbol{d} = (d_1, \ldots, d_N)$, we set $\boldsymbol{d}' = (d_1', \ldots, d_N')$ and $\boldsymbol{d}^{(i)}$ the i-th composition of $(-)'$ on $\boldsymbol{d}$ for $i \geq 1$. Note that $\boldsymbol{d}^{(s)} = \boldsymbol{d}$.

We arrange elements of $S_{\boldsymbol{d}}=\{u_{\boldsymbol{d}}(1),\dots,u_{\boldsymbol{d}}(N!V(g))\}$ by $w(u_{\boldsymbol{d}}(1))\le\cdots\le w(u_{\boldsymbol{d}}(N!V(g)))$, and we repeat this ordering for $S_{\boldsymbol{d}'},\dots,S_{\boldsymbol{d}^{(s-1)}}$. For an integer $\ell\ge 0$, we set [loc. cit., Theorem 3.17]

$$W(\ell)=\operatorname{card}\left\{j\;\middle|\;\sum_{i=0}^{s-1}w(u_{\boldsymbol{d}^{(i)}}(j))=\frac{\ell}{M}\right\}.$$

When $\ell>sNM$, we have $W(\ell)=0$.

Definition 4.1.3 (Adolphson–Sperber). The Hodge polygon $\mathrm{HP}(\Delta(g)_{\boldsymbol{d}})$ is defined by the convex polygon in $\mathbb{R}^2$ with vertices $(0,0)$ and

$$\left(\sum_{\ell=0}^{m}W(\ell),\ \frac{1}{sM}\sum_{\ell=0}^{m}\ell W(\ell)\right),\qquad m=0,1,\dots,sNM.$$

Theorem 4.1.4 [Adolphson and Sperber 1993, Corollary 3.18]. *If g is nondegenerate and* $\dim(\Delta(g))=N$, *the Frobenius Newton polygon of $L(\chi,g;T)^{(-1)^{N-1}}$ lies above the Hodge polygon* $\mathrm{HP}(\Delta(g)_{\boldsymbol{d}})$, *and their endpoints coincide.*

Definition 4.1.5. We say that (g,χ) is *ordinary* if these two polygons coincide. When the character χ is trivial, we simply say g is *ordinary*.

4.1.6. In the following, we apply the above theory to the case of hypergeometric sums at the beginning of Section 4. We may assume that χ_1 is trivial (i.e., $\alpha_1=0$). Let a be an element of $k^\times$. We take $N=n+m-1$, $\boldsymbol{d}=(\bar a_2,\dots,\bar a_n,b_1,\dots,b_m)$, and g to be the nondegenerate function (2.2.1.1)

$$f_a=a\frac{y_1\cdots y_m}{x_2\cdots x_n}+x_2+\cdots+x_n-y_1-\cdots-y_m.$$

Then, we recover the hypergeometric sum (4.0.0.1) from (4.1.1.1).

Proposition 4.1.7. *If (χ,ρ) is nonresonant and the orders of the characters χ_i and ρ_j divide $p-1$, then the Hodge polygon* $\mathrm{HP}(\Delta(f_a)_{\boldsymbol{d}})$ *coincides with the irregular Hodge polygon defined by* (4.0.0.2) *associated to*

$$\left(0,\alpha_2=\frac{a_2}{p-1},\dots,\alpha_n=\frac{a_n}{p-1}\right),\quad\left(\beta_1=\frac{b_1}{p-1},\dots,\beta_m=\frac{b_m}{p-1}\right).$$

Proof. Since α_i and β_j have denominators dividing $p-1$, the numbers $\boldsymbol{d}^{(i)}$ are equal to $\boldsymbol{d}$ for every $i\ge 1$. In particular, the multiset of slopes of $\mathrm{HP}(\Delta(f_a)_{\boldsymbol{d}})$ coincides with $w(S_{\boldsymbol{d}})=\{\omega(u)\mid u\in S_{\boldsymbol{d}}\}$.

