Pacific Journal of Mathematics

Volume 271 No. 2 October 2014

PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

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

The subscription price for 2014 is US $410/year for the electronic version, and $535/year for print and electronic.
Subscriptions, requests for back issues and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.257

MONOIDS OF MODULES AND ARITHMETIC OF DIRECT-SUM DECOMPOSITIONS

NICHOLAS R. BAETH AND ALFRED GEROLDINGER

Let R be a (possibly noncommutative) ring and let $\mathcal{C}$ be a class of finitely generated (right) R-modules which is closed under finite direct sums, direct summands, and isomorphisms. Then the set $\mathcal{V}(\mathcal{C})$ of isomorphism classes of modules is a commutative semigroup with operation induced by the direct sum. This semigroup encodes all possible information about direct sum decompositions of modules in $\mathcal{C}$. If the endomorphism ring of each module in $\mathcal{C}$ is semilocal, then $\mathcal{V}(\mathcal{C})$ is a Krull monoid. Although this fact was observed nearly a decade ago, the focus of study thus far has been on ring- and module-theoretic conditions enforcing that $\mathcal{V}(\mathcal{C})$ is Krull. If $\mathcal{V}(\mathcal{C})$ is Krull, its arithmetic depends only on the class group of $\mathcal{V}(\mathcal{C})$ and the set of classes containing prime divisors. In this paper we provide the first systematic treatment to study the direct-sum decompositions of modules using methods from factorization theory of Krull monoids. We do this when $\mathcal{C}$ is the class of finitely generated torsion-free modules over certain one- and two-dimensional commutative Noetherian local rings.

1. Introduction

The study of direct-sum decompositions of finitely generated modules is a classical topic in module theory dating back over a century. In the early 1900s, Wedderburn, Remak, Krull, and Schmidt proved unique direct-sum decomposition results for various classes of groups (see [Maclagan-Wedderburn 1909; Remak 1911; Krull 1925; Schmidt 1929]). A few decades later Azumaya [1950] proved uniqueness of (possibly infinite) direct-sum decomposition of modules provided that each indecomposable module has a local endomorphism ring. In the commutative setting, Evans [1973] gave an example due to Swan illustrating a nonunique direct-sum decomposition of a finitely generated module over a local ring. The past decade has seen a new semigroup-theoretical approach. This approach was first introduced by

Baeth was a Fulbright-NAWI Graz Visiting Professor in the Natural Sciences and supported by the Austrian-American Education Commission. Geroldinger was supported by the Austrian Science Fund FWF, Project Number P21576-N18.
MSC2010: 13C14, 16D70, 20M13.
Keywords: Krull monoids, sets of lengths, direct-sum decompositions, indecomposable modules.

Facchini and Wiegand [2004] and has been used by several authors (for example, see [Baeth 2007; 2009; Baeth and Luckas 2011; Baeth and Saccon 2012; Diracca 2007; Facchini 2002; 2006; 2012; Facchini and Halter-Koch 2003; Facchini et al. 2006; Facchini and Wiegand 2004; Hassler et al. 2007; Herbera and Příhoda 2010; Levy and Odenthal 1996]). Let R be a ring and let $\mathcal{C}$ be a class of right R-modules which is closed under finite direct sums, direct summands, and isomorphisms. For a module M in $\mathcal{C}$, let $[M]$ denote the isomorphism class of M. Let $\mathcal{V}(C)$ denote the set of isomorphism classes of modules in $\mathcal{C}$. (We assume here that $\mathcal{V}(C)$ is indeed a set, and note that this hypothesis holds for all examples we study.) Then $\mathcal{V}(\mathcal{C})$ is a commutative semigroup with operation defined by $[M]+[N]=[M \oplus N]$ and all information about direct-sum decomposition of modules in $\mathcal{C}$ can be studied in terms of factorization of elements in the semigroup $\mathcal{V}(\mathcal{C})$. In particular, the direct-sum decompositions in $\mathcal{C}$ are (essentially) unique (in other words, the Krull–Remak–Schmidt–Azumaya theorem — KRSA — holds) if and only if $\mathcal{V}(C)$ is a free abelian monoid. This semigroup-theoretical point of view was justified by Facchini [2002] who showed that $\mathcal{V}(\mathcal{C})$ is a reduced Krull monoid provided that the endomorphism ring $\operatorname{End}_R(M)$ is semilocal for all modules M in $\mathcal{C}$. This result allows one to describe the direct-sum decomposition of modules in terms of factorization of elements in Krull monoids, a well-studied class of commutative monoids.

However, thus far much of the focus in this direction has been on the study of module-theoretic conditions which guarantee that all endomorphism rings are semilocal, as well as on trying to describe the monoid $\mathcal{V}(\mathcal{C})$ in terms of various ring- and module-theoretic conditions. Although some factorization-theoretic computations have been done in various settings (e.g., the study of elasticity in [Baeth 2009; Baeth and Luckas 2011; Baeth and Saccon 2012] and the study of the ω-invariant in [Diracca 2007]), the general emphasis has not been on the arithmetic of the monoid $\mathcal{V}(\mathcal{C})$. Our intent is to use known module-theoretic results along with factorization-theoretic techniques in order to give detailed descriptions of the arithmetic of direct-sum decompositions of finitely generated torsion-free modules over certain one- and two-dimensional local rings. We hope that this systematic approach will not only serve to inspire others to consider more detailed and abstract factorization-theoretic approaches to the study of direct-sum decompositions, but to provide new and interesting examples for zero-sum theory over torsion-free groups. We refer to [Facchini 2003] and to the opening paragraph in the recent monograph [Leuschke and Wiegand 2012] for broad information on the Krull–Remak–Schmidt–Azumaya theorem, and to the surveys [Facchini 2012; Baeth and Wiegand 2013] promoting this semigroup-theoretical point of view. More details and references will be given in Section 3.

Krull monoids, both their ideal theory and their arithmetic, are well-studied; see [Geroldinger and Halter-Koch 2006] for a thorough treatment. A reduced

Krull monoid is uniquely determined (up to isomorphism) by its class group G, the set of classes $G_{\mathcal{P}} \subset G$ containing prime divisors, and the number of prime divisors in each class. Let $\mathcal{V}(C)$ be a monoid of modules and suppose $\mathcal{V}(\mathcal{C})$ is Krull with class group G and with set of classes containing prime divisors $G_{\mathcal{P}}$. We are interested in determining what this information tells us about direct-sum decompositions of modules. Let M be a module in $\mathcal{C}$ and let $M = M_1 \oplus \cdots \oplus M_\ell$ where $M_1, \ldots, M_\ell$ are indecomposable right R-modules. Then ℓ is called the length of this factorization (decomposition into indecomposables), and the set of lengths $\mathsf{L}(M) \subset \mathbb{N}$ is defined as the set of all possible factorization lengths. Then KRSA holds if and only if $|G| = 1$. Moreover, it is easy to check that $|\mathsf{L}(M)| = 1$ for all M in $\mathcal{C}$ provided that $|G| \le 2$. Clearly, sets of lengths are a measure how badly KRSA fails. Assuming that $\mathcal{V}(\mathcal{C})$ is Krull, M has at least one direct-sum decomposition in terms of indecomposable right R-modules, and, up to isomorphism, only finitely many distinct decompositions. In particular, all sets of lengths are finite and nonempty. Without further information about the class group G and the subset $G_{\mathcal{P}} \subset G$, this is all that can be said. Indeed, there is a standing conjecture that for every infinite abelian group G there is a Krull monoid with class group G and set $G_{\mathcal{P}}$ such that every set of lengths has cardinality one (see [Geroldinger and Göbel 2003]). On the other hand, if the class group of a Krull monoid is infinite and every class contains a prime divisor, then every finite subset of $\mathbb{N}_{\ge 2}$ occurs as a set of lengths (see Proposition 6.2).

Thus an indispensable prerequisite for the study of sets of lengths (and other arithmetical invariants) in Krull monoids is detailed information about not only the class group G, but also on the set $G_{\mathcal{P}} \subset G$ of classes containing prime divisors. For the monoid $\mathcal{V}(\mathcal{C})$, this is of course a module-theoretic task which depends on both the ring R and the class $\mathcal{C}$ of R-modules. Early results gave only extremal sets $G_{\mathcal{P}}$ and thus no further arithmetical investigations were needed. In Sections 4 and 5 we determine, based on deep module-theoretic results, the class group G of $\mathcal{V}(\mathcal{C})$. We then exhibit well-structured sets $G_{\mathcal{P}}$ providing a plethora of arithmetically interesting direct-sum decompositions. In particular, we study the classes of finitely generated modules, finitely generated torsion-free modules, and maximal Cohen–Macaulay modules over one- and two-dimensional commutative Noetherian local rings. We restrict, if necessary, to specific families of rings in order to obtain explicit results for $G_{\mathcal{P}}$, since it is possible that even slightly different sets $G_{\mathcal{P}}$ can induce completely different behavior in terms of the sets of lengths. Given this information, we use transfer homomorphisms, a key tool in factorization theory and introduced in Section 3, which make it possible to study sets of lengths and other arithmetical invariants of general Krull monoids instead in an associated monoid of zero-sum sequences (see Lemma 3.4). These monoids can be studied using methods from additive (group and number) theory (see [Geroldinger 2009]).

Factorization theory describes the nonuniqueness of factorizations of elements in rings and semigroups into irreducible elements by arithmetical invariants such as sets of lengths, catenary, and tame degrees. We will define each of these invariants in Section 2. The goal is to relate the arithmetical invariants with algebraic parameters (such as class groups) of the objects under consideration. The study of sets of lengths in Krull monoids is a central topic in factorization theory. However, since much of this theory was motivated by examples in number theory (such as holomorphy rings in global fields), most of the focus so far has been on Krull monoids with finite class group and with each class containing a prime divisor. This is in contrast to Krull monoids stemming from module theory which often have infinite class group (see Section 4). A key result in Section 6 shows that the arithmetic of these two types of Krull monoids can have drastically different arithmetic.

In combination with the study of various arithmetical invariants of a given Krull monoid, the following dual question has been asked since the beginning of factorization theory: Are arithmetical phenomena characteristic for a given Krull monoid (inside a given class of Krull monoids)? Affirmative answers have been given for the class of Krull monoids with finitely generated class groups where every class contains a prime divisor. Since sets of lengths are the most investigated invariants in factorization theory, the emphasis in the last decade has been on the following question: Within the class of Krull monoids having finite class group and such that every class contains a prime divisor, does the system of sets of lengths of a monoid H characterize the class group of H? A survey of these problems can be found in [Geroldinger and Halter-Koch 2006, Sections 7.1 and 7.2]. For recent progress, see [Schmid 2009b; 2009a; Baginski et al. 2013]. In Theorem 6.8 we exhibit that for many Krull monoids stemming from the module theory of Sections 4 and 5, the system of sets of lengths and the behavior of absolutely irreducible elements characterizes the class group of these monoids.

In Section 2 we introduce some of the main arithmetical invariants studied in factorization theory as well as their relevance to the study of direct-sum decompositions. Our focus is on sets of lengths and on parameters controlling their structure, but we will also need other invariants such as catenary and tame degrees. Section 3 gives a brief introduction to Krull monoids, monoids of modules, and transfer homomorphisms. Sections 4 and 5 provide explicit constructions stemming from module theory of class groups and distribution of prime divisors in the classes. Finally, in Section 6, we present our results on the arithmetic of direct-sum decomposition in the Krull monoids discussed in Sections 4 and 5.

We use standard notation from commutative algebra and module theory (see [Leuschke and Wiegand 2012]) and we follow the notation of [Geroldinger and Halter-Koch 2006] for factorization theory. All monoids of modules $\mathcal{V}(\mathcal{C})$ are written additively, while all abstract Krull monoids are written multiplicatively.

This follows the tradition in factorization theory, and makes sense also here because our crucial tool, the monoid of zero-sum sequences, is written multiplicatively. In particular, our arithmetical results in Section 6 are written in a multiplicative setting but they are derived for the additive monoids of modules discussed in Sections 4 and 5. Since we hope that this article is readable both for those working in ring and module theory as well as those working in additive theory and factorization theory, we often recall concepts of both areas which are well known to the specialists in the respective fields.

2. Arithmetical preliminaries

In this section we gather together the concepts central to describing the arithmetic of nonfactorial monoids. In particular, we exhibit the arithmetical invariants which will be studied in Section 6 and which will give a measure of nonunique direct sum decompositions of classes of modules studied in Sections 4 and 5. When possible, we recall previous work in the area of direct-sum decompositions for which certain invariants have been studied. For more details on nonunique factorization, see [Geroldinger and Halter-Koch 2006]. First we record some preliminary terminology.

Notation. We denote by $\mathbb{N}$ the set of positive integers and set $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For every $n \in \mathbb{N}$, C_n denotes a cyclic group of order n. For real numbers $a, b \in \mathbb{R}$ we set $[a, b] = \{x \in \mathbb{Z} : a \le x \le b\}$. We use the convention that $\sup \varnothing = \max \varnothing = \min \varnothing = 0$.

Subsets of the integers. Let $L, L' \subset \mathbb{Z}$. We denote by $L + L' = \{a + b : a \in L, b \in L'\}$ the *sumset* of L and L'. If $\varnothing \ne L \subset \mathbb{N}$, we call

$$\rho(L) = \sup\left\{\frac{m}{n} : m, n \in L\right\} = \frac{\sup L}{\min L} \in \mathbb{Q}_{\ge 1} \cup \{\infty\}$$

the *elasticity* of L. In addition, we set $\rho(\{0\}) = 1$. Distinct elements $k, l \in L$ are called *adjacent* if $L \cap [\min\{k, l\}, \max\{k, l\}] = \{k, l\}$. A positive integer $d \in \mathbb{N}$ is called a *distance* of L if there exist adjacent elements $k, l \in L$ with $d = |k - l|$. We denote by $\Delta(L)$ the *set of distances* of L. Note that $\Delta(L) = \varnothing$ if and only if $|L| \le 1$, and that L is an arithmetical progression with difference $d \in \mathbb{N}$ if and only if $\Delta(L) \subset \{d\}$.

Monoids and rings. By a *monoid* H we always mean a commutative semigroup with identity 1 which satisfies the cancellation law; that is, if a, b, and c are elements of the H with $ab = ac$, then $b = c$.

Let H be a monoid. We denote by $\mathcal{A}(H)$ the set of atoms (irreducible elements) of H, by $\mathsf{q}(H)$ a quotient group of H with $H \subset \mathsf{q}(H) = \{a^{-1}b : a, b \in H\}$, and by $H^\times$ the set of invertible elements of H. We say that H is *reduced* if $H^\times = \{1\}$, and we denote by $H_{\mathrm{red}} = H/H^\times = \{aH^\times : a \in H\}$ the associated reduced monoid.

Let $H' \subset H$ be a subset. We say that H' is *divisor-closed* if $a \in H'$ and $b \in H$ with $b \mid a$ implies that $b \in H'$. Denote by $[H'] \subset H$ the submonoid generated by H'.

A monoid F is called *free abelian with basis* $\mathcal{P} \subset F$ if every $a \in F$ has a unique representation of the form

$$a = \prod_{p \in \mathcal{P}} p^{\mathsf{v}_p(a)} \quad \text{with } \mathsf{v}_p(a) \in \mathbb{N}_0 \text{ and } \mathsf{v}_p(a) = 0 \text{ for almost all } p \in \mathcal{P}.$$

If F is free abelian with basis $\mathcal{P}$, we set $F = \mathcal{F}(\mathcal{P})$ and call

$$|a| = \sum_{p \in \mathcal{P}} \mathsf{v}_p(a)$$

the *length* of a, and

$$\operatorname{supp}(a) = \{p \in \mathcal{P} : \mathsf{v}_p(a) > 0\}$$

the *support* of a. The multiplicative monoid $\mathcal{F}(\mathcal{P})$ is, of course, isomorphic to the additive monoid $(\mathbb{N}_0^{(\mathcal{P})}, +)$.

Throughout this manuscript, all rings have a unit element and, apart from a few motivating remarks in Section 3, all rings are commutative. Let R be a ring. Then we let $R^\bullet = R \setminus \{0\}$ denote the nonzero elements of R and let $R^\times$ denote its group of units. Note that if R is a domain, then $R^\bullet$ is a monoid as defined above. By the dimension of a ring we always mean its Krull dimension.

Abelian groups. Let G be an additive abelian group and let $G_0 \subset G$ a subset. Then $-G_0 = \{-g : g \in G_0\}$, $G_0^\bullet = G_0 \setminus \{0\}$, and $\langle G_0 \rangle \subset G$ denotes the subgroup generated by G_0. A family $(e_i)_{i \in I}$ of elements of G is said to be *independent* if $e_i \neq 0$ for all $i \in I$ and, for every family $(m_i)_{i \in I} \in \mathbb{Z}^{(I)}$,

$$\sum_{i \in I} m_i e_i = 0 \text{ implies } m_i e_i = 0 \quad \text{for all } i \in I.$$

The family $(e_i)_{i \in I}$ is called a *basis* for G if $G = \bigoplus_{i \in I} \langle e_i \rangle$. The *total rank* $\mathsf{r}^*(G)$ is the supremum of the cardinalities of independent subsets of G. Thus $\mathsf{r}^*(G) = \mathsf{r}_0(G) + \sum_{p \in \mathbb{P}} \mathsf{r}_p(G)$, where $\mathsf{r}_0(G)$ is the torsion-free rank of G and $\mathsf{r}_p(G)$ is the p-rank of G for every prime $p \in \mathbb{P}$.