The cohomology classes $\omega_{r,\ell}=g_{r,\ell}\cdot\eta$ in Proposition 3.2.1 form a basis of the de Rham cohomology group $\mathrm{H}^{n+m-1}_{\mathrm{dR}}(U_a,f_a)^{(G,\tilde\chi\times\rho)}$. By the calculation of cohomology groups [loc. cit., §3, Theorem 3.14], the functions $\{g_{r,\ell}\}$ also form a basis of $\overline{V}_{\bar{\boldsymbol{d}}}$, with $\bar{\boldsymbol{d}}=(a_2,\dots,a_n,\bar b_1,\dots,\bar b_m)$. Hence

$$w(S_{\bar{\boldsymbol{d}}})=\{w(g_{r,\ell})\mid 0\le r\le m,\ 1\le\ell\le s_{r+1}-s_r\}.$$

By (3.1.1.1), Lemma 3.3.4 and the duality (3.3.6.4), the set of weights $w(S_{\boldsymbol{d}})$ coincides with the set of irregular Hodge numbers (4.0.0.2). Then, the proposition follows. □

4.2. ***Frobenius slopes of hypergeometric sums: proof of Theorem 4.0.2.*** We proceed by induction on n. Suppose the theorem holds when the rank of the hypergeometric F-isocrystal is less than n.

4.2.1. *Resonant case.* We first show that we can deduce the assertion in the resonant case from the induction hypothesis. We assume there exists i and j such that $\alpha_i = \beta_j$.

We slightly modify our convention on α and β by replacing those $\alpha_i, \beta_j = 0$ by 1 and then arranging them as $0 < \alpha_1 \le \alpha_2 \le \cdots \le \alpha_n \le 1$ and $0 < \beta_1 \le \cdots \le \beta_m \le 1$. Note that this modification does not change the multiset $\{\theta(1), \ldots, \theta(n)\}$ of irregular Hodge numbers. After twisting by a multiplicative character, we may assume that $\chi_n = \rho_m = 1$ are the trivial characters (i.e., $\alpha_n = \beta_m = 1$). Then we have the following identities:

$$\begin{aligned}\mathrm{Hyp}_{(n,m)}(\chi;\rho)(a) &= \sum_{x_i, y_j \in k^\times} \psi\Bigl(\sum_{i=1}^{n-1} x_i + a\frac{y_1\cdots y_m}{x_1\cdots x_{n-1}} - \sum_{j=1}^{m} y_j\Bigr)\cdot\prod_{i=1}^{n-1}\chi_i(x_i)\prod_{j=1}^{m-1}\rho_j^{-1}(y_j)\\
&= \sum_{x_i, y_j \in k^\times, y_m\in k} \psi\Bigl(\sum_{i=1}^{n-1} x_i - \sum_{j=1}^{m-1} y_j + y_m\Bigl(a\frac{y_1\cdots y_{m-1}}{x_1\cdots x_{n-1}} - 1\Bigr)\Bigr)\cdot\prod_{i=1}^{n-1}\chi_i(x_i)\prod_{j=1}^{m-1}\rho_j^{-1}(y_j)\\
&\qquad - \sum_{x_i, y_j \in k^\times} \psi\Bigl(\sum_{i=1}^{n-1} x_i - \sum_{j=1}^{m-1} y_j\Bigr)\cdot\prod_{i=1}^{n-1}\chi_i(x_i)\prod_{j=1}^{m-1}\rho_j^{-1}(y_j)\\
&= q\,\mathrm{Hyp}_{(n-1,m-1)}(\chi';\rho')(a) - \psi(-1)^{m-1}\prod_{i=1}^{n-1} G(\psi,\chi_i)\prod_{j=1}^{m-1} G(\psi,\rho_j^{-1}),\end{aligned} \tag{4.2.1.1}$$

where $\chi' = (\chi_1, \ldots, \chi_{n-1})$ and $\rho' = (\rho_1, \ldots, \rho_{m-1})$, and

$$G(\psi, \chi_i) = \sum_{x\in k^\times} \psi(x)\chi_i(x)$$

denotes the Gauss sum. In particular, the above sum can be written as a sum of n Frobenius eigenvalues by induction.

Let θ' be the function (4.0.0.2) defined by rational numbers $\alpha_1, \ldots, \alpha_{n-1}$ and $\beta_1, \ldots, \beta_{m-1}$. Then, we have

$$\theta(k) = \theta'(k) + 1 \quad \text{for all } 1 \le k \le n-1$$

and

$$\theta(n) = \sum_{i=1}^{n}(1-\alpha_i) + \sum_{\beta_j<1}\beta_j = \mathrm{ord}_q\Bigl(\prod_{i=1}^{n-1} G(\psi,\chi_i)\prod_{j=1}^{m-1} G(\psi,\rho_j^{-1})\Bigr),$$

where the second identity follows from Stickelberger's theorem, saying that

$$\mathrm{ord}_q\, G(\psi, \omega^{-k}) = \frac{k}{p-1}.$$

Then, the theorem in the resonant case follows from the induction hypothesis and decomposition (4.2.1.1).