Factorizations. Let H be a monoid. The free abelian monoid $\mathsf{Z}(H) = \mathcal{F}(\mathcal{A}(H_{\mathrm{red}}))$ is called the *factorization monoid* of H, and the unique homomorphism

$$\pi : \mathsf{Z}(H) \to H_{\mathrm{red}} \text{ satisfying } \pi(u) = u \text{ for each } u \in \mathcal{A}(H_{\mathrm{red}})$$

is called the *factorization homomorphism* of H. For $a \in H$ and $k \in \mathbb{N}$,

- $\mathsf{Z}_H(a) = \mathsf{Z}(a) = \pi^{-1}(aH^\times) \subset \mathsf{Z}(H)$ is the *set of factorizations* of a,
- $\mathsf{Z}_k(a) = \{z \in \mathsf{Z}(a) : |z| = k\}$ is the *set of factorizations* of a of length k,

- $\mathsf{L}_H(a) = \mathsf{L}(a) = \{|z| : z \in \mathsf{Z}(a)\} \subset \mathbb{N}_0$ is the *set of lengths* of a, and
- $\mathcal{L}(H) = \{\mathsf{L}(b) : b \in H\}$ is the *system of sets of lengths* of H.

By definition, we have $\mathsf{Z}(a) = \{1\}$ and $\mathsf{L}(a) = \{0\}$ for all $a \in H^\times$. If H is assumed to be Krull, as is the case in the monoids of modules we study, and $a \in H$, then the set of factorizations $\mathsf{Z}(a)$ is finite and nonempty and hence $\mathsf{L}(a)$ is finite and nonempty. Suppose that there is $a \in H$ with $|\mathsf{L}(a)| > 1$ with distinct $k, l \in \mathsf{L}(a)$. Then for all $N \in \mathbb{N}$, $\mathsf{L}(a^N) \supset \{(N-i)k + il : i \in [0, N]\}$ and hence $|\mathsf{L}(a^N)| > N$. Thus, whenever there is an element $a \in H$ that has at least two factorizations of distinct lengths, there exist elements of H having arbitrarily many factorizations of distinct lengths. This motivates the need for more refined measures of nonunique factorization.

Several invariants such as elasticity and the Δ-set measure nonuniqueness in terms of sets of lengths. Other invariants such as the catenary degree provide an even more subtle measurement in terms of the distinct factorizations of elements. However, these two approaches cannot easily be separated and it is often the case that a factorization-theoretical invariant is closely related to an invariant of the set of lengths. Thus the exposition that follows will introduce invariants as they are needed and so that the relations between these invariants can be made as clear as possible.

The monoid H is called

- *atomic* if $\mathsf{Z}(a) \neq \varnothing$ for all $a \in H$,
- *factorial* if $|\mathsf{Z}(a)| = 1\,k$ for all $a \in H$ (equivalently, H_{red} is free abelian), and
- *half-factorial* if $|\mathsf{L}(a)| = 1$ for all $a \in H$.

Let $z,\ z' \in \mathsf{Z}(H)$. Then we can write

$$z = u_1 \cdot \ldots \cdot u_l v_1 \cdot \ldots \cdot v_m \quad \text{and} \quad z' = u_1 \cdot \ldots \cdot u_l w_1 \cdot \ldots \cdot w_n,$$

where $l, m, n \in \mathbb{N}_0$ and $u_1, \ldots, u_l,\ v_1, \ldots, v_m,\ w_1, \ldots, w_n \in \mathcal{A}(H_{\mathrm{red}})$ satisfy

$$\{v_1, \ldots, v_m\} \cap \{w_1, \ldots, w_n\} = \varnothing.$$

Then $\gcd(z, z') = u_1 \cdot \ldots \cdot u_l$, and we call

$$\mathsf{d}(z, z') = \max\{m,\ n\} = \max\{|z \gcd(z, z')^{-1}|, |z' \gcd(z, z')^{-1}|\} \in \mathbb{N}_0$$

the *distance* between z and z'. If $\pi(z) = \pi(z')$ and $z \neq z'$, then clearly

$$2 + \big||z| - |z'|\big| \le \mathsf{d}(z, z').$$

For subsets $X, Y \subset \mathsf{Z}(H)$, we set

$$\mathsf{d}(X, Y) = \min\{\mathsf{d}(x, y) : x \in X,\ y \in Y\},$$

and thus $\mathsf{d}(X, Y) = 0$ if and only if ($X \cap Y \neq \varnothing$, $X = \varnothing$, or $Y = \varnothing$).

From this point on, we will assume all monoids to be atomic. Since the monoids described in Sections 4 and 5 are of the form $\mathcal{V}(\mathcal{C})$ for $\mathcal{C}$ a subclass of finitely generated modules over a commutative Noetherian ring, they are Krull and hence atomic.

The set of distances and chains of factorizations. We now recall the Δ-set of a monoid H, an invariant which describes the sets of lengths of elements in H, and illustrate its relationship with distances between factorizations of elements in H. We denote by

$$\Delta(H) = \bigcup_{L \in \mathcal{L}(H)} \Delta(L) \subset \mathbb{N}$$

the *set of distances* of H. By definition, $\Delta(H) = \emptyset$ if and only if H is half-factorial. For a more thorough investigation of factorizations in H, we will need a distinguished subset of the set of distances. Let $\Delta^*(H)$ denote the set of all $d = \min \Delta(S)$ for some divisor-closed submonoid $S \subset H$ with $\Delta(S) \neq \emptyset$. By definition, we have $\Delta^*(H) \subset \Delta(H)$.

Suppose that H is not factorial. Then there exists an element $a \in H$ with $|\mathsf{Z}(a)| > 1$, and so there exist distinct $z, z' \in \mathsf{Z}(a)$. Then, for $N \in \mathbb{N}$, we have $\mathsf{Z}(a^N) \supset \{z^{N-i}(z')^i : i \in [0, N]\}$. Although $\mathsf{d}(z^N, (z')^N) = N d(z, z') > N$ suggests that the factorizations z^N and $(z')^N$ of a^N are very different,

$$\mathsf{d}(z^{N-i}(z')^i, z^{N-i+1}(z')^{i-1}) = \mathsf{d}(z, z')$$

for each $i \in [1, N]$. This illustrates that the distance alone is too coarse of an invariant, and motivates the study of the catenary degree as a way of measuring how distinct two factorizations are. As will be described below, there is a structure theorem for the set of lengths of a Krull monoid. However, except in very simple situations, there is no known structure theorem for the set of factorizations of an element in a Krull monoid. Thus we use the catenary degree, its many variations, the tame degree, and other invariants help to measure the subtle distinctions between factorizations.

Let $a \in H$ and $N \in \mathbb{N}_0 \cup \{\infty\}$. A finite sequence $z_0, \ldots, z_k \in \mathsf{Z}(a)$ is called a (*monotone*) *N-chain of factorizations* if $\mathsf{d}(z_{i-1}, z_i) \le N$ for all $i \in [1, k]$ and ($|z_0| \le \cdots \le |z_k|$ or $|z_0| \ge \cdots \ge |z_k|$ respectively). We denote by $\mathsf{c}(a)$ (or by $\mathsf{c}_{\mathrm{mon}}(a)$ respectively) the smallest $N \in \mathbb{N}_0 \cup \{\infty\}$ such that any two factorizations $z,\ z' \in \mathsf{Z}(a)$ can be concatenated by an N-chain (or by a monotone N-chain respectively). Then

$$\mathsf{c}(H) = \sup\{\mathsf{c}(b) : b \in H\} \in \mathbb{N}_0 \cup \{\infty\},$$

$$\mathsf{c}_{\mathrm{mon}}(H) = \sup\{\mathsf{c}_{\mathrm{mon}}(b) : b \in H\} \in \mathbb{N}_0 \cup \{\infty\}$$

denote the *catenary degree* and the *monotone catenary degree* of H. The monotone catenary degree is studied by using the two auxiliary notions of the equal and the adjacent catenary degrees. Let $\mathsf{c}_{\mathrm{eq}}(a)$ denote the smallest $N \in \mathbb{N}_0 \cup \{\infty\}$ such that any two factorizations $z, z' \in \mathsf{Z}(a)$ with $|z| = |z'|$ can be concatenated by a monotone N-chain. We call

$$\mathsf{c}_{\mathrm{eq}}(H) = \sup\{\mathsf{c}_{\mathrm{eq}}(b) : b \in H\} \in \mathbb{N}_0 \cup \{\infty\}$$

the *equal catenary degree* of H. We set

$$\mathsf{c}_{\mathrm{adj}}(a) = \sup\{\mathsf{d}(\mathsf{Z}_k(a), \mathsf{Z}_l(a)) : k, l \in \mathsf{L}(a) \text{ are adjacent}\},$$

and the *adjacent catenary degree* of H is defined as

$$\mathsf{c}_{\mathrm{adj}}(H) = \sup\{\mathsf{c}_{\mathrm{adj}}(b) : b \in H\} \in \mathbb{N}_0 \cup \{\infty\}.$$

Obviously, we have

$$\mathsf{c}(a) \le \mathsf{c}_{\mathrm{mon}}(a) = \sup\{\mathsf{c}_{\mathrm{eq}}(a), \mathsf{c}_{\mathrm{adj}}(a)\} \le \sup \mathsf{L}(a) \quad \text{for all } a \in H,$$

and hence

$$\mathsf{c}(H) \le \mathsf{c}_{\mathrm{mon}}(H) = \sup\{\mathsf{c}_{\mathrm{eq}}(H), \mathsf{c}_{\mathrm{adj}}(H)\}.$$

Note that $\mathsf{c}_{\mathrm{adj}}(H) = 0$ if and only if H is half-factorial, and if H is not half-factorial, then $2 + \sup \Delta(H) \le \mathsf{c}(H)$. Moreover, $\mathsf{c}_{\mathrm{eq}}(H) = 0$ if and only if for all $a \in H$ and all $k \in \mathsf{L}(a)$ we have $|\mathsf{Z}_k(a)| = 1$. Corollary 2.12 of [Coykendall and Smith 2011] implies that if D is a domain, we have that $\mathsf{c}_{\mathrm{eq}}(D^\bullet) = 0$ if and only if $D^\bullet$ is factorial.

We call

$$\sim_{H,\mathrm{eq}} = \big\{(x, y) \in \mathsf{Z}(H) \times \mathsf{Z}(H) : \pi(x) = \pi(y) \text{ and } |x| = |y|\big\}$$

the *monoid of equal-length relations* of H. Let $Z \subset \mathsf{Z}(H)$ be a subset. We say that an element $x \in Z$ is *minimal* in Z if for all elements $y \in Z$ with $y \mid x$ it follows that $x = y$. We denote by $\mathrm{Min}(Z)$ the *set of minimal elements* in Z. Let $x \in Z$. Since the number of elements $y \in Z$ with $y \mid x$ is finite, there exists an $x^* \in \mathrm{Min}(Z)$ with $x^* \mid x$.

Lemma 2.1. *Let H be an atomic monoid.*

(1) $\mathsf{c}_{\mathrm{eq}}(H) \le \sup\big\{|x| : (x, y) \in \mathcal{A}(\sim_{H,\mathrm{eq}})$ *for some* $y \in \mathsf{Z}(H) \setminus \{x\}\big\}$.

(2) *For* $d \in \Delta(H)$ *let* $A_d = \{x \in \mathsf{Z}(H) : |x| - d \in \mathsf{L}(\pi(x))\}$. *Then* $\mathsf{c}_{\mathrm{adj}}(H) \le \sup\{|x| : x \in \mathrm{Min}(A_d),\ d \in \Delta(H)\}$.

Proof. See [Blanco et al. 2011, Proposition 4.4]. □

Unions of sets of lengths and the refined elasticities. We now return to studying sets of lengths. We note that the elasticity of certain monoids of modules were studied in [Baeth and Luckas 2011; Baeth and Saccon 2012], but that in Section 6 we will provide results which generalize these results to larger classes of Krull monoids. In addition, we will fine tune these results by also computing the refined elasticities. Let $k, l \in \mathbb{N}$. If $H \neq H^\times$, then

$$\mathcal{U}_k(H) = \bigcup_{\substack{k \in L \\ L \in \mathcal{L}(H)}} L$$

is the union of all sets of lengths containing k. In other words, $\mathcal{U}_k(H)$ is set of all $m \in \mathbb{N}$ for which there exist $u_1, \ldots, u_k, v_1, \ldots, v_m \in \mathcal{A}(H)$ with $u_1 \cdot \ldots \cdot u_k = v_1 \cdot \ldots \cdot v_m$. When $H^\times = H$, we set $\mathcal{U}_k(H) = \{k\}$. In both cases, we define $\rho_k(H) = \sup \mathcal{U}_k(H) \in \mathbb{N} \cup \{\infty\}$ and $\lambda_k(H) = \min \mathcal{U}_k(H) \in [1, k]$. Clearly, we have $\mathcal{U}_1(H) = \{1\}$, $k \in \mathcal{U}_k(H)$, and since $\mathcal{U}_k(H) + \mathcal{U}_l(H) \subset \mathcal{U}_{k+l}(H)$, it follows that

$$\lambda_{k+l}(H) \le \lambda_k(H) + \lambda_l(H) \le k + l \le \rho_k(H) + \rho_l(H) \le \rho_{k+l}(H).$$

The *elasticity* $\rho(H)$ of H is defined as

$$\rho(H) = \sup\{\rho(L) : L \in \mathcal{L}(H)\} \in \mathbb{R}_{\ge 1} \cup \{\infty\},$$

and it is not difficult to verify that

$$\rho(H) = \sup\left\{\frac{\rho_k(H)}{k} : k \in \mathbb{N}\right\} = \lim_{k \to \infty} \frac{\rho_k(H)}{k}.$$

The structure of sets of lengths. To describe the structure of sets of lengths and of their unions, we need the concept of arithmetical progressions as well as various generalizations. Let $l, M \in \mathbb{N}_0$, $d \in \mathbb{N}$, and $\{0, d\} \subset \mathcal{D} \subset [0, d]$. We set

$$P_l(d) = d\mathbb{Z} \cap [0, ld] = \{0, d, 2d, \ldots, ld\}.$$

Thus a subset $L \subset \mathbb{Z}$ is an arithmetical progression (with *difference* $d \in \mathbb{N}$ and *length* $l \in \mathbb{N}_0$) if $L = \min L + P_l(d)$. A subset $L \subset \mathbb{Z}$ is called an *almost arithmetical multiprogression* (AAMP for short) with *difference* d, *period* $\mathcal{D}$, and *bound* M, if

$$L = y + (L' \cup L^* \cup L'') \subset y + \mathcal{D} + d\mathbb{Z},$$

where

- L^* is finite and nonempty with $\min L^* = 0$ and $L^* = (\mathcal{D} + d\mathbb{Z}) \cap [0, \max L^*]$,
- $L' \subset [-M, -1]$ and $L'' \subset \max L^* + [1, M]$, and
- $y \in \mathbb{Z}$.

Note that an AAMP is finite and nonempty, and that an AAMP with period $\{0, d\}$ and bound $M = 0$ is a (usual) arithmetical progression with difference d.

The ω-invariant and the tame degrees. We now study the ω-invariant as well as local and global tame degrees. We note that these notions have been studied in specific noncommutative module-theoretic situations in terms of the so-called *semiexchange property* (see [Diracca 2007]). Moreover, when describing the sets of lengths of elements within a Krull monoid H in terms of AAMPs (see Proposition 6.2), the bound M (described above) is a tame degree related to the monoid H. We begin with the definition. For an atom $u \in H$, let $\omega(H, u)$ denote the smallest $N \in \mathbb{N} \cup \{\infty\}$ having the following property:

> For any multiple a of u and any factorization $a = v_1 \cdot \ldots \cdot v_n$ of a, there exists a subset $\Omega \subset [1, n]$ such that $|\Omega| \le N$ and
>
> $$u \text{ divides } \prod_{\nu \in \Omega} v_\nu .$$

Furthermore, we set

$$\omega(H) = \sup\{\omega(H, u) : u \in \mathcal{A}(H)\} \in \mathbb{N} \cup \{\infty\}.$$

An atom $u \in H$ is prime if and only if $\omega(H, u) = 1$, and thus H is factorial if and only if $\omega(H) = 1$. If H satisfies the ascending chain condition on divisorial ideals (in particular, H is a Krull monoid or a Noetherian domain), then $\omega(H, u) < \infty$ for all $u \in \mathcal{A}(H)$ [Geroldinger and Hassler 2008, Theorem 4.2]. Roughly speaking, the tame degree $\mathsf{t}(H, u)$ is the maximum of $\omega(H, u)$ and a factorization length of $u^{-1} \prod_{\nu \in \Omega} v_\nu$ in H. More precisely, for an atom $u \in H$, the local tame degree $\mathsf{t}(H, u)$ is the smallest $N \in \mathbb{N}_0 \cup \{\infty\}$ having the following property:

> For any multiple a of u and any factorization $a = v_1 \cdot \ldots \cdot v_n$ of a which does not contain u, there is a short subproduct which is a multiple of u, say $v_1 \cdot \ldots \cdot v_m$, and a refactorization of this subproduct which contains u, say $v_1 \cdot \ldots \cdot v_m = u u_2 \cdot \ldots \cdot u_\ell$, such that $\max\{\ell, m\} \le N$.

Thus the local tame degree $\mathsf{t}(H, u)$ measures the distance between any factorization of a multiple a of u and a factorization of a which contains u. As before, we set

$$\mathsf{t}(H) = \sup\{\mathsf{t}(H, u) : u \in \mathcal{A}(H)\} \in \mathbb{N}_0 \cup \{\infty\}.$$

We conclude this section with the following lemma (see [Geroldinger and Halter-Koch 2006, Chapter 1; Geroldinger and Kainrath 2010]) which illustrates how the primary invariants measure the nonuniqueness of factorizations and show that all of these invariants are trivial if the monoid is factorial.

Lemma 2.2. *Let H be an atomic monoid.*

(1) *H is half-factorial if and only if $\rho(H) = 1$ if and only if $\rho_k(H) = k$ for every $k \in \mathbb{N}$.*

(2) *H is factorial if and only if $\mathsf{c}(H) = \mathsf{t}(H) = 0$ if and only if $\omega(H) = 1$.*

(3) $\mathsf{c}(H) = 0$ *or* $\mathsf{c}(H) \ge 2$, *and if* $\mathsf{c}(H) \le 2$, *then* H *is half-factorial.*

(4) $\mathsf{c}(H) \le \omega(H) \le \mathsf{t}(H) \le \omega(H)^2$, *and if* H *is not factorial*, *then*

$$\max\{2, \rho(H)\} \le \omega(H).$$

(5) *If* $\mathsf{c}(H) = 3$, *every* $L \in \mathcal{L}(H)$ *is an arithmetical progression with difference* 1.

3. Krull monoids, monoids of modules, and transfer homomorphisms

The theory of Krull monoids is presented in detail in the monographs [Halter-Koch 1998; Geroldinger and Halter-Koch 2006]. Here we gather the terminology required for our treatment. We then present an introduction to monoids of modules — the key objects of our study. Finally, we recall important terminology and results about monoids of zero-sum sequences and transfer homomorphisms — the key tools in our arithmetical investigations.

Krull monoids. Let H and D be monoids. A monoid homomorphism $\varphi : H \to D$ is called

- a *divisor homomorphism* if $\varphi(a) \mid \varphi(b)$ implies that $a \mid b$ for all $a, b \in H$.
- *cofinal* if for every $a \in D$ there exists some $u \in H$ such that $a \mid \varphi(u)$.
- a *divisor theory* (for H) if $D = \mathcal{F}(\mathcal{P})$ for some set $\mathcal{P}$, φ is a divisor homomorphism, and for every $a \in \mathcal{F}(\mathcal{P})$, there exists a finite nonempty subset $X \subset H$ satisfying $a = \gcd(\varphi(X))$.

We call $\mathcal{C}(\varphi) = \mathsf{q}(D)/\mathsf{q}(\varphi(H))$ the class group of φ, use additive notation for this group, and for $a \in \mathsf{q}(D)$, we denote by $[a] = a\,\mathsf{q}(\varphi(H)) \in \mathsf{q}(D)/\mathsf{q}(\varphi(H))$ the class containing a. Clearly $D/H = \{[a] : a \in D\} \subset \mathcal{C}(\varphi)$ is a submonoid with quotient group $\mathcal{C}(\varphi)$. The homomorphism φ is cofinal if and only if $\mathcal{C}(\varphi) = D/H$ and, by definition, every divisor theory is cofinal. Let $\varphi : H \to D = \mathcal{F}(\mathcal{P})$ be a divisor homomorphism. Then $\varphi(H) = \{a \in D : [a] = [1]\}$ and

$$G_{\mathcal{P}} = \{[p] = p\mathsf{q}(\varphi(H)) : p \in \mathcal{P}\} \subset \mathcal{C}(\varphi)$$

is called the *set of classes containing prime divisors*. Moreover, $\langle G_{\mathcal{P}} \rangle = \mathcal{C}(\varphi)$ and $[G_{\mathcal{P}}] = \{[a] : a \in D\}$.

The monoid H is called a *Krull monoid* if it satisfies one of the following equivalent conditions:

(a) H is completely integrally closed and satisfies the accending chain condition on divisorial ideals.

(b) H has a divisor theory.

(c) H has a divisor homomorphism into a free abelian monoid.

If H is a Krull monoid, then a divisor theory is unique up to unique isomorphism, and the class group associated to a divisor theory depends only on H. It is called the class group of H and will be denoted by $\mathcal{C}(H)$. Moreover, a reduced Krull monoid H with divisor theory $H \hookrightarrow \mathcal{F}(\mathcal{P})$ is uniquely determined up to isomorphism by its *characteristic* $(G, (m_g)_{g\in G})$ where G is an abelian group together with an isomorphism $\Phi : G \to \mathcal{C}(H)$ and with family $(m_g)_{g\in G}$ of cardinal numbers $m_g = |\mathcal{P} \cap \Phi(g)|$ (see [Geroldinger and Halter-Koch 2006, Theorem 2.5.4], and the forthcoming Lemma 3.4).

It is well known that a domain R is a Krull domain if and only if its multiplicative monoid $R^\bullet$ is a Krull monoid, and we set the class group of R to be $\mathcal{C}(R) = \mathcal{C}(R^\bullet)$. Property (a) shows that a Noetherian domain is Krull if and only if it is integrally closed. In addition, many well-studied classes of commutative monoids such as regular congruence monoids in Krull domains and Diophantine monoids are Krull. The focus of the present paper is on Krull monoids stemming from module theory.

Monoids of modules. Let R be a (not necessarily commutative) ring and $\mathcal{C}$ a class of (right) R-modules. We say that $\mathcal{C}$ is *closed under finite direct sums, direct summands, and isomorphisms* provided the following holds: Whenever M, M_1 and M_2 are R-modules with $M \cong M_1 \oplus M_2$, we have $M \in \mathcal{C}$ if and only if $M_1, M_2 \in \mathcal{C}$. We say that $\mathcal{C}$ satisfies the KRSA theorem if the following holds:

> If $k, l \in \mathbb{N}$ and $M_1, \dots, M_k, N_1, \dots, N_l$ are indecomposable modules in $\mathcal{C}$ with $M_1 \oplus \cdots \oplus M_k \cong N_1 \oplus \cdots \oplus N_l$, then $l = k$ and, after a possible reordering of terms, $M_i \cong N_i$ for all $i \in [1, k]$.

Suppose that $\mathcal{C}$ is closed under finite direct sums, direct summands, and isomorphisms. For a module $M \in \mathcal{C}$, we denote by $[M]$ its isomorphism class, and by $\mathcal{V}(\mathcal{C})$ the set of isomorphism classes. (For our purposes here, we tacitly assume that this is actually a set. For the classes of modules studied in Sections 4 and 5 this is indeed the case. For the involved set-theoretical problems in a more general context, see [Facchini 2012, Section 2].) Then $\mathcal{V}(\mathcal{C})$ is a commutative semigroup with operation $[M] + [N] = [M \oplus N]$ and all information about direct-sum decompositions of modules in $\mathcal{C}$ can be studied in terms of factorizations in the semigroup $\mathcal{V}(\mathcal{C})$. By definition, $\mathcal{C}$ satisfies KRSA if and only if $\mathcal{V}(\mathcal{C})$ is a free abelian monoid, which holds if $\mathrm{End}_R(M)$ is local for each indecomposable M in $\mathcal{C}$ (see [Leuschke and Wiegand 2012, Theorem 1.3]).

If the endomorphism ring $\mathrm{End}_R(M)$ is semilocal for all modules M in $\mathcal{C}$, then $\mathcal{V}(\mathcal{C})$ is a Krull monoid ([Facchini 2002, Theorem 3.4]). There is an abundance of recent work which provides examples of rings and classes of modules over these rings for which all endomorphism rings are semilocal (see [Facchini 2004; 2006; 2012]. For monoids of modules, a characterization of when the class group is a torsion group is given in [Facchini and Halter-Koch 2003]).

Suppose that $\mathcal{V}(\mathcal{C})$ is a Krull monoid. Then to understand the structure of direct-sum decompositions of modules in $\mathcal{C}$ is to understand the arithmetic of the reduced Krull monoid $\mathcal{V}(\mathcal{C})$. Since any reduced Krull monoid H is uniquely determined by its class group and by the distribution of prime divisors (that is, the characteristic of H), one must study these parameters.

In the present paper we will focus on the following classes of modules over a commutative Noetherian local ring S, each closed under finite direct sums, direct summands, and isomorphisms. For a commutative Noetherian local ring S, we denote by

- $\mathcal{M}(S)$ the semigroup of isomorphism classes of finitely generated S-modules,
- $\mathcal{T}(S)$ the semigroup of isomorphism classes of finitely generated torsion-free S-modules, and
- $\mathfrak{C}(S)$ the semigroup of isomorphism classes of maximal Cohen–Macaulay (MCM) S-modules.

Note that in order to make $\mathfrak{C}(S)$ a semigroup, we insist that $[0_S] \in \mathfrak{C}(S)$, even though the zero module is not MCM. We say that a commutative Noetherian local ring S has finite representation type if there are, up to isomorphism, only finitely many indecomposable MCM S-modules. Otherwise we say that S has infinite representation type.

Throughout, let $(R, \mathfrak{m})$ be a commutative Noetherian local ring with maximal ideal $\mathfrak{m}$, and let $(\widehat{R}, \widehat{\mathfrak{m}})$ denote its $\mathfrak{m}$-adic completion. Let $\mathcal{V}(R)$ and $\mathcal{V}(\widehat{R})$ be any of the above three semigroups. If M is an R-module such that $[M] \in \mathcal{V}(R)$, then $\widehat{M} \cong M \otimes_R \widehat{R}$ is an $\widehat{R}$-module with $[\widehat{M}] \in \mathcal{V}(\widehat{R})$, and every such $\widehat{R}$-module is called *extended*. Note that R has finite representation type if and only if $\widehat{R}$ has finite representation type (see [Leuschke and Wiegand 2012, Chapter 10]), and that the dimension of R is equal to the dimension of $\widehat{R}$. The following crucial result shows that the monoid $\mathcal{V}(R)$ is Krull.

Lemma 3.1. *Let $(R, \mathfrak{m})$ be a commutative Noetherian local ring with maximal ideal $\mathfrak{m}$, and let $(\widehat{R}, \widehat{\mathfrak{m}})$ denote its $\mathfrak{m}$-adic completion.*

(1) *For each indecomposable finitely generated $\widehat{R}$-module M, $\mathrm{End}_{\widehat{R}}(M)$ is local, and therefore $\mathcal{M}(\widehat{R})$, $\mathcal{T}(\widehat{R})$, and $\mathfrak{C}(\widehat{R})$ are free abelian monoids.*

(2) *The embedding $\mathcal{M}(R) \hookrightarrow \mathcal{M}(\widehat{R})$ is a divisor homomorphism. It is cofinal if and only if every finitely generated $\widehat{R}$-module is a direct summand of an extended module.*

(3) *The embeddings $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ and $\mathfrak{C}(R) \hookrightarrow \mathfrak{C}(\widehat{R})$ are divisor homomorphisms.*

In particular, $\mathcal{M}(R), \mathcal{T}(R)$, and $\mathfrak{C}(R)$ are reduced Krull monoids. Moreover, the embeddings in (2) *and* (3) *are injective and map R-modules onto the submonoid of extended $\widehat{R}$-modules.*

Proof. Property (1) holds by the KRSA theorem (see [Leuschke and Wiegand 2012, Chapter 1]).

Wiegand [2001] proved that the given embedding is a divisor homomorphism (see also [Baeth and Wiegand 2013, Theorem 3.6]). The characterization of cofinality follows from the definition and thus (2) holds.

Let M, N be R-modules such that either $[M], [N] \in \mathcal{V}(R)$ where $\mathcal{V}(R)$ denotes either $\mathcal{T}(R)$ or $\mathfrak{C}(R)$ and suppose that $[\widehat{M}]$ divides $[\widehat{N}]$ in $\mathcal{V}(\widehat{R})$. Then we have divisibility in $\mathcal{M}(\widehat{R})$, and hence in $\mathcal{M}(R)$ by (2). Since $\mathcal{V}(R) \subset \mathcal{M}(R)$ is divisor-closed, it follows that $[M]$ divides $[N]$ in $\mathcal{V}(R)$, proving (3).

Together, (2) and (3) show that $\mathcal{M}(R), \mathcal{T}(R)$, and $\mathfrak{C}(R)$ satisfy Property (c) in the definition of Krull monoids. Since each of these monoids is reduced, the maps induced by $[M] \mapsto [\widehat{M}]$ are injective. □

Note that the embedding $\mathcal{M}(R) \hookrightarrow \mathcal{M}(\widehat{R})$ is not necessarily cofinal, as is shown in [Hassler and Wiegand 2009; Frankild et al. 2008]. In Sections 4 and 5 we will study in detail the class group and the distribution of prime divisors of these Krull monoids, in the case of one-dimensional and two-dimensional commutative Noetherian local rings.

Monoids of zero-sum sequences. We now introduce Krull monoids having a combinatorial flavor which are used to model arbitrary Krull monoids. Let G be an additive abelian group and let $G_0 \subset G$ be a subset. Following the tradition in additive group and number theory, we call the elements of $\mathcal{F}(G_0)$ *sequences* over G_0. Thus a sequence $S \in \mathcal{F}(G_0)$ will be written in the form

$$S = g_1 \cdot \ldots \cdot g_l = \prod_{g \in G_0} g^{\mathsf{v}_g(S)}.$$

We will use all notions (such as the length) as in general free abelian monoids. We set $-S = (-g_1) \cdot \ldots \cdot (-g_l)$, and call $\sigma(S) = g_1 + \cdots + g_l \in G$ the *sum* of S. The monoid

$$\mathcal{B}(G_0) = \{S \in \mathcal{F}(G_0) : \sigma(S) = 0\}$$

is called the *monoid of zero-sum sequences* over G_0, and its elements are called *zero-sum sequences* over G_0. Obviously, the inclusion $\mathcal{B}(G_0) \hookrightarrow \mathcal{F}(G_0)$ is a divisor homomorphism, and hence $\mathcal{B}(G_0)$ is a reduced Krull monoid by Property (c) in the definition of Krull monoids. By definition, the inclusion $\mathcal{B}(G_0) \hookrightarrow \mathcal{F}(G_0)$ is cofinal if and only if for every $g \in G_0$ there is an $S \in \mathcal{B}(G_0)$ with $g \mid S$; equivalently, there is no proper subset $G_0' \subsetneq G_0$ such that $\mathcal{B}(G_0') = \mathcal{B}(G_0)$. If $|G| \neq 2$, then $\mathcal{C}(\mathcal{B}(G)) \cong G$, and every class contains precisely one prime divisor.

For every arithmetical invariant $*(H)$, as defined for a monoid H in Section 2, it is usual to write $*(G_0)$ instead of $*(\mathcal{B}(G_0))$ (whenever the meaning is clear from the context). In particular, we set $\mathcal{A}(G_0) = \mathcal{A}(\mathcal{B}(G_0))$, $\mathcal{L}(G_0) = \mathcal{L}(\mathcal{B}(G_0))$, $\mathsf{c}_{\text{mon}}(G_0) = \mathsf{c}_{\text{mon}}(\mathcal{B}(G_0))$, etc.

The study of sequences, subsequence sums, and zero-sums is a flourishing subfield of additive group and number theory (see, for example, [Gao and Geroldinger 2006; Geroldinger and Ruzsa 2009; Grynkiewicz 2013]). The *Davenport constant* $\mathsf{D}(G_0)$, defined as

$$\mathsf{D}(G_0) = \sup\{|U| : U \in \mathcal{A}(G_0)\} \in \mathbb{N}_0 \cup \{\infty\},$$

is among the most studied invariants in additive theory and will play a crucial role in the computations of arithmetical invariants (see the discussion after Lemma 3.4). We will need the following two simple lemmas which we present here so as to not clutter the exposition of Section 6.