4.2.2. *Nonresonant case.* By the previous argument, we may assume that the assertion for the hypergeometric sum of type (n, m) defined by a resonant pair (α, β) is already proved. It suffices to treat the nonresonant case. We may assume $\chi_1 = 1$ is trivial.

We set $\tilde f_a(x_2, \ldots, x_n, y_1, \ldots, y_m) = f_a(x_2^{p-1}, \ldots, x_n^{p-1}, y_1^{p-1}, \ldots, y_m^{p-1})$. We first prove the ordinariness of exponential sums associated to $\tilde f_a$ (Definition 4.1.5) using a theorem of Wan [1993].

Let $\delta_1, \ldots, \delta_{m+n}$ be all the facets of $\Delta = \Delta(\tilde{f}_a)$ which do not contain the origin. Let $\tilde{f}_a^{\delta_i}$ be the restriction of $\tilde{f}_a$ to δ_i [Wan 2004, §1.1], which is also nondegenerate [Wan 2004, §3.1]. By [Wan 2004, Theorem 3.1], $\tilde{f}_a$ is ordinary if and only if each $\tilde{f}_a^{\delta_i}$ is ordinary.

Each Laurent polynomial $\tilde{f}_a^{\delta_i}$ is diagonal; that is, $\tilde{f}_a^{\delta_i}$ has exactly $n+m-1$ nonconstant terms of monomials and $\Delta(\tilde{f}_a^{\delta_i})$ is $(n+m-1)$-dimensional [Wan 2004, §2]. Indeed, if $V_1, \ldots, V_{m+n-1}$ denote the vertex of δ_i written as column vectors, the set $S(\delta_i)$ of solutions of

$$(V_1, \ldots, V_{m+n-1}) \begin{pmatrix} r_1 \\ \vdots \\ r_{m+n-1} \end{pmatrix} \equiv 0 \pmod 1, \quad r_i \text{ rational}, \quad 0 \le r_i < 1,$$

forms an abelian group, which is isomorphic to $(\mathbb{Z}/(p-1)\mathbb{Z})^{n+m-1}$. We deduce that, for each δ_i, $\tilde{f}_a^{\delta_i}$ is ordinary by [Wan 2004, Corollary 2.6].

We have a decomposition of exponential sums as follows:

$$\sum_{x_i, y_j \in k^\times} \psi(\tilde{f}_a(x_i, y_j)) = \sum_{\chi_i, \rho_j} \mathrm{Hyp}_{(n,m)}(\chi, \rho)(a), \tag{4.2.2.1}$$

where the sum is taken over all multiplicative characters χ_i and ρ_j with $2 \le i \le n$ and $1 \le j \le m$ of orders dividing $p-1$. We have a similar decomposition for $S_{\boldsymbol{d}}$ (Section 4.1.2) given by

$$S_1(\tilde{f}_a) = \bigsqcup_{\boldsymbol{d}} S_{\boldsymbol{d}}(f_a),$$

where $1 = (0, 0, \ldots, 0)$ and $\boldsymbol{d}$ is taken over all $(n+m-1)$-tuples of rational numbers with denominators $p-1$ in $[0, 1)$.

On the left-hand side of (4.2.2.1), we have shown “Newton equals Hodge” (i.e., the ordinariness of $\tilde{f}_a$). Together with the “Newton above Hodge” for each hypergeometric sum (Theorem 4.1.4), we deduce that “Newton equals Hodge” for each component of the right-hand side. Then, the assertion in the nonresonant case follows from Proposition 4.1.7. □

In particular, our proof shows Proposition 4.1.7 in the resonant case.

Corollary 4.2.3. *Proposition 4.1.7 holds without the nonresonant assumption.*

Proof. In the resonant case, the Frobenius Newton polygon equals the irregular Hodge polygon by Section 4.2.1. By the proof in Section 4.2.2, the Frobenius Newton polygon equals the (combinatorial) Hodge polygon defined by Adolphson–Sperber. Then, the assertion follows. □

Acknowledgements

The authors thank Alberto Castaño Domínguez, Javier Fresán, Lei Fu, Shun Ohkubo, Claude Sabbah, Christian Sevenheck, Daqing Wan, and Jeng-Daw Yu for their valuable discussions. We are also grateful to an anonymous referee for careful reading and valuable comments.

Qin acknowledges the financial support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 101020009) for part of this work. Xu acknowledges financial support from the National Natural Science Foundation of China (grant nos. 12222118 and 12288201) and the CAS Project for Young Scientists in Basic Research, grant no. YSBR-033. Part of the work was done when Qin was staying at Morningside Center of Mathematics, and he would like to thank Morningside Center of Mathematics for its hospitality.