Lemma 3.2. *Suppose that the inclusion $\mathcal{B}(G_0) \hookrightarrow \mathcal{F}(G_0)$ is cofinal. The following are equivalent.*

(a) *There exist nontrivial submonoids $H_1, H_2 \subset \mathcal{B}(G_0)$ such that $\mathcal{B}(G_0) = H_1 \times H_2$.*

(b) *There exist nonempty subsets $G_1, G_2 \subset G_0$ such that $G_0 = G_1 \uplus G_2$ and $\mathcal{B}(G_0) = \mathcal{B}(G_1) \times \mathcal{B}(G_2)$.*

(c) *There exist nonempty subsets $G_1, G_2 \subset G_0$ such that $G_0 = G_1 \uplus G_2$ and $\mathcal{A}(G_0) = \mathcal{A}(G_1) \uplus \mathcal{A}(G_2)$.*

Proof. Clearly (b) implies (a). The converse follows from [Geroldinger and Halter-Koch 2006, Proposition 2.5.6]. The implication (b) implies (c) is obvious. We now show that (c) implies (b). Let $B \in \mathcal{B}(G_0)$. Since $\mathcal{B}(G_0)$ is a Krull monoid, it is atomic and hence $B = U_1 \cdot \ldots \cdot U_l$ with $U_1, \ldots, U_l \in \mathcal{A}(G_0)$. After renumbering (if necessary), we can find $k \in [0, l]$ such that $U_1, \ldots, U_k \in \mathcal{A}(G_1)$ and $U_{k+1}, \ldots, U_l \in \mathcal{A}(G_2)$. Thus $\mathcal{B}(G_0) = \mathcal{B}(G_1)\mathcal{B}(G_2)$. If $B \in \mathcal{B}(G_1) \cap \mathcal{B}(G_2)$, then B is a product of atoms from $\mathcal{A}(G_1)$ and a product of atoms from $\mathcal{A}(G_2)$. Since their intersection is empty, both products are empty. Therefore $B = 1$ and hence $\mathcal{B}(G_0) = \mathcal{B}(G_1) \times \mathcal{B}(G_2)$. $\square$

Lemma 3.2.(c) shows that $\mathcal{B}(G^\bullet)$ is not a direct product of submonoids. Suppose that $0 \in G_0$. Then $0 \in \mathcal{B}(G_0)$ is a prime element and $\mathcal{B}(G_0) = \mathcal{B}(\{0\}) \times \mathcal{B}(G_0^\bullet)$. But $\mathcal{B}(\{0\}) = \mathcal{F}(\{0\}) \cong (\mathbb{N}_0, +)$, and thus all the arithmetical invariants measuring the nonuniqueness of factorizations of $\mathcal{B}(G_0)$ and of $\mathcal{B}(G_0^\bullet)$ coincide. Therefore we can assume that $0 \notin G_0$ whenever it is convenient.

Lemma 3.3. *Let G be an abelian group and let $G_0 \subset G$ be a subset such that $1 < \mathsf{D}(G_0) < \infty$.*

(1) *For all $k \in \mathbb{N}$,*

$$\rho(G_0) \le \mathsf{D}(G_0)/2, \quad k \le \rho_k(G_0) \le k\rho(G_0), \quad \rho(G_0)^{-1}k \le \lambda_k(G_0) \le k.$$

(2) *Suppose that $\rho_2(G_0) = \mathsf{D}(G_0)$. Then $\rho(G_0) = \mathsf{D}(G_0)/2$, and for all $k \in \mathbb{N}$,*

$$\rho_{2k}(G_0) = k\mathsf{D}(G_0) \quad \textit{and} \quad k\mathsf{D}(G_0) + 1 \le \rho_{2k+1}(G_0) \le k\mathsf{D}(G_0) + \frac{\mathsf{D}(G_0)}{2}.$$

Moreover, if $j, l \in \mathbb{N}_0$ are such that $l\mathsf{D}(G_0) + j \ge 1$, then

$$2l + \frac{2j}{\mathsf{D}(G_0)} \le \lambda_{l\mathsf{D}(G_0)+j}(G_0) \le 2l + j.$$

Proof. By definition, $\lambda_k(G_0) \le k \le \rho_k(G_0)$. Since $\rho(G_0) = \sup\{\rho_k(G_0)/k : k \in \mathbb{N}\}$, it follows that $\rho_k(G_0) \le k\rho(G_0)$ and $k \le \rho(G_0)\lambda_k(G_0)$. Furthermore, $2\rho_k(G_0) \le k\mathsf{D}(G_0)$ for all $k \in \mathbb{N}$ implies that $\rho(G_0) \le \mathsf{D}(G_0)/2$. This gives (1).

We now prove (2). Since $\rho_k(G_0) + \rho_l(G_0) \le \rho_{k+l}(G_0)$ for every $k, l \in \mathbb{N}$, (1) implies that

$$k\mathsf{D}(G_0) = k\rho_2(G_0) \le \rho_{2k}(G_0) \le (2k)\frac{\mathsf{D}(G_0)}{2} = k\mathsf{D}(G_0),$$

and hence

$$\begin{aligned} k\mathsf{D}(G_0) + 1 = \rho_{2k}(G_0) + \rho_1(G_0) \le \rho_{2k+1}(G_0) &\le (2k+1)\rho(G_0) \\ &\le k\mathsf{D}(G_0) + \frac{\mathsf{D}(G_0)}{2}. \end{aligned}$$

Let $j, l \in \mathbb{N}_0$ be such that $l\mathsf{D}(G_0)+j \ge 1$. For convenience, set $\rho_0(G_0) = \lambda_0(G_0) = 0$. Since

$$2l = \frac{2}{\mathsf{D}(G_0)} l\mathsf{D}(G_0) \le \lambda_{l\mathsf{D}(G_0)}(G_0) \quad \text{and} \quad \rho_{2l}(G_0) = l\mathsf{D}(G_0),$$

it follows that $\lambda_{l\mathsf{D}(G_0)}(G_0) = 2l$, and hence

$$\begin{aligned} 2l + \frac{2j}{\mathsf{D}(G_0)} &= \frac{2}{\mathsf{D}(G_0)}\big(l\mathsf{D}(G_0) + j\big) = \rho(G_0)^{-1}\big(l\mathsf{D}(G_0) + j\big) \\ &\le \lambda_{l\mathsf{D}(G_0)+j}(G_0) \le \lambda_{l\mathsf{D}(G_0)}(G_0) + \lambda_j(G_0) \le 2l + j. \qquad \square \end{aligned}$$

Transfer homomorphisms. Transfer homomorphisms are a central tool in factorization theory. In order to study a given monoid H, one constructs a transfer homomorphism $\theta : H \to B$ to a simpler monoid B, studies factorizations in B, and then lifts arithmetical results from B to H. In the case of Krull monoids, transfer homomorphisms allow one to study nearly all of the arithmetical invariants introduced in Section 2 in an associated monoid of zero-sum sequences. We now gather the basic tools necessary for this approach.

A monoid homomorphism $\theta : H \to B$ is called a *transfer homomorphism* if it has the following properties:

(T1) $B = \theta(H)B^\times$ and $\theta^{-1}(B^\times) = H^\times$.

(T2) If $u \in H$, $b, c \in B$ and $\theta(u) = bc$, then there exist $v, w \in H$ such that $u = vw$, $\theta(v) \simeq b$ and $\theta(w) \simeq c$.

The next result provides the link between the arithmetic of Krull monoids and additive group and number theory. This interplay is highlighted in the survey [Geroldinger 2009].

Lemma 3.4. *Let H be a Krull monoid, $\varphi : H \to D = \mathcal{F}(\mathcal{P})$ a cofinal divisor homomorphism, $G = \mathcal{C}(\varphi)$ its class group, and $G_\mathcal{P} \subset G$ the set of classes containing prime divisors. Let $\widetilde{\boldsymbol{\beta}} : D \to \mathcal{F}(G_\mathcal{P})$ denote the unique homomorphism defined by $\widetilde{\boldsymbol{\beta}}(p) = [p]$ for all $p \in \mathcal{P}$.*

(1) *The inclusion $\mathcal{B}(G_\mathcal{P}) \hookrightarrow \mathcal{F}(G_\mathcal{P})$ is cofinal, and the homomorphism*
$$\boldsymbol{\beta} = \widetilde{\boldsymbol{\beta}} \circ \varphi : H \to \mathcal{B}(G_\mathcal{P})$$
is a transfer homomorphism.

(2) *For all $a \in H$, $\mathsf{L}_H(a) = \mathsf{L}_{\mathcal{B}(G_\mathcal{P})}(\boldsymbol{\beta}(a))$. In particular, $\mathcal{L}(H) = \mathcal{L}(G_\mathcal{P})$, $\Delta(H) = \Delta(G_\mathcal{P})$, $\mathcal{U}_k(H) = \mathcal{U}_k(G_\mathcal{P})$, $\rho_k(H) = \rho_k(G_\mathcal{P})$, and $\lambda_k(H) = \lambda_k(G_\mathcal{P})$ for each $k \in \mathbb{N}$.*

(3) *Suppose that H is not factorial. Then $\mathsf{c}(H) = \mathsf{c}(G_\mathcal{P})$, $\mathsf{c}_{\text{adj}}(H) = \mathsf{c}_{\text{adj}}(G_\mathcal{P})$, $\mathsf{c}_{\text{mon}}(H) = \mathsf{c}_{\text{mon}}(G_\mathcal{P})$, $\Delta^*(H) = \Delta^*(G_\mathcal{P})$, and $\omega(H) \le \mathsf{D}(G_\mathcal{P})$.*

Proof. See [Geroldinger and Halter-Koch 2006, Section 3.4] for details pertaining to most of the invariants. For the statements on the monotone catenary degree, see [Geroldinger et al. 2010]. Roughly speaking, all of the statements in (2) are straightforward, but the statements in (3) are more subtle. Note that a statement corresponding to (3) does not hold true for the tame degree (see [Gao et al. 2015]). □

In summary, if the monoid of modules $\mathcal{V}(R)$ is Krull with class group G and set $G_\mathcal{P}$ of classes containing prime divisors, then the arithmetic of direct-sum decompositions can be studied in the monoid $\mathcal{B}(G_\mathcal{P})$ of zero-sum sequences over $G_\mathcal{P}$. In particular, if $H = \mathcal{M}(R)$ and $D = \mathcal{M}(\widehat{R})$ as in Lemma 3.1 and all notation is as in Lemma 3.4, then

$$\mathsf{D}(G_\mathcal{P}) = \sup\bigl\{l : \widehat{M} \cong N_1 \oplus \cdots \oplus N_l \text{ with } [M] \in \mathcal{A}(H) \text{ and } [N_i] \in \mathcal{A}(D) \ \forall i \in [1, l]\bigr\}.$$

4. Monoids of modules: class groups and distribution of prime divisors, I

Throughout this section we use the following setup:

(S) $(R, \mathfrak{m})$ *denotes a one-dimensional analytically unramified commutative Noetherian local ring with unique maximal ideal* $\mathfrak{m}$, $k = R/\mathfrak{m}$ *its residue field,* $\widehat{R}$ *its* $\mathfrak{m}$*-adic completion, and* $\mathrm{spl}(R) = |\mathrm{spec}(\widehat{R})| - |\mathrm{spec}(R)|$ *the splitting number of* R.

In this section we investigate the characteristic of the Krull monoids $\mathcal{M}(R)$ and $\mathcal{T}(R)$ for certain one-dimensional local rings. This study is based on deep module-theoretic work achieved over the past several decades. We gather together module-theoretic information and proceed using a recent construction (see 4.4) to obtain results on the class group and on the set $G_{\mathcal{P}}$ of classes containing prime divisors. The literature does not yet contain a systematic treatment along these lines. Indeed, early results (see Theorem 4.2 below) indicated only the existence of extremal sets $G_{\mathcal{P}}$ which imply either trivial direct-sum decompositions or that all arithmetical invariants describing the direct-sum decompositions are infinite. In either case there was no need for further arithmetical study. Here we reveal that finite and well-structured sets $G_{\mathcal{P}}$ occur in abundance. Thus, as we will see in Section 6, the arithmetical behavior of direct-sum decompositions is well-structured.

We first gather basic ring and module-theoretic properties. By definition, $\widehat{R}$ and R are both reduced and the integral closure of R is a finitely generated R-module. Moreover, we have $\mathfrak{C}(R) = \mathcal{T}(R)$. Let M be a finitely generated R-module. If $\mathfrak{p}$ is a minimal prime ideal of R, then $R_{\mathfrak{p}}$ is a field, $M_{\mathfrak{p}}$ is a finite-dimensional $R_{\mathfrak{p}}$-vector space, and we set $\mathrm{rank}_{\mathfrak{p}}(M) = \dim_{R_{\mathfrak{p}}}(M_{\mathfrak{p}})$. If $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ are the minimal prime ideals of R, then $\mathrm{rank}(M) = (r_1, \ldots, r_s)$ where $r_i = \mathrm{rank}_{\mathfrak{p}_i}(M)$ for all $i \in [1, s]$. The module M is said to have constant rank if $r_1 = \cdots = r_s$.

We start with a beautiful result of Levy and Odenthall, which gives us a tool to determine which finitely generated $\widehat{R}$-modules are extended from R-modules.

Proposition 4.1 [Levy and Odenthal 1996, Theorem 6.2]. *Let M be a finitely generated torsion-free $\widehat{R}$-module. Then M is extended if and only if* $\mathrm{rank}_{\mathfrak{p}}(M) = \mathrm{rank}_{\mathfrak{q}}(M)$ *whenever $\mathfrak{p}$ and $\mathfrak{q}$ are minimal prime ideals of $\widehat{R}$ with $\mathfrak{p} \cap R = \mathfrak{q} \cap R$. In particular, if R is a domain, then M is extended if and only if its rank is constant.*

We start our discussion with a result which completely determines the characteristic of the Krull monoid $\mathcal{M}(R)$. The arithmetic of this monoid is studied in Proposition 6.2.2.

Theorem 4.2 [Hassler et al. 2007, Theorem 6.3]. *Let G denote the class group of $\mathcal{M}(R)$ and let $G_{\mathcal{P}} \subset G$ denote the set of classes containing prime divisors.*

(1) *If R is not Dedekind-like, then G is free abelian of rank* $\mathrm{spl}(R)$ *and each class contains $|k|\aleph_0$ prime divisors.*

(2) *If R is a DVR, then $G = 0$.*
(3) *If R is Dedekind-like but not a DVR, then either*
   (a) $\mathrm{spl}(R) = 0$ *and* $G = 0$, *or*
   (b) $\mathrm{spl}(R) = 1$, *G is infinite cyclic with $G = \langle e \rangle$ and $G_{\mathcal{P}} = \{-e, 0, e\}$. Each of the classes e and $-e$ contain $\aleph_0$ prime divisors and the class 0 contains $|k|\aleph_0$ prime divisors.*

Thus, for the rest of this section, we focus our attention on $\mathcal{T}(R)$. To determine if the divisor homomorphism $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ is a divisor theory, we will require additional information. For now, we easily show that it is always cofinal.

Proposition 4.3. *The embedding $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ is a cofinal divisor homomorphism.*

Proof. By Lemma 3.1 the embedding is a divisor homomorphism. If M is a finitely generated torsion-free $\widehat{R}$-module, we can consider its rank, $\mathrm{rank}(M) = (r_1, \ldots, r_t)$, where t is the number of minimal primes of $\widehat{R}$. If $r_1 = \cdots = r_t$, then by Proposition 4.1 M is extended, say $M = \widehat{N}$ for some finitely generated torsion-free R-module N and the result is trivial. If the rank of M is not constant, set $r = \max\{r_1, \ldots, r_t\}$ and consider the $\widehat{R}$-module

$$L = (\widehat{R}/\mathfrak{q}_1)^{r-r_1} \oplus \cdots \oplus (\widehat{R}/\mathfrak{q}_t)^{r-r_t},$$

where $\mathfrak{q}_1, \ldots, \mathfrak{q}_t$ denote the minimal primes of $\widehat{R}$. Then $\mathrm{rank}(N \oplus L) = (r, \ldots, r)$ is constant and hence $N \oplus L$ is extended, say $N \oplus L \cong \widehat{P}$ for some finitely generated torsion-free R-module P. Clearly M is isomorphic to a direct summand of $\widehat{P}$ and the result follows. $\square$

Since $\mathcal{T}(\widehat{R})$ is free abelian, we can identify it with the free abelian monoid $\mathbb{N}_0^{(\mathcal{P})}$, where $\mathcal{P}$ is an index set for the isomorphism classes of indecomposable finitely generated torsion-free $\widehat{R}$-modules. We then use Proposition 4.1 to describe $\mathcal{T}(R)$ in detail. The following construction has been used numerous times (see, for example, [Baeth and Luckas 2011; Baeth and Saccon 2012; Facchini et al. 2006]).

Construction 4.4. • Let $\mathfrak{p}_1, \ldots, \mathfrak{p}_s$ be the distinct minimal prime ideals of R. For each $i \in [1, s]$, let $\mathfrak{q}_{i,1}, \ldots, \mathfrak{q}_{i,t_i}$ be the minimal primes of $\widehat{R}$ lying over $\mathfrak{p}_i$. Note that $\mathrm{spl}(R) = \sum_{i=1}^{s}(t_i - 1)$.

- Let $\mathcal{P}$ be the set of isomorphism classes of indecomposable finitely generated torsion-free $\widehat{R}$-modules.
- Let $\mathsf{A}(R)$ be the $\mathrm{spl}(R) \times |\mathcal{P}|$ matrix whose column indexed by the isomorphism class $[M] \in \mathcal{P}$ is

$$\begin{bmatrix} r_{1,1} - r_{1,2} & \cdots & r_{1,1} - r_{1,t_1} & \cdots & r_{s,1} - r_{s,2} & \cdots & r_{s,1} - r_{s,t_s} \end{bmatrix}^T,$$

where $r_{i,j} = \mathrm{rank}_{\mathfrak{q}_{i,j}}(M)$.

Then $\mathcal{T}(R) \cong \ker(\mathsf{A}(R)) \cap \mathbb{N}^{(\mathcal{P})} \subset \mathbb{N}_0^{(\mathcal{P})}$ is a Diophantine monoid.

If one has a complete description of how the minimal prime ideals of $\widehat{R}$ lie over the minimal prime ideals of R together with the ranks of all indecomposable finitely generated torsion-free $\widehat{R}$-modules, then Construction 4.4 completely describes the monoid $\mathcal{T}(R)$. In certain cases (e.g., Section 4A) we are able to obtain all of this information. Other times we know only some of the ranks that occur for indecomposable $\widehat{R}$-modules and thus have only a partial description for $\mathcal{T}(R)$. However, as was shown in [Baeth and Saccon 2012], the ranks of indecomposable cyclic $\widehat{R}$-modules gives enough information about the columns of $\mathsf{A}(R)$ to prove that $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ is nearly always a divisor theory. First we recall that if $\mathfrak{q}_1, \ldots, \mathfrak{q}_t$ are the minimal primes of $\widehat{R}$, and $E \subset [1, t]$. Then

$$\operatorname{rank}\left(\frac{\widehat{R}}{\bigcap_{i \in E} \mathfrak{q}_i}\right) = (r_1, \ldots, r_t), \qquad \text{where } r_i = \begin{cases} 1 & \text{if } i \in E, \\ 0 & \text{if } i \notin E. \end{cases}$$

Thus *every* nontrivial t-tuple of zeros and ones can be realized as the rank of a nonzero (necessarily indecomposable) cyclic $\widehat{R}$-module. Thus we have the following:

Construction 4.5. Let all notation be as in Construction 4.4. After renumbering if necessary, there is $p \in [0, s]$ such that $t_1, \ldots, t_p \geq 2$ and such that $t_i = 1$ for each $i \in [p+1, s]$. Then $\operatorname{spl}(R) = \sum_{j=1}^{p} t_j - p$. For each $i \in [1, p]$, let A_i be the set of $(t_i - 1) \times 1$ column vectors all of whose entries are either 0 or 1, and let B_i be the set of $(t_i - 1) \times 1$ column vectors all of whose entries are either 0 or -1.

We now define $\mathcal{T}$ to be the $\operatorname{spl}(R) \times \prod_{i=1}^{p} (2^{t_i} - 1)$ matrix, each of whose columns has the form

$$\left[\begin{array}{c} T_1 \\ \hline \vdots \\ \hline T_p \end{array}\right], \qquad \text{where } T_i \in A_i \cup B_i \text{ for each } i \in [1, p].$$

With the notation as in Constructions 4.4 and 4.5, we give a realization result which shows that the matrix $\mathcal{T}$ occurs as a submatrix of $\mathsf{A}(R)$.

Proposition 4.6 [Baeth and Saccon 2012, Proposition 3.7]. *For each column $\boldsymbol{\alpha}$ of $\mathcal{T}$, there exist nonnegative integers $r_{i,j}$ and an indecomposable torsion-free $\widehat{R}$-module $M_{\boldsymbol{\alpha}}$ of rank*

$$(r_{1,1}, \ldots, r_{1,t_1}, \ldots, r_{p,1}, \ldots, r_{p,t_p}, r_{p+1,1}, \ldots, r_{s,1})$$

such that

$$\boldsymbol{\alpha} = \begin{bmatrix} r_{1,1} - r_{1,2} & \cdots & r_{1,1} - r_{1,t_1} & \cdots & r_{p,1} - r_{p,2} & \cdots & r_{p,1} - r_{p,t_p} \end{bmatrix}^T.$$

In particular, the matrix $\mathsf{A}(R)$ nearly always satisfies the hypotheses of the following lemma.

Lemma 4.7 [Baeth and Saccon 2012, Lemma 4.1]. *Fix an integer $q \geq 1$, and let I_q denote the $q \times q$ identity matrix. Let $\mathcal{P}$ be an index set, and let $\mathcal{D}$ be a $q \times |\mathcal{P}|$ integer matrix whose columns are indexed by $\mathcal{P}$. Assume $\mathcal{D} = \big[D_1 \mid D_2\big]$, where D_1 is the $q \times (2q+2)$ integer matrix*

$$\left[\begin{array}{c|c|cc} & & 1 & -1 \\ I_q & -I_q & \vdots & \vdots \\ & & 1 & -1 \end{array}\right],$$

and D_2 is an arbitrary integer matrix with q rows (and possibly infinitely many columns). Let $H = \ker(\mathcal{D}) \cap \mathbb{N}_0^{(\mathcal{P})}$.

(1) *The map $\mathcal{D} : \mathbb{Z}^{(\mathcal{P})} \to \mathbb{Z}^{(q)}$ is surjective.*

(2) *The natural inclusion $H \hookrightarrow \mathbb{N}_0^{(\mathcal{P})}$ is a divisor theory.*

(3) $\ker(\mathcal{D}) = \mathsf{q}(H)$.

(4) *$\mathcal{C}(H) \cong \mathbb{Z}^{(q)}$, and this isomorphism maps the set of classes containing prime divisors onto the set of distinct columns of $\mathcal{D}$.*

In particular, we observe the following: Given a fixed column α of $\mathcal{D}$, the cardinality of $\{\beta : \beta$ is a column of $\mathcal{D}$ and $\beta = \alpha\}$ is equal to the cardinality of prime divisors in the class corresponding to α. Therefore, the characteristic of the Krull monoid H is completely given by the matrix $\mathcal{D}$.

Based on the previous results, one easily obtains the following theorem which provides the framework for our study of the characteristic of $\mathcal{T}(R)$.

Theorem 4.8. (1) *If $\operatorname{spl}(R) = 0$, then $\mathcal{T}(R) \cong \mathcal{T}(\widehat{R})$ is free abelian.*

(2) *If $\operatorname{spl}(R) \geq 2$ then the embedding $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ is a divisor theory. Moreover,*

  (a) $\mathcal{T}(R) \cong \ker(\mathsf{A}(R)) \cap \mathbb{N}_0^{(\mathcal{P})}$,

  (b) *$\mathcal{C}(\mathcal{T}(R)) \cong \mathbb{Z}^{(\operatorname{spl}(R))}$, and this isomorphism maps the set of classes containing prime divisors onto the set of distinct columns of $\mathsf{A}(R)$.*

Suppose that $\operatorname{spl}(R) = 1$. The embedding $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ is a divisor theory if and only if the defining matrix $\mathsf{A}(R)$ contains at least two positive and at least two negative entries (see Proposition 6.1.2).

In many cases, computing the ranks of indecomposable $\widehat{R}$-modules and hence the columns of the defining matrix $\mathsf{A}(R)$ is difficult. However, an additional hypotheses on R implies that the set of classes containing prime divisors satisfies $G_{\mathcal{P}} = -G_{\mathcal{P}}$, a crucial property for all arithmetical investigations (see Proposition 6.2 and the subsequent remarks).

Corollary 4.9. *Suppose in addition that $\widehat{R} \cong S/(f)$ where $(S, \mathfrak{n})$ is a hypersurface, that is, a regular Noetherian local ring of dimension two and where $0 \neq f \in \mathfrak{n}$. If G is the class group of $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ and $G_{\mathcal{P}}$ is the set of classes containing prime divisors, then $G_{\mathcal{P}} = -G_{\mathcal{P}}$.*

Proof. With the hypotheses given, we can apply [Baeth and Saccon 2012, Proposition 6.2] to see that if M is any indecomposable $\widehat{R}$-module with rank $(r_1, \dots, r_t)$, then there is an indecomposable $\widehat{R}$-module N with rank $(m-r_1, m-r_2, \dots, m-r_t)$ for some $m \geq \max\{r_1, \dots, r_t\}$. Using Construction 4.4 we see that if $\alpha = \begin{bmatrix} a_1 \cdots a_q \end{bmatrix}$ is the column of $\mathsf{A}(R)$ indexed by M, then $-\alpha$ is the column indexed by N. Therefore, since $G_{\mathcal{P}}$ corresponds to the distinct columns of $\mathsf{A}(R)$, $G_{\mathcal{P}} = -G_{\mathcal{P}}$. ☐

Remark 4.10. Although the system of equations developed in Construction 4.4 is somehow natural, it is not the only system of equations which can be used to define $\mathcal{T}(R)$. Indeed, the matrix $\mathsf{A}(R)$ can be adjusted by performing any set of elementary row operations. If J is an elementary matrix corresponding to such a set of row operations, then $\mathcal{T}(R) \cong \ker(\mathsf{A}(R)) \cap \mathbb{N}_0^{(\mathcal{P})} \cong \ker(J\mathsf{A}(R)) \cap \mathbb{N}_0^{(\mathcal{P})}$. Moreover, this isomorphism gives rise to an automorphism of $\mathcal{C}(\mathcal{T}(R))$ mapping the set of classes containing prime divisors to another set of classes containing prime divisors. Example 4.20 illustrates the usefulness of considering an alternate defining matrix for $\mathcal{T}(R)$.

4A. ***Finite representation type.*** *Throughout this subsection, let R be as in Setup (S), and suppose in addition that R has finite representation type.*

Decades of work, going back to [Green and Reiner 1978], and including [Wiegand and Wiegand 1994; Cimen 1998; Arnavut et al. 2007; Baeth 2007], culminated in a precise classification of tuples that can occur as the ranks of indecomposable torsion-free R-modules [Baeth and Luckas 2009]. We note that since R has finite representation type, both R and $\widehat{R}$ have at most three minimal primes (see [Cimen et al. 1995, Theorem 0.5]).

Proposition 4.11 [Baeth and Luckas 2009, Main Theorem 1.2]. (1) *If $\widehat{R}$ is a domain, then every indecomposable finitely generated torsion-free $\widehat{R}$-module has rank* 1, 2 *or* 3.

(2) *If $\widehat{R}$ has exactly two minimal prime ideals, then every indecomposable finitely generated torsion-free $\widehat{R}$-module has rank* $(0,1)$, $(1,0)$, $(1,1)$, $(1,2)$, $(2,1)$ *or* $(2,2)$.

(3) *If $\widehat{R}$ has exactly two minimal prime ideals, then every indecomposable finitely generated torsion-free $\widehat{R}$-module has rank* $(0,0,1)$, $(0,1,0)$, $(1,0,0)$, $(0,1,1)$, $(1,0,1)$, $(1,1,0)$, $(1,1,1)$ *or* $(2,1,1)$.

Note the lack of symmetry in case (3): With a predetermined order on the minimal prime ideals of $\widehat{R}$, there is an indecomposable module of rank $(2,1,1)$, but not of rank $(1,2,1)$ or $(1,1,2)$. As is stated in [Baeth and Luckas 2009, Remark 5.2], even for a fixed number of minimal primes, not each of these tuples will occur as the rank of an indecomposable module for each ring. However, since when applying Construction 4.4 we cannot distinguish between an indecomposable of rank $(2,1)$ and one of rank $(1,0)$, and since all nontrivial tuples of zeros and ones occur as ranks of indecomposable cyclic modules, we have [Baeth and Luckas 2011, Proposition 3.3]:

(1) If $\mathrm{spl}(R)=1$, then $\mathsf{A}(R)=\begin{bmatrix}1 & \cdots & 1 & -1 & \cdots & -1 & 0 & \cdots & 0\end{bmatrix}$.

(2) If $\mathrm{spl}(R)=2$, then $\mathsf{A}(R)=\begin{bmatrix}0 & -1 & 1 & -1 & 1 & 0 & 0 & 1 & \cdots\\ -1 & 0 & 1 & -1 & 0 & 1 & 0 & 1 & \cdots\end{bmatrix}$.

When $\mathrm{spl}(R)=1$, we are guaranteed at least one entry for each of 1, -1, and 0, coming from the ranks of indecomposable cyclic $\widehat{R}$-modules. If we have at most one 1 or at most one -1 in the defining matrix $\mathsf{A}(R)$, then it must be the case that R is a domain, $\widehat{R}$ has exactly two minimal primes $\mathfrak{p}$ and $\mathfrak{q}$, and up to isomorphism either $\widehat{R}/\mathfrak{p}$ is the only indecomposable torsion-free $\widehat{R}$-module of rank (r,s) with $r-s=1$ or the $\widehat{R}/\mathfrak{q}$ is the only indecomposable torsion-free $\widehat{R}$-module of rank (r,s) with $r-s=-1$. If this is the situation, we say that R satisfies condition (†). In case $\mathrm{spl}(R)=2$, we are guaranteed that each column listed appears at least once as a column of $\mathsf{A}(R)$.

We then have the following refinement of Theorem 4.8 when R has finite representation type. The arithmetic of this monoid is studied in Proposition 6.2.2, Theorem 6.4, and Corollary 6.10.

Theorem 4.12 [Baeth and Luckas 2011, Proposition 3.3]. (1) *If* $\mathrm{spl}(R)=1$ *and* R *satisfies condition* (†) *then* $\mathcal{T}(R)\hookrightarrow\mathcal{T}(\widehat{R})$ *is not a divisor theory but* $\mathcal{T}(R)$ *is free abelian.*

(2) *If* $\mathrm{spl}(R)=1$ *and* R *does not satisfy condition* (†), *then* $\mathcal{T}(R)\hookrightarrow\mathcal{T}(\widehat{R})$ *is a divisor theory with infinite cyclic class group* $G=\langle e\rangle$, *and* $G_{\mathcal{P}}=\{-e,0,e\}$.

(3) *If* $\mathrm{spl}(R)=2$, *then* $\mathcal{T}(R)\hookrightarrow\mathcal{T}(\widehat{R})$ *is a divisor theory and* $\mathcal{C}(\mathcal{T}(R))\cong\mathbb{Z}^{(2)}$. *Moreover, this isomorphism maps the set of classes containing prime divisors onto*

$$\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}-1\\-1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}-1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}0\\-1\end{bmatrix}\right\}.$$

4B. *Infinite representation type.* *Throughout this subsection, let* R *be as in Setup* (S), *and suppose in addition that* R *has infinite representation type.*

Unfortunately, in this case, there is no known complete list of the tuples that can occur as ranks of indecomposable finitely generated torsion-free R-modules. Thus we cannot give a full description of $\mathcal{T}(R)$ using Construction 4.4. However, with the additional assumption that $\widehat{R}/\mathfrak{q}$ has infinite representation type for some minimal prime ideal $\mathfrak{q}$ of $\widehat{R}$, we can produce a wide variety of interesting ranks and can provide a partial description of $\mathcal{T}(R)$. This information is enough to show that, very much unlike the finite representation type case of Section 4A, all of the arithmetical invariants we study are infinite.

Proposition 4.13 [Saccon 2010, Theorem 3.4.1]. *Let S be a one-dimensional analytically unramified commutative Noetherian local ring with residue field K, and with t minimal prime ideals $\mathfrak{q}_1, \dots, \mathfrak{q}_t$ such that $S/\mathfrak{q}_1$ has infinite representation type. Let $(r_1, \dots, r_t)$ be a nonzero t-tuple of nonnegative integers with $r_i \le 2r_1$ for all $i \in [2, t]$.*

(1) *There exists an indecomposable torsion-free S-module of rank $(r_1, \dots, r_t)$.*

(2) *If the residue field K is infinite, then the set of isomorphism classes of indecomposable torsion-free S-modules of rank $(r_1, \dots, r_t)$ has cardinality $|K|$.*

By Proposition 4.13 the conditions of Lemma 4.7 are satisfied. Therefore the map $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ is a divisor theory and the class group $\mathcal{C}(\mathcal{T}(R))$ is free abelian of rank $\operatorname{spl}(R)$. Our main result of this subsection is a refinement of Theorem 4.8. Its arithmetical consequences are given in Proposition 6.2.1, strongly improving the arithmetical characterizations given in [Baeth and Saccon 2012].

Theorem 4.14. *Suppose that* $\operatorname{spl}(R) \ge 1$ *and that there is at least one minimal prime ideal $\mathfrak{q}$ of $\widehat{R}$ such that $\widehat{R}/\mathfrak{q}$ has infinite representation type. Then $\mathcal{C}(\mathcal{T}(R))$ is free abelian of rank* $\operatorname{spl}(R)$ *and the set of classes containing prime divisors contains an infinite cyclic subgroup.*

Proof. Let $\mathfrak{p}_1, \dots, \mathfrak{p}_s$ denote the minimal primes of R and, for each $i \in [1, s]$, let $\mathfrak{q}_{i,1}, \dots, \mathfrak{q}_{i,t_i}$ denote the set of minimal primes of $\widehat{R}$ lying over $\mathfrak{p}_i$. Without loss of generality, assume that $\widehat{R}/\mathfrak{q}_{1,1}$ has infinite representation type. If $t_1 = 1$, then without loss of generality, $t_2 > 1$. From Proposition 4.13 there is, for each pair (r, s) of nonnegative integers (not both zero), an indecomposable $\widehat{R}$-module M with $\operatorname{rank}_{\mathfrak{q}_{2,1}}(M) = r$, $\operatorname{rank}_{\mathfrak{q}_{2,2}}(M) = s$, and $\operatorname{rank}_{\mathfrak{q}_{i,j}}(M) = 0$ for all $(i, j) \notin \{(1,1), (2,1), (2,2)\}$. Now suppose that $t_1 > 1$. Then we have, from Proposition 4.13, for each pair (r, s) of nonnegative integers (not both zero) satisfying $r - s \ge -s$, an indecomposable $\widehat{R}$-module M with $\operatorname{rank}_{\mathfrak{q}_{1,1}}(M) = r$, $\operatorname{rank}_{\mathfrak{q}_{1,2}}(M) = s$, and $\operatorname{rank}_{\mathfrak{q}_{i,j}}(M) = 0$ for all $(i, j) \notin \{(1,1), (1,2)\}$. In either case, using Construction 4.4 we see that the set

$$\left\{\begin{bmatrix} x & 0 & \cdots & 0 \end{bmatrix}^T : x \in \mathbb{Z}\right\}$$

occurs as a set of columns for $\mathsf{A}(R)$ and hence occurs as a subset of the set of classes containing prime divisors. $\square$

4C. *Divisor-closed submonoids of $\mathcal{T}(R)$.* Suppose that R has infinite representation type but, in contrast to Theorem 4.14, suppose that $\widehat{R}/\mathfrak{q}$ has finite representation type for each minimal prime $\mathfrak{q}$ of $\widehat{R}$. Then there is no known classification of all ranks of indecomposable finitely generated torsion-free R-modules. Specific rings have been studied in the literature, but even in these settings, a complete solution has been unattainable. We now give such an example which we will return to in Section 4D.