References

[Adolphson and Sperber 1987] A. Adolphson and S. Sperber, "Twisted Kloosterman sums and p-adic Bessel functions, II: Newton polygons and analytic continuation", *Amer. J. Math.* **109**:4 (1987), 723–764. MR

[Adolphson and Sperber 1989] A. Adolphson and S. Sperber, "Exponential sums and Newton polyhedra: cohomology and estimates", *Ann. of Math.* (2) **130**:2 (1989), 367–406. MR

[Adolphson and Sperber 1991] A. Adolphson and S. Sperber, "On twisted exponential sums", *Math. Ann.* **290**:4 (1991), 713–726. MR

[Adolphson and Sperber 1993] A. Adolphson and S. Sperber, "Twisted exponential sums and Newton polyhedra", *J. Reine Angew. Math.* **443** (1993), 151–177. MR

[Adolphson and Sperber 1997] A. Adolphson and S. Sperber, "On twisted de Rham cohomology", *Nagoya Math. J.* **146** (1997), 55–81. MR

[Beilinson and Drinfeld 1997] A. Beilinson and V. Drinfeld, "Quantization of Hitchin's integrable system and Hecke eigensheaves", preprint, 1997, available at https://math.uchicago.edu/~drinfeld/langlands/QuantizationHitchin.pdf.

[Berthelot 1984] P. Berthelot, "Cohomologie rigide et théorie de Dwork: le cas des sommes exponentielles", pp. 17–49 in *p-adic cohomology*, Astérisque **119-120**, Soc. Math. France, Paris, 1984. MR

[Berthelot and Ogus 1978] P. Berthelot and A. Ogus, *Notes on crystalline cohomology*, Princeton Univ. Press, 1978. MR

[Beukers and Heckman 1989] F. Beukers and G. Heckman, "Monodromy for the hypergeometric function ${}_nF_{n-1}$", *Invent. Math.* **95**:2 (1989), 325–354. MR

[Castaño Domínguez and Sevenheck 2021] A. Castaño Domínguez and C. Sevenheck, "Irregular Hodge filtration of some confluent hypergeometric systems", *J. Inst. Math. Jussieu* **20**:2 (2021), 627–668. MR

[Castaño Domínguez et al. 2019] A. Castaño Domínguez, T. Reichelt, and C. Sevenheck, "Examples of hypergeometric twistor $\mathscr{D}$-modules", *Algebra Number Theory* **13**:6 (2019), 1415–1442. MR

[Deligne 2007a] P. Deligne, "Théorie de Hodge irrégulière (version originelle)", pp. 109–114 in *Singularités irrégulières: correspondance et documents*, Doc. Math. (Paris) **5**, Soc. Math. France, Paris, 2007.

[Deligne 2007b] P. Deligne, "Théorie de Hodge irrégulière (août 2006)", pp. 115–128 in *Singularités irrégulières: correspondance et documents*, Doc. Math. (Paris) **5**, Soc. Math. France, Paris, 2007.

[Drinfeld and Kedlaya 2017] V. Drinfeld and K. S. Kedlaya, "Slopes of indecomposable F-isocrystals", *Pure Appl. Math. Q.* **13**:1 (2017), 131–192. MR

[Dwork 1974] B. Dwork, "Bessel functions as p-adic functions of the argument", *Duke Math. J.* **41** (1974), 711–738. MR

[Esnault et al. 2017] H. Esnault, C. Sabbah, and J.-D. Yu, "E_1-degeneration of the irregular Hodge filtration", *J. Reine Angew. Math.* **729** (2017), 171–227. MR

[Fedorov 2018] R. Fedorov, "Variations of Hodge structures for hypergeometric differential operators and parabolic Higgs bundles", *Int. Math. Res. Not.* **2018**:18 (2018), 5583–5608. MR

[Fresán et al. 2022] J. Fresán, C. Sabbah, and J.-D. Yu, "Hodge theory of Kloosterman connections", *Duke Math. J.* **171**:8 (2022), 1649–1747. MR

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Ann. of Math. Stud. **131**, Princeton Univ. Press, 1993. MR

[Hotta et al. 2008] R. Hotta, K. Takeuchi, and T. Tanisaki, *D-modules, perverse sheaves, and representation theory*, Progr. Math. **236**, Birkhäuser, Boston, MA, 2008. MR

[Kamgarpour and Yi 2021] M. Kamgarpour and L. Yi, "Geometric Langlands for hypergeometric sheaves", *Trans. Amer. Math. Soc.* **374**:12 (2021), 8435–8481. MR