Example 4.15. Let K be an algebraically closed field of characteristic zero. Consider the ring $S = K[\![x, y]\!]/(x^4 - xy^7)$ which has exactly two minimal primes xS and $(x^3 - y^7)S$. Detailed constructions in [Karr and Wiegand 2011; Saccon 2010] show that S has indecomposable modules of ranks (m, m), $(m+1, m)$, and $(m+2, m)$ for each positive integer m. Moreover, [Baeth and Saccon 2012, Proposition 6.2] guarantees indecomposable modules of ranks $(s-(m+1), s-m)$ and $(t-(m+2), t-m)$, where $s \geq m+1$ and $t \geq m+2$ are positive integers. Determining what other tuples occur as ranks of indecomposable torsion-free S-modules appears to be quite difficult.

Thus, since studying $\mathcal{T}(R)$ as a whole is out of reach at the present state of knowledge, we pick finitely many R-modules $M_1, \dots, M_n$, and study the direct-sum relations among them. In more technical terms, instead of studying the full Krull monoid $\mathcal{T}(R)$, we focus on divisor-closed submonoids. Suppose that H is a Krull monoid and $H \hookrightarrow \mathcal{F}(\mathcal{P})$ a cofinal divisor homomorphism. If $H' \subset H$ is a divisor-closed submonoid, then $H' \hookrightarrow H \hookrightarrow \mathcal{F}(\mathcal{P})$ is a divisor homomorphism. For each of the arithmetical invariants $*(\cdot)$ introduced in Section 2, we have $*(H') \leq *(H)$ or $*(H') \subset *(H)$; for example we have $\mathsf{c}(H') \leq \mathsf{c}(H)$, $\mathcal{L}(H') \subset \mathcal{L}(H)$, and so on. Moreover, if H' is the smallest divisor-closed submonoid containing finitely many elements $a_1, \dots, a_k \in H$, it is also the smallest divisor-closed submonoid containing $a_1 \cdot \ldots \cdot a_k$.

For the rest of Section 4, we study divisor-closed submonoids of $\mathcal{T}(R)$ generated by a single R-module M, regardless of whether R has finite or infinite representation type. We denote this monoid by $\operatorname{add}(M)$. Before discussing specific examples in Section 4D, we carefully recall the consequences of our main Construction 4.4 for such submonoids.

Construction 4.16. Let R and $\widehat{R}$ be as in Construction 4.4. Let M be a finitely generated torsion-free R-module. Then $\operatorname{add}(M)$ consists of all isomorphism classes $[N] \in \mathcal{T}(R)$ such that N is isomorphic to a direct summand of $M^{(n)}$ for some finite positive integer n.

Write $\widehat{M} = L_1^{(n_1)} \oplus \cdots \oplus L_k^{(n_k)}$, where the L_i are pairwise nonisomorphic indecomposable finitely generated torsion-free $\widehat{R}$-modules and the n_i are positive integers. If $[N] \in \operatorname{add}(M)$, then $[\widehat{N}] \in \operatorname{add}(\widehat{M})$ and thus, since direct-sum decomposition is essentially unique over $\widehat{R}$,

$$\widehat{N} \cong L_1^{(a_1)} \oplus \cdots \oplus L_k^{(a_k)},$$

with each a_i a nonnegative integer at most n_i. Thus there is a divisor homomorphism $\Psi : \mathrm{add}(M) \to \mathbb{N}_0^{(k)}$ given by $[N] \mapsto (a_1, \ldots, a_k)$. We identify $\mathrm{add}(M)$ with the saturated submonoid $\Gamma(M) = \Psi(\mathrm{add}(M))$ of $\mathbb{N}_0^{(k)}$.

Moreover, if $\mathsf{A}(M)$ is the $\mathrm{spl}(R) \times k$ integer-valued matrix for which the l-th column is the transpose of the row vector

$$\left[r_{1,1} - r_{1,2} \ \cdots \ r_{1,1} - r_{1,t_1} \ \cdots \ r_{s,1} - r_{s,2} \ \cdots \ r_{s,1} - r_{s,t_s} \right],$$

where $r_{i,j} = \mathrm{rank}_{\mathfrak{q}_{i,j}}(V_l)$, then $\mathrm{add}(M) \cong \Gamma(M) = \ker(\mathsf{A}(M)) \cap \mathbb{N}_0^{(k)}$.

We now state a corollary of Theorem 4.8 for $\mathrm{add}(M)$.

Corollary 4.17. *Let M be a finitely generated torsion-free R-module as in Construction 4.16.*

(1) *If* $\mathrm{spl}(R) = 0$, *then* $\mathrm{add}(M) \cong \mathrm{add}(\widehat{M})$ *is free abelian.*

(2) *If* $\mathrm{spl}(R) \geq 1$ *and* $\mathsf{A}(M)$ *satisfies the conditions of Lemma 4.7, then the inclusion* $\Gamma(M) \subset \mathbb{N}_0^{(k)}$ *is a divisor theory. Moreover:*

  (a) $\mathrm{add}(M) \cong \ker(\mathsf{A}(M)) \cap \mathbb{N}_0^{(k)}$.

  (b) $\mathcal{C}(\mathrm{add}(M)) \cong \mathbb{Z}^{(\mathrm{spl}(R))}$, *and this isomorphism maps the set of classes containing prime divisors onto the set of distinct columns of* $\mathsf{A}(M)$.

Before considering explicit examples, we give a realization result (see also [Leuschke and Wiegand 2012, Chapter 1]).

Proposition 4.18. *Let H be a reduced Krull monoid with free abelian class group G of rank q and let $G_{\mathcal{P}} \subset G$ denote the set of classes containing prime divisors. Suppose that $G_{\mathcal{P}}$ is finite and that G has a basis $(e_1, \ldots, e_q)$ such that*

$$G_0 = \{e_0 = e_1 + \cdots + e_q, e_1, \ldots, e_q, -e_0, \ldots, -e_q\} \subset G_{\mathcal{P}}.$$

Then there exists an analytically unramified commutative Noetherian local domain S and a finitely generated torsion-free S-module M such that $\mathrm{add}(M) \cong H$.

Proof. Let $\Phi : G \to \mathbb{Z}^{(q)}$ denote the isomorphism which maps $(e_1, \ldots, e_q)$ onto the standard basis of $\mathbb{Z}^{(q)}$. Let S be an analytically unramified Noetherian local domain with completion $\widehat{S}$ having $q+1$ minimal primes $Q_0, \ldots, Q_q$ such that $\widehat{S}/Q_0$ has infinite representation type. For $\boldsymbol{s} = [s_1 \ \cdots \ s_q] \in \Phi(G_{\mathcal{P}})$, set $r_0 = s_1 + \cdots + s_q$ and $r_i = \sum_{j \neq i} s_j$ for each $i \in [1, q]$. By Proposition 4.13 there exists an indecomposable finitely generated torsion-free $\widehat{S}$-module $N_{\boldsymbol{s}}$ such that $\mathrm{rank}(N_{\boldsymbol{s}}) = (r_0, \ldots, r_q)$. Set

$$N = \bigoplus_{\boldsymbol{s} \in \Phi(G_{\mathcal{P}})} N_{\boldsymbol{s}},$$

and write $\operatorname{rank}(N) = (a_0, \dots, a_q)$. Set $a = \max\{a_0, \dots, a_q\}$ and

$$L = \bigoplus_{i=0}^{q} (\widehat{S}/Q_i)^{(a-a_i)}.$$

Then $N \oplus L$ is a finitely generated torsion-free $\widehat{S}$-module with constant rank and is thus extended from a finitely generated torsion-free S-module M. By Construction 4.16 and Corollary 4.17 we see that $\operatorname{add}(M)$ has class group isomorphic to $\mathbb{Z}^{(q)}$ and this isomorphism maps the set of prime divisors onto the elements of the set $\Phi(G_{\mathcal{P}})$. $\square$

4D. *Examples.* In this section we provide the constructions of naturally occurring monoids $\operatorname{add}(M)$ where M is a finitely generated torsion-free R-module. In particular, we construct specific modules M, whose completion $\widehat{M}$ is often a direct sum of indecomposable cyclic $\widehat{R}$-modules and we determine the class group G of $\operatorname{add}(M)$ and the set of classes $G_{\mathcal{P}} \subset G$ containing prime divisors. Note that the Krull monoids $\mathcal{M}(R)$ of *all* finitely generated R-modules and $\mathcal{T}(R)$ of *all* finitely generated torsion-free R-modules have class groups $G' \supset G$ and a set $G'_{\mathcal{P}}$ of classes containing prime divisors such that $G'_{\mathcal{P}} \supset G_{\mathcal{P}}$. Since $\operatorname{add}(M)$ is a divisor-closed submonoid of both $\mathcal{M}(R)$ and of $\mathcal{T}(R)$, a study of the arithmetic of $\operatorname{add}(M)$ provides a partial description of $\mathcal{M}(R)$ and $\mathcal{T}(R)$. Moreover, the values of arithmetical invariants of $\operatorname{add}(M)$ give lower bounds on the same arithmetical invariants of $\mathcal{M}(R)$ and $\mathcal{T}(R)$.

In each of the following examples we construct an $\widehat{R}$-module $L = L_1^{n_1} \oplus \cdots \oplus L_k^{n_k}$ of constant rank, where $L_1, \dots, L_k$ are pairwise nonisomorphic indecomposable $\widehat{R}$-modules. Then, by Corollary 4.17 with $\widehat{M} \cong L$ for some R-module M,

$$\operatorname{add}(M) \cong \ker\big(\mathsf{A}(M)\big) \cap \mathbb{N}_0^{(k)} \subset \mathbb{N}_0^{(k)} \cong \operatorname{add}(L).$$

In particular, we do so in such a way that the natural map $\operatorname{add}(M) \hookrightarrow \operatorname{add}(L)$ is a divisor theory with class group isomorphic to $\mathbb{Z}^{(\operatorname{spl}(R))}$ and where the set of classes containing prime divisors maps onto the distinct columns of $\mathsf{A}(M)$.

Example 4.19. We now construct a monoid of modules whose arithmetic will be studied in Proposition 6.12. Let S be as in Example 4.15. Then there are indecomposable torsion-free S-modules M_1, M_{-1}, M_2, M_{-2}, N_1, N_{-1}, N_2, and N_{-2} with ranks (respectively) $(2,1)$, $(1,2)$, $(3,1)$, $(1,3)$, $(3,2)$, $(3,2)$, $(2,3)$, $(4,2)$, and $(2,4)$. Set L to be the direct sum of these eight indecomposable S-modules. By Lech's theorem [1986], there exists a Noetherian local domain $(R, \mathfrak{m})$ with $\mathfrak{m}$-adic completion $\widehat{R} \cong S$. Since L has constant rank, L is extended from some R-module M, and $\operatorname{add}(M) \hookrightarrow \operatorname{add}(L) \cong \mathbb{N}_0^{(8)}$ is a divisor theory with infinite cyclic class group G and with $G_{\mathcal{P}} = \{-2e, -e, e, 2e\}$ where $G = \langle e \rangle$.

Example 4.20. We now provide an example that illustrates the convenience of choosing an alternate defining matrix for $\operatorname{add}(M)$, as is described in Remark 4.10. Its arithmetic is given in Theorem 6.4. Suppose that R has two minimal prime ideals $\mathfrak{p}_1, \mathfrak{p}_2$ and that $\widehat{R}$ has five minimal prime ideals $\mathfrak{q}_{(1,1)}, \mathfrak{q}_{(1,2)}, \mathfrak{q}_{(1,3)}, \mathfrak{q}_{(2,1)}$, and $\mathfrak{q}_{(2,2)}$, with $\mathfrak{q}_{(i,j)}$ lying over $\mathfrak{p}_i$ for each $i \in [1,2]$ and for each j. Set

$$L = \frac{\widehat{R}}{\mathfrak{q}_{(1,1)} \cap \mathfrak{q}_{(1,3)} \cap \mathfrak{q}_{(2,2)}} \oplus \frac{\widehat{R}}{\mathfrak{q}_{(1,1)} \cap \mathfrak{q}_{(1,2)} \cap \mathfrak{q}_{(2,1)}} \oplus \frac{\widehat{R}}{\mathfrak{q}_{(1,2)} \cap \mathfrak{q}_{(2,1)}} \oplus \frac{\widehat{R}}{\mathfrak{q}_{(1,1)}} \oplus \frac{\widehat{R}}{\mathfrak{q}_{(1,3)} \cap \mathfrak{q}_{(2,2)}} \oplus \frac{\widehat{R}}{\mathfrak{q}_{(1,2)} \cap \mathfrak{q}_{(1,3)} \cap \mathfrak{q}_{(2,1)} \cap \mathfrak{q}_{(2,2)}}.$$

Since L has constant rank 3, there is an R-module M such that $\widehat{M} \cong L$. Then $\operatorname{add}(M) \cong \ker(\mathsf{A}(M)) \cap \mathbb{N}_0^{(6)}$ where

$$\mathsf{A}(M) = \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & -1 & -1 \\ -1 & 1 & 1 & 0 & -1 & 0 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & -1 \\ -1 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = J\mathsf{A}(M).$$

Thus

$$\operatorname{add}(M) \cong \ker(J\mathsf{A}(M)) \cap \mathbb{N}_0^{(6)} \cong \ker \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & -1 & -1 \end{bmatrix} \cap \mathbb{N}_0^{(6)}.$$

Since the matrix $\mathsf{A}(M)$ has rank two, the representation of $\operatorname{add}(M)$ as a Diophantine matrix defined by two equations more clearly describes this monoid. Moreover, since the map from $\mathbb{Z}^{(6)}$ to $\mathbb{Z}^{(2)}$ is surjective (the map $\mathsf{A}(M) : \mathbb{Z}^{(6)} \to \mathbb{Z}^{(3)}$ is not surjective), we immediately see that $\mathcal{C}(\operatorname{add}(M)) \cong \mathbb{Z}^{(2)}$, and this isomorphism maps the set of classes containing prime divisors onto

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}.$$

Example 4.21. We now consider a monoid $\operatorname{add}(M)$ which generalizes the monoid $\mathcal{T}(R)$ when R has finite representation type, and its arithmetic is studied in Theorem 6.7 and Corollary 6.10. Suppose that $\widehat{R}$ has $q+1$ minimal primes $\mathfrak{q}_1, \ldots, \mathfrak{q}_{q+1}$, and set

$$L = \bigoplus_{\substack{I \subset [1,q+1] \\ I \neq \varnothing}} \frac{\widehat{R}}{\bigcap_{i \in I} \mathfrak{q}_i}.$$

From the symmetry of the set of ranks of the indecomposable cyclic $\widehat{R}$-modules

$$\frac{\widehat{R}}{\bigcap_{i\in I} \mathfrak{q}_i}$$

we immediately see that L has constant rank $(2^q,\dots,2^q)$ and is therefore extended from some R-module M. Then

$$\operatorname{add}(M)\cong \ker(\mathsf{A}(M))\cap \mathbb{N}_0^{(q)},$$

where $\mathsf{A}(M)$ is an $q\times 2^{q+1}-1$ integer-valued matrix with columns $[\epsilon_1 \ \cdots \ \epsilon_q]^T$ where either $\epsilon_i\in\{0,1\}$ for all $i\in[1,q]$ or $\epsilon_i\in\{0,-1\}$ for all $i\in[1,q]$.