[Kamgarpour et al. 2023] M. Kamgarpour, D. Xu, and L. Yi, "Hypergeometric sheaves for classical groups via geometric Langlands", *Trans. Amer. Math. Soc.* **376**:5 (2023), 3585–3640. MR

[Katz 1990] N. M. Katz, *Exponential sums and differential equations*, Ann. of Math. Stud. **124**, Princeton Univ. Press, 1990. MR

[Kontsevich and Soibelman 2011] M. Kontsevich and Y. Soibelman, "Cohomological Hall algebra, exponential Hodge structures and motivic Donaldson–Thomas invariants", *Commun. Number Theory Phys.* **5**:2 (2011), 231–352. MR

[Martin 2021] N. Martin, "Middle multiplicative convolution and hypergeometric equations", *J. Singul.* **23** (2021), 194–204. MR

[Mazur 1972] B. Mazur, "Frobenius and the Hodge filtration", *Bull. Amer. Math. Soc.* **78** (1972), 653–667. MR

[Miyatani 2020] K. Miyatani, "p-adic generalized hypergeometric equations from the viewpoint of arithmetic $\mathscr{D}$-modules", *Amer. J. Math.* **142**:4 (2020), 1017–1050. MR

[Mochizuki 2025] T. Mochizuki, "Rescalability of integrable mixed twistor D-modules", pp. 13–207 in *Perspectives on four decades of algebraic geometry*, *II*, edited by A. Albano et al., Progr. Math. **352**, Springer, 2025.

[Qin 2024] Y. Qin, "Hodge numbers of motives attached to Kloosterman and Airy moments", *J. Reine Angew. Math.* **808** (2024), 143–192. MR

[Sabbah 2010] C. Sabbah, "Fourier–Laplace transform of a variation of polarized complex Hodge structure, II", pp. 289–347 in *New developments in algebraic geometry, integrable systems and mirror symmetry* (Kyoto, 2008), edited by M.-H. Saito et al., Adv. Stud. Pure Math. **59**, Math. Soc. Japan, Tokyo, 2010. MR

[Sabbah 2018] C. Sabbah, *Irregular Hodge theory*, Mém. Soc. Math. Fr. (N.S.) **156**, Soc. Math. France, Paris, 2018. MR

[Sabbah and Yu 2015] C. Sabbah and J.-D. Yu, "On the irregular Hodge filtration of exponentially twisted mixed Hodge modules", *Forum Math. Sigma* **3** (2015), e9. MR

[Sabbah and Yu 2019] C. Sabbah and J.-D. Yu, "Irregular Hodge numbers of confluent hypergeometric differential equations", *Épijournal Géom. Algébrique* **3** (2019), art. id. 7. MR

[Sabbah and Yu 2023] C. Sabbah and J.-D. Yu, "Hodge properties of Airy moments", *Tunis. J. Math.* **5**:2 (2023), 215–271. MR

[Saito 1990] M. Saito, "Mixed Hodge modules", *Publ. Res. Inst. Math. Sci.* **26**:2 (1990), 221–333. MR

[Simpson 1990] C. T. Simpson, "Harmonic bundles on noncompact curves", *J. Amer. Math. Soc.* **3**:3 (1990), 713–770. MR

[Sperber 1977] S. Sperber, "p-adic hypergeometric functions and their cohomology", *Duke Math. J.* **44**:3 (1977), 535–589. MR

[Wan 1993] D. Q. Wan, "Newton polygons of zeta functions and L functions", *Ann. of Math.* (2) **137**:2 (1993), 249–293. MR

[Wan 2004] D. Wan, "Variation of p-adic Newton polygons for L-functions of exponential sums", *Asian J. Math.* **8**:3 (2004), 427–471. MR

[Yu 2014] J.-D. Yu, "Irregular Hodge filtration on twisted de Rham cohomology", *Manuscripta Math.* **144**:1-2 (2014), 99–133. MR

[Zhu 2017] X. Zhu, "Frenkel–Gross' irregular connection and Heinloth–Ngô–Yun's are the same", *Selecta Math.* (*N.S.*) **23**:1 (2017), 245–274. MR

Communicated by Hélène Esnault
Received 2024-05-10 Revised 2024-10-14 Accepted 2024-12-05

yichenqin@fudan.edu.cn — *School of Mathematical Sciences, Fudan University, Shanghai, China*

daxin.xu@amss.ac.cn — *Morningside Center of Mathematics and State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China*

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 19 No. 12 2025