Since the columns of $\mathsf{A}(M)$ contain a basis for $\mathbb{Z}^{(q)}$, $\operatorname{add}(M)\hookrightarrow \operatorname{add}(L)\cong \mathbb{N}_0^{(q)}$ is a divisor theory with class group $\mathcal{C}(\operatorname{add}(M))\cong \mathbb{Z}^{(q)}$, and this isomorphism maps the set of classes containing prime divisors onto

$$\Big(\big\{[\epsilon_1 \ \cdots \ \epsilon_q]^T : \epsilon_i\in\{0,1\}\big\}\cup\big\{[\epsilon_1 \ \cdots \ \epsilon_q]^T : \epsilon_i\in\{0,-1\}\big\}\Big)\setminus\big\{[0 \ \cdots \ 0]\big\}.$$

Example 4.22. In this example we construct a monoid $\operatorname{add}(M)$ which generalizes the monoid of Example 4.21 by including all vectors having entries in $\{-1,0,1\}$ in the set $G_{\mathcal{P}}$. This larger set of classes containing prime divisors adds much complexity to the arithmetic. Suppose that R has q minimal primes and that $\widehat{R}$ has $2q$ minimal primes

$$\mathfrak{q}_{(1,1)},\mathfrak{q}_{(1,2)},\mathfrak{q}_{(2,1)},\dots,\mathfrak{q}_{(q,2)},$$

where $\mathfrak{q}_{(i,j)}\cap R=\mathfrak{q}_{(i',j')}\cap R$ if and only if $i=i'$. As in the previous example, let

$$L=\bigoplus_{\substack{I\subset\{(1,1),\dots,(q,2)\}\\ I\neq\varnothing}} \frac{\widehat{R}}{\bigcap_{(i,j)\in I}\mathfrak{q}_{(i,j)}}.$$

From the symmetry of the set of ranks of the indecomposable cyclic $\widehat{R}$-modules $\widehat{R}/\bigcap_{\{i,j\}\in I}\mathfrak{q}_{i,j}$, we immediately see that L has constant rank $(2^{2q-1},\dots,2^{2q-1})$ and is therefore extended from some R-module M. Then

$$\operatorname{add}(M)\cong \ker(\mathsf{A}(M))\cap \mathbb{N}_0^{(q)},$$

where $\mathsf{A}(M)$ is an $q\times 2^{2q}-1$ integer-valued matrix with columns of the form

$$\big[r_{(1,1)}-r_{(1,2)} \quad r_{(2,1)}-r_{(2,2)} \quad \cdots \quad r_{(q,1)}-r_{(q,2)}\big]^T,$$

where $(r_{(1,1)},r_{(1,2)},\dots,r_{(q,2)})$ is the rank of one of the $2^{2q}-1$ indecomposable cyclic $\widehat{R}$-modules — that is, any one of the q-tuples of 1s and 0s (not all 0). In other words, the columns of $\mathsf{A}(M)$ are exactly the 3^q columns $[\epsilon_1 \ \cdots \ \epsilon_q]^T$, where

$\epsilon_i \in \{-1, 0, 1\}$ for all $i \in [1, q]$, repeated with some multiplicity. For example, the column of all zeros occurs for each of the indecomposable cyclic $\widehat{R}$-modules

$$\frac{\widehat{R}}{\bigcap_{(i,j)\in I} \mathfrak{q}_{(i,j)}},$$

where $(i, 1) \in I$ if and only if $(i, 2) \in I$.

Since the columns of $\mathsf{A}(M)$ contain a basis for $\mathbb{Z}^{(q)}$,

$$\operatorname{add}(M) \hookrightarrow \operatorname{add}(L) \cong \mathbb{N}_0^{(2^{2q}-1)}$$

is a divisor theory whose class group $G \cong \mathbb{Z}^{(q)}$, and this isomorphism maps the set of classes containing prime divisors onto

$$\bigl\{[\epsilon_1 \ \cdots \ \epsilon_q]^T : \epsilon_i \in \{-1, 0, 1\}\bigr\}.$$

Example 4.23. In this example we consider $\operatorname{add}(M)$ when the completion of M is isomorphic to a direct sum of some (but not all) of the indecomposable cyclic $\widehat{R}$-modules. In this case, the example is constructed in such a way that $\mathcal{B}(G_{\mathcal{P}})$ is a direct product of nontrivial submonoids (see Lemma 3.2). Suppose that R has q minimal primes and that $\widehat{R}$ has $3q$ minimal primes

$$\bigl\{\mathfrak{q}_{(i,j)} : i \in [1, q],\ j \in [1, 3]\bigr\},$$

where $\mathfrak{q}_{(i,j)} \cap R = \mathfrak{q}_{(i',j')} \cap R$ if and only if $i = i'$. Let L be the $\widehat{R}$-module

$$\bigoplus_{i=1}^{q}\Bigl(\widehat{R}/\mathfrak{q}_{(i,1)} \oplus \widehat{R}/\mathfrak{q}_{(i,2)} \oplus \widehat{R}/\mathfrak{q}_{(i,3)} \oplus \widehat{R}/(\mathfrak{q}_{(i,1)} \cap \mathfrak{q}_{(i,2)}) \oplus \widehat{R}/(\mathfrak{q}_{(i,1)} \cap \mathfrak{q}_{(i,3)}) \oplus \widehat{R}/(\mathfrak{q}_{(i,2)} \cap \mathfrak{q}_{(i,3)})\Bigr).$$

We see immediately that L has constant rank $(3, \ldots, 3)$ and thus L is extended from some R-module M. Then $\operatorname{add}(M) \cong \ker(\mathsf{A}(M)) \cap \mathbb{N}_0^{(2q)}$ where $\mathsf{A}(M)$ is an $2q \times 6q$ integer-valued matrix with columns

$$\bigl\{e_{2k-1}, e_{2k}, e_{2k-1} + e_{2k}, -e_{2k-1}, -e_{2k}, -e_{2k-1} - e_{2k} : k \in [1, q]\bigr\},$$

where $(e_1, \ldots, e_{2q})$ denotes the canonical basis of $\mathbb{Z}^{(2q)}$.

For $k \in [1, q]$, we set

$$G_k = \bigl\{e_{2k-1}, e_{2k}, e_{2k-1} + e_{2k}, -e_{2k-1}, -e_{2k}, -e_{2k-1} - e_{2k}\bigr\}.$$

Then $G_{\mathcal{P}} = \biguplus_{k\in[1,q]} G_k$ is the set of classes containing prime divisors and $\mathcal{B}(G_{\mathcal{P}}) = \mathcal{B}(G_1) \times \cdots \times \mathcal{B}(G_q)$. From Proposition 6.1 we will see that $\mathcal{B}(G_k) \hookrightarrow \mathcal{F}(G_k)$

is a divisor theory, whence $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow \mathcal{F}(G_{\mathcal{P}})$ and $\mathrm{add}(M) \hookrightarrow \mathrm{add}(L)$ are divisor theories. The arithmetic of this monoid is studied in Proposition 6.12 and Corollary 6.15.

Example 4.24. As in Example 4.23, suppose that R has q minimal primes and suppose that the completion $\widehat{R}$ of R has $3q$ minimal primes

$$\{\mathfrak{q}_{(i,j)} : i \in [1,q], j \in [1,3]\},$$

where $\mathfrak{q}_{(i,j)} \cap R = \mathfrak{q}_{(i',j')} \cap R$ if and only if $i = i'$. Further suppose that $\widehat{R} = S/(f)$ where $(S, \mathfrak{n})$ is a regular Noetherian local ring of dimension two and where $0 \neq f \in \mathfrak{n}$ and that $\widehat{R}/\mathfrak{q}_{(i,j)}$ has infinite representation type for all pairs (i, j). By Proposition 4.13, for each $k \in [1,q]$ there are indecomposable finitely generated torsion-free $\widehat{R}$-modules M_k and N_k of ranks $(r_{1,1}, \ldots, r_{q,3})$ and $(s_{1,1}, \ldots, s_{q,3})$ where

$$r_{i,j} = \begin{cases} 0 & \text{if } i \neq k, \\ 2 & \text{if } i = k, j \in [1,2], \\ 0 & \text{if } i = k, j = 3, \end{cases} \quad \text{and} \quad s_{i,j} = \begin{cases} 0 & \text{if } i \neq k, \\ 3 & \text{if } i = k, j = 1, \\ 2 & \text{if } i = k, j = 2, \\ 0 & \text{if } i = k, j = 3. \end{cases}$$

Moreover, by Corollary 4.9, for each $k \in [1,q]$ there are constant $t_k \geq 2$ and $t'_k \geq 3$ and indecomposable finitely generated torsion-free $\widehat{R}$-modules M'_k and N'_k having ranks $(r'_{1,1}, \ldots, r'_{q,3})$ and $(s'_{1,1}, \ldots, s'_{q,3})$ where

$$r'_{i,j} = \begin{cases} t_k & \text{if } i \neq k, \\ t_k - 2 & \text{if } i = k, j \in [1,2], \\ t_k & \text{if } i = k, j = 3, \end{cases} \quad \text{and} \quad s'_{i,j} = \begin{cases} t'_k & \text{if } i \neq k, \\ t'_k - 3 & \text{if } i = k, j = 1, \\ t'_k - 2 & \text{if } i = k, j = 2, \\ t'_k & \text{if } i = k, j = 3. \end{cases}$$

Let

$$L = \widehat{R} \oplus \Bigg(\bigoplus_{k=1}^{q} (M_k \oplus N_k \oplus M'_k \oplus N'_k) \\ \oplus \widehat{R}/(\mathfrak{q}_{(i,1)} \cap \mathfrak{q}_{(i,3)}) \oplus \widehat{R}/(\mathfrak{q}_{(i,1)} \cap \mathfrak{q}_{(i,2)}) \oplus \widehat{R}/\mathfrak{q}_{(i,2)} \oplus \widehat{R}/\mathfrak{q}_{(i,3)} \Bigg).$$

Since L has constant rank, L is extended from an R-module M. Then $\mathrm{add}(M) \cong \ker(\mathsf{A}(M)) \cap \mathbb{N}_0^{(2q)}$ where $\mathsf{A}(M)$ is an $2q \times (8q+1)$ integer-valued matrix with columns 0 and

$$\{e_{2k-1}, e_{2k}, 2e_{2k}, e_{2k-1} + 2e_{2k}, -e_{2k-1}, -e_{2k}, -2e_{2k}, -e_{2k-1} - 2e_{2k} : k \in [1,q]\},$$

where $(e_1, \ldots, e_{2q})$ denotes the canonical basis of $\mathbb{Z}^{(2q)}$.

For $k \in [1, q]$, we set

$$G_k = \{e_{2k-1}, e_{2k}, 2e_{2k}, e_{2k-1} + 2e_{2k}, -e_{2k-1}, -e_{2k}, -2e_{2k}, -e_{2k-1} - 2e_{2k}\}.$$

Then $G_{\mathcal{P}} = \biguplus_{k \in [1,q]} G_k$ is the set of classes containing prime divisors and $\mathcal{B}(G_{\mathcal{P}}) = \mathcal{B}(G_1) \times \cdots \times \mathcal{B}(G_q)$. From Proposition 6.1 we will see that $\mathcal{B}(G_k) \hookrightarrow \mathcal{F}(G_k)$ is a divisor theory, whence $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow \mathcal{F}(G_{\mathcal{P}})$ and $\operatorname{add}(M) \hookrightarrow \operatorname{add}(N)$ are divisor theories. The arithmetic of this monoid is studied in Proposition 6.13 and Corollary 6.15.

Example 4.25. In our final example we construct a tuple $(G, G_{\mathcal{P}})$ which generalizes the monoid $\mathcal{T}(R)$ when R has finite representation type (see Theorem 4.12). The arithmetic of such Krull monoids is studied in Theorem 6.4 and Corollary 6.10. Suppose that $\widehat{R}$ has $q+1$ minimal primes $\mathfrak{q}_1, \dots, \mathfrak{q}_{q+1}$, and set

$$L = \bigoplus_{j=1}^{q+1} \Big(\big(\widehat{R} / \bigcap_{i \neq j} \mathfrak{q}_i \big) \oplus \big(\widehat{R} / \mathfrak{q}_j \big) \Big).$$

Note that L has constant rank $(q, \dots, q)$ and is hence extended from some R-module M. Then $\operatorname{add}(M) \cong \ker(\mathsf{A}(M)) \cap \mathbb{N}_0^{(q)}$ where $\mathsf{A}(M)$ is an $q \times 2q$ integer-valued matrix with columns

$$e_1, \dots, e_q, e_0 = e_1 + e_2 + \cdots + e_q, -e_1, \dots, -e_q, -e_0.$$

By Proposition 6.1, $\operatorname{add}(M) \hookrightarrow \operatorname{add}(L) \cong \mathbb{N}_0^{(2q)}$ is a divisor theory with class group $\mathcal{C}(\operatorname{add}(M)) \cong \mathbb{Z}^{(q)}$, and this isomorphism maps the set of classes containing prime divisors onto

$$\{e_0 = e_1 + \cdots + e_q, e_1, \dots, e_q, -e_0, \dots, -e_q\}.$$

5. Monoids of modules: class groups and distribution of prime divisors, II

In this section we investigate the characteristic of the Krull monoids $\mathcal{T}(R)$ and $\mathfrak{C}(R)$ for two-dimensional Noetherian local Krull domains (see Theorem 5.4). We will show that, apart from a well-described exceptional case, their class groups are both isomorphic to the factor group $\mathcal{C}(\widehat{R})/\iota(\mathcal{C}(R))$, where $\iota : \mathcal{C}(R) \to \mathcal{C}(\widehat{R})$ is the natural homomorphism between the class groups of R and $\widehat{R}$. In a well-studied special case where R is factorial and $\widehat{R}$ is a hypersurface with finite representation type, this factor group is a finite cyclic group (see Theorem 5.5). This is in strong contrast to the results on one-dimensional rings in the previous section where all class groups are torsion-free.

Let S be a Krull domain and let $\mathcal{I}_v^*(S)$ denote the monoid of nonzero divisorial ideals. Then $\varphi : S \to \mathcal{I}_v^*(S)$, defined by $a \mapsto aS$, is a divisor theory. In this section we view $\mathcal{C}(S)$ as the class group of this specific divisor theory. First we give a classical result (see [Bourbaki 1988, Chapter VII, Section 4.7]).

Lemma 5.1. *Let S be a Noetherian Krull domain. One can associate to each finitely generated S-module M a class $\boldsymbol{c}(M) \in \mathcal{C}(S)$ in such a way that*

(1) *If $0 \to M' \to M \to M'' \to 0$ is an exact sequence of finitely generated S-modules, then $\boldsymbol{c}(M) = \boldsymbol{c}(M') + \boldsymbol{c}(M'')$.*

(2) *If I is a fractional ideal of S and I_v the divisorial ideal generated by I, then $\boldsymbol{c}(I) = \boldsymbol{c}(I_v)$.*

Note that if S is any Noetherian domain, every ideal of S is obviously an indecomposable finitely generated torsion-free S-module. If, in addition, the ring has dimension two, then we have the following stronger result.

Lemma 5.2. *Let S be a Noetherian local Krull domain of dimension two.*

(1) *Every divisorial ideal of S is an indecomposable MCM S-module* [Evans and Griffith 1985, Lemma 1.1 and Theorem 3.6].

(2) *In addition, assume that the $\mathfrak{m}$-adic completion $\widehat{S}$ of S is a Krull domain. Then a finitely generated torsion-free $\widehat{S}$-module N is extended from an S-module if and only if $\boldsymbol{c}(N)$ is in the image of the natural homomorphism $\iota : \mathcal{C}(S) \to \mathcal{C}(\widehat{S})$* [Rotthaus et al. 1999, Proposition 3].

We now give a result on abstract Krull monoids which encapsulates the structure of the monoids of modules described in Theorem 5.4.

Lemma 5.3. *Let $D = \mathcal{F}(\mathcal{P})$ be a free abelian monoid, G an additive abelian group, $\psi : D \to G$ a homomorphism, and $H = \psi^{-1}(0) \subset D$.*

(1) *If $H \subset D$ is cofinal, then the inclusion $H \hookrightarrow D$ is a divisor homomorphism and $\overline{\psi} : D/H \to \psi(D) \subset G$ given by $\overline{\psi}([a]) = \psi(a)$ is an isomorphism.*

(2) *The inclusion $H \hookrightarrow D$ is a divisor theory if and only if $\langle \psi(\mathcal{P}) \rangle = [\psi(\mathcal{P} \setminus \{q\})]$ for every $q \in \mathcal{P}$. If this is the case, then $\overline{\psi} : D/H \to \psi(D)$ is an isomorphism and, for every $g \in \psi(D)$, the set $\mathcal{P} \cap \psi^{-1}(g)$ is the set of prime divisors in the class $\overline{\psi}^{-1}(g)$.*

(3) *If the restriction $\psi \mid_{\mathcal{P}} : \mathcal{P} \to G$ of ψ to $\mathcal{P}$ is an epimorphism, then $H \hookrightarrow D$ is cofinal. Moreover, it is a divisor theory apart from the following exception:*

$$G = \{0, g\} \quad \textit{and} \quad |\mathcal{P} \cap \psi^{-1}(g)| = 1.$$

If $H \hookrightarrow D$ is not a divisor theory, then H is factorial.

Proof. For the proofs of (1) and (2), see [Geroldinger and Halter-Koch 2006, Proposition 2.5.1]. We now consider the proof of (3).

Let $a \in D$. Since $\psi \mid_{\mathcal{P}} : \mathcal{P} \to G$ is an epimorphism, there exists $p \in \mathcal{P} \subset D$ such that $\psi(p) = -\psi(a)$. Therefore $ap \in H$ and the inclusion $H \hookrightarrow D$ is cofinal. In order to show that $H \hookrightarrow D$ is a divisor theory we distinguish three cases. First suppose that $|G| = 1$. Then $|D/H| = 1$, and hence $H = D$. Next suppose that $|G| > 2$. By (2) we must verify that

$$\psi(q) \in [\psi(\mathcal{P} \setminus \{q\})] \quad \text{for every } q \in \mathcal{P}.$$

Let $q \in \mathcal{P}$. Since $|G| > 2$, there exist $g_1, g_2 \in G \setminus \{0, \psi(q)\}$ with $\psi(q) = g_1 + g_2$. Since the restriction $\psi \mid_{\mathcal{P}} : \mathcal{P} \to G$ is an epimorphism, there exist $p_1, p_2 \in \mathcal{P} \setminus \{q\}$ with $\psi(p_i) = g_i$ for $i \in [1, 2]$. Therefore

$$\psi(q) = g = g_1 + g_2 = \psi(p_1) + \psi(p_2) \in [\psi(\mathcal{P} \setminus \{q\})].$$

Finally, suppose that $|G| = 2$. Then $H \hookrightarrow D$ is a divisor theory if and only if $\langle \psi(\mathcal{P}) \rangle = [\psi(\mathcal{P} \setminus \{q\})]$ for every $q \in \mathcal{P}$ if and only if there exist distinct $q_1, q_2 \in \mathcal{P}$ such that $\psi(q_i) \neq 0$ for $i \in [1, 2]$. Clearly, if $q \in \mathcal{P}$ is the unique element of $\mathcal{P}$ with $\psi(q) \neq 0$, then H is free abelian with basis $\mathcal{P} \setminus \{q\} \cup \{q^2\}$. $\square$

We are now able to determine both the class group and the set of classes containing prime divisors for the monoids $\mathcal{T}(R)$ and $\mathfrak{C}(R)$. This generalizes and refines the results of [Baeth 2009]. Since each divisorial ideal over a two-dimensional local ring is MCM and thus finitely generated and torsion-free, Theorem 5.4 can be stated in parallel both for $\mathcal{T}(R)$ and $\mathfrak{C}(R)$.

Theorem 5.4. *Let $(R, \mathfrak{m})$ be a Noetherian local Krull domain of dimension two whose $\mathfrak{m}$-adic completion $\widehat{R}$ is also a Krull domain. Let $\mathcal{V}(R)$ (respectively $\mathcal{V}(\widehat{R})$) denote either $\mathcal{T}(R)$ (respectively $\mathcal{T}(\widehat{R})$) or $\mathfrak{C}(R)$ (respectively $\mathfrak{C}(\widehat{R})$), and let $\iota : \mathcal{C}(R) \to \mathcal{C}(\widehat{R})$ be the natural map.*

(1) *The embedding $\mathcal{V}(R) \hookrightarrow \mathcal{V}(\widehat{R})$ is a cofinal divisor homomorphism. The class group of this divisor homomorphism is isomorphic to $G = \mathcal{C}(\widehat{R})/\iota(\mathcal{C}(R))$ and every class contains a prime divisor. Moreover the embedding is a divisor theory except if $\widehat{R}$ satisfies the following condition*:

   (E) *$|G| = 2$ and, up to isomorphism, there is precisely one nonextended indecomposable $\widehat{R}$-module M with $[M] \in \mathcal{V}(\widehat{R})$.*

   In particular, $\mathcal{V}(R)$ satisfies KRSA if and only if either $|G| = 1$ or $\widehat{R}$ satisfies (E).

(2) *Suppose that the embeddings $\mathcal{T}(R) \hookrightarrow \mathcal{T}(\widehat{R})$ and $\mathfrak{C}(R) \hookrightarrow \mathfrak{C}(\widehat{R})$ are both divisor theories. Then their class groups are isomorphic. If $(G, (m_g)_{g \in G})$ is the characteristic of $\mathcal{T}(R)$ and $(G, (n_g)_{g \in G})$ is the characteristic of $\mathfrak{C}(R)$, then $m_g \geq n_g$ for all $g \in G$. Moreover, $\sum_{g \in G} m_g$ infinite.*

(A_n)	$k[\![x,y,z]\!]/(x^2+y^2+z^{n+1})$,	$n \geq 1$
(D_n)	$k[\![x,y,z]\!]/(x^2z+y^2+z^{n-1})$,	$n \geq 4$
(E_6)	$k[\![x,y,z]\!]/(x^3+y^2+z^4)$	
(E_7)	$k[\![x,y,z]\!]/(x^3+xz^3+y^2)$	
(E_8)	$k[\![x,y,z]\!]/(x^2+y^3+z^5)$	

Table 1. Two-dimensional rings with finite representation type.

Proof. We set $D = \mathcal{F}(\mathcal{P}) = \mathcal{V}(\widehat{R})$. By (1) of Lemma 5.1 there is a homomorphism

$$\psi : \mathcal{V}(\widehat{R}) \to \mathcal{C}(\widehat{R})/\iota(\mathcal{C}(R)) = G \quad \text{given by } [M] \mapsto \boldsymbol{c}(M) + \iota(\mathcal{C}(R)).$$

By (1) of Lemma 5.2, every divisorial ideal of $\widehat{R}$ is an indecomposable MCM $\widehat{R}$-module. That is, the class of each divisorial ideal of $\widehat{R}$ in $\mathcal{C}(\widehat{R})$ is the image of some $[M] \in \mathcal{V}(\widehat{R})$, where M is an indecomposable MCM $\widehat{R}$-module. In other words, ψ restricted to $\mathcal{P} = \mathcal{A}(D)$ is an epimorphism and $\psi(\mathcal{A}(D)) = \mathcal{C}(\widehat{R})/\iota(\mathcal{C}(R))$.

By Lemma 5.3, the inclusion $H = \psi^{-1}(0) \subset D$ is a cofinal divisor homomorphism. By (2) of Lemma 5.2, H is the image of the embedding $\mathcal{V}(R) \hookrightarrow \mathcal{V}(\widehat{R})$, and thus the embedding $\mathcal{V}(R) \hookrightarrow \mathcal{V}(\widehat{R})$ is a divisor theory if and only if the inclusion $H \hookrightarrow D$ is a divisor theory. By Lemma 5.3 this always holds apart from the described Exception (E). A Krull monoid is factorial if and only if its class group is trivial. Thus, if $\mathcal{V}(R) \hookrightarrow \mathcal{V}(\widehat{R})$ is a divisor theory then KRSA holds for $\mathcal{V}(R)$ if and only if $|G| = 0$. If $\mathcal{V}(R) \hookrightarrow \mathcal{V}(\widehat{R})$ is not a divisor theory, then the inclusion $H \hookrightarrow D$ is not a divisor theory. By Lemma 5.3, H is factorial, whence $\mathcal{V}(R)$ is factorial.

Since each MCM R-module is finitely generated and torsion-free, it is clear that $m_g \geq n_g$ for each $g \in G$. From [Bass 1962] we know that there infinitely many nonisomorphic indecomposable finitely generated torsion-free $\widehat{R}$-modules, and therefore $\sum_{g \in G} m_g$ infinite. □

Let R be as in the above theorem and assume in addition that

- R contains a field and $k = R/\mathfrak{m}$ is algebraically closed with characteristic zero;
- $\widehat{R}$ is a hypersurface, that is, $\widehat{R}$ is isomorphic to a three-dimensional regular Noetherian local ring modulo a regular element;
- R has finite representation type.

Such rings were classified in [Buchweitz et al. 1987; Knörrer 1987] and are given, up to isomorphism, in Table 1. Note that since $\widehat{R}$ has finite representation type and each divisorial ideal of $\widehat{R}$ is and indecomposable MCM $\widehat{R}$-module, $\mathcal{C}(\widehat{R})$ and hence $\mathfrak{C}(R)$ is finite.

An amazing theorem of Heitmann [1993] gives the existence of a local factorial domain whose completion is a ring as in Table 1. In this situation we can determine the characteristic of $\mathfrak{C}(R)$.

Theorem 5.5. *Let* $(R, \mathfrak{m})$ *be a Noetherian local factorial domain with* $\mathfrak{m}$*-adic completion* $\widehat{R}$ *isomorphic to a ring in Table 1.*

(1) *If* $\widehat{R}$ *is a ring of type* (A_n), *then* $\mathcal{C}(\mathfrak{C}(R))$ *is cyclic of order* $n+1$ *and each class contains exactly one prime divisor.*

(2) *Suppose* $\widehat{R}$ *is a ring of type* (D_n).

  (a) *If* n *is even, then* $\mathcal{C}(\mathfrak{C}(R)) \cong C_2 \oplus C_2$. *The trivial class contains* $n/2$ *prime divisors. Two nontrivial classes each contain a single prime divisor and their sum contains* $(n-2)/2$ *prime divisors.*

  (b) *If* n *is odd, then* $\mathcal{C}(\mathfrak{C}(R))$ *is cyclic of order four. The classes of order four each contain a single prime. The remaining classes each contain* $(n-1)/2$ *prime divisors.*

(3) *If* $\widehat{R}$ *is a ring of type* (E_6), *then* $\mathcal{C}(\mathfrak{C}(R))$ *is cyclic of order three. The trivial class contains three prime divisors, while each remaining class contains two prime divisors.*

(4) *If* $\widehat{R}$ *is a ring of type* (E_7), *then* $\mathcal{C}(\mathfrak{C}(R))$ *is cyclic of order two. The trivial class contains five prime divisors and the nontrivial class contains three prime divisors.*

(5) *If* $\widehat{R}$ *is a ring of type* (E_8), *then* $\mathcal{C}(\mathfrak{C}(R))$ *is trivial, with the trivial class containing all nine prime divisors.*

Proof. The class groups $\mathcal{C}(\widehat{R})$ for $\widehat{R}$ a ring listed in Table 1 were given in [Brieskorn 1967–1968]. Since R is factorial, $\mathcal{C}(R) = 0$ and by Theorem 5.4, $\mathcal{C}(\mathfrak{C}(R)) \cong \mathcal{C}(\widehat{R})$. Following the proof of [Baeth 2009, Theorem 4.3] one can compute the class of each indecomposable $\widehat{R}$-module in $\mathcal{C}(\widehat{R})$ by using the Auslander–Reiten sequence for $\widehat{R}$. The result follows by considering the map $\overline{\psi}$ defined in the proof of Theorem 5.4. □

The above theorem completely determines the characteristic of the monoid $\mathfrak{C}(R)$. With this information, in addition to being able to completely describe the arithmetic of $\mathfrak{C}(R)$ as we do in Theorem 6.8, we can easily enumerate the atoms of $\mathfrak{C}(R)$ (the nonisomorphic indecomposable MCM modules). We now illustrate this ability with an example. If $\boldsymbol{\beta} : \mathfrak{C}(R) \to \mathcal{B}(G_{\mathcal{P}})$ is the transfer homomorphism of Lemma 3.4, then $\mathcal{A}(\mathfrak{C}(R)) = \boldsymbol{\beta}^{-1}\big(\mathcal{A}(\mathcal{B}(G_{\mathcal{P}}))\big)$. Suppose that $\widehat{R}$ is a ring of type (D_n) with n even. Then $\widehat{R}$ has exactly $n+1$ nonisomorphic indecomposable MCM modules. If $C_2 \oplus C_2 = \{0, e_1, e_2, e_1 + e_2\}$, then

$$\mathcal{A}(C_2 \oplus C_2) = \{0, e_1^2, e_2^2, (e_1+e_2)^2, e_1e_2(e_1+e_2)\},$$

and hence R has exactly

$$|\mathcal{A}(\mathfrak{C}(R))| = \frac{n}{2} + 1\cdot 1 + 1\cdot 1 + \frac{n-2}{2}\cdot\frac{n-2}{2} + 1\cdot 1\cdot\frac{n-2}{2} = \frac{n^2+8}{4}$$

nonisomorphic indecomposable MCM modules.

We conclude this section by noting that a two-dimensional local Krull domain $(R, \mathfrak{m})$ having completion isomorphic to a ring in Table 1 may not be factorial. However, Theorem 5.4 implies that $\mathcal{C}(\mathfrak{C}(R))$ is a factor group of a group given in Theorem 5.5. In particular, $\mathcal{C}(\mathfrak{C}(R))$ is a finite cyclic group such that every class contains a prime divisor, and thus the arithmetic of $\mathfrak{C}(R)$ is described in Theorem 6.8.

6. The arithmetic of monoids of modules

In this section we study the arithmetic of the Krull monoids that have been discussed in Sections 4 and 5. Thus, using the transfer properties presented in Section 2, we describe the arithmetic of direct-sum decompositions of modules. Suppose that H is a Krull monoid having a divisor homomorphism $\varphi : H \to \mathcal{F}(\mathcal{P})$ and let $G_{\mathcal{P}} \subset \mathcal{C}(\varphi)$ be the set of classes containing prime divisors. The first subsection deals with quite general sets $G_{\mathcal{P}}$ and provides results on the finiteness or nonfiniteness of various arithmetical parameters. The second subsection studies three specific sets $G_{\mathcal{P}}$, provides explicit results on arithmetical parameters, and establishes a characterization result (Theorems 6.4, 6.7, 6.8, and Corollary 6.10). The third subsection completely determines the system of sets of lengths in case of small subsets $G_{\mathcal{P}}$. It shows that small subsets in torsion groups and in torsion-free groups can have the same systems of sets of lengths, and it reveals natural limits for arithmetical characterization results (Corollary 6.15).

6A. ***General sets $G_{\mathcal{P}}$ of classes containing prime divisors.*** In this subsection we consider the algebraic and arithmetic structure of Krull monoids with respect to $G_{\mathcal{P}}$. We will often assume that $G_{\mathcal{P}} = -G_{\mathcal{P}}$, a property which has a strong influence on the arithmetic of H. Recall that $G_{\mathcal{P}} = -G_{\mathcal{P}}$ holds in many of the (finite and infinite representation type) module-theoretic contexts described in Sections 4 and 5. More generally, all configurations $(G, G_{\mathcal{P}})$ occur for certain monoids of modules (see Proposition 4.18, [Herbera and Příhoda 2010] and [Leuschke and Wiegand 2012, Chapter 1]) and, by Claborn's realization theorem, all configurations $(G, G_{\mathcal{P}})$ occur for Dedekind domains (see [Geroldinger and Halter-Koch 2006, Theorem 3.7.8]). In addition, every abelian group can be realized as the class group of a Dedekind domain which is a quadratic extension of a principal ideal domain, and in this case we have $G_{\mathcal{P}} = -G_{\mathcal{P}}$ (see [Leedham-Green 1972]).

Proposition 6.1. *Let H be a Krull monoid, $\varphi : H \to \mathcal{F}(\mathcal{P})$ a divisor homomorphism with class group $G = \mathcal{C}(\varphi)$, and let $G_{\mathcal{P}} \subset G$ denote the set of classes containing prime divisors.*

(1) *If $G_{\mathcal{P}}$ is finite, then $\mathcal{A}(G_{\mathcal{P}})$ is finite and hence $\mathsf{D}(G_{\mathcal{P}}) < \infty$. If G has finite total rank, then $G_{\mathcal{P}}$ is finite if and only if $\mathcal{A}(G_{\mathcal{P}})$ is finite if and only if $\mathsf{D}(G_{\mathcal{P}}) < \infty$.*

(2) *If $G_{\mathcal{P}} = -G_{\mathcal{P}}$, then $[G_{\mathcal{P}}] = G$. Moreover, the map $\varphi : H \to D$ and the inclusion $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow \mathcal{F}(G_{\mathcal{P}})$ are both cofinal.*

(3) *Suppose that G is infinite cyclic, say $G = \langle e \rangle$, and that $\{-e, e\} \subset G_{\mathcal{P}}$. Then $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow \mathcal{F}(G_{\mathcal{P}})$ is a divisor theory if and only if there exist $k, l \in \mathbb{N}_{\ge 2}$ such that $-ke, le \in G_{\mathcal{P}}$.*

(4) *Let $r, \alpha \in \mathbb{N}$ with $r + \alpha > 2$. Let $(e_1, \dots, e_r) \in G_{\mathcal{P}}^r$ be independent and let $e_0 \in G_{\mathcal{P}}$ such that $\alpha e_0 = e_1 + \dots + e_r$, $\{-e_0, \dots, -e_r\} \subset G_{\mathcal{P}}$, and $\langle e_0, \dots, e_r \rangle = G$.*
   (a) *The map $\varphi : H \to \mathcal{F}(\mathcal{P})$ and the inclusion $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow \mathcal{F}(G_{\mathcal{P}})$ are divisor theories with class group isomorphic to G.*
   (b) *If $0 \notin G_{\mathcal{P}}$, then $\mathcal{B}(G_{\mathcal{P}})$ is not a direct product of nontrivial submonoids.*

Proof. (1) follows from [Geroldinger and Halter-Koch 2006, Theorem 3.4.2].

If $G_{\mathcal{P}} = -G_{\mathcal{P}}$, then $[G_{\mathcal{P}}] = \langle G_{\mathcal{P}} \rangle = G$. By Lemma 3.4, (2) follows once we verify that φ is cofinal. If $p \in \mathcal{P}$, then there is a $q \in \mathcal{P}$ with $q \in -[p]$, whence there is an $a \in H$ with $\varphi(a) = pq$, and so φ is cofinal.

If $\{-ke : k \in \mathbb{N}\} \cap G_{\mathcal{P}} = \{-e\}$ or $\{ke : k \in \mathbb{N}\} \cap G_{\mathcal{P}} = \{e\}$, then $\mathcal{B}(G_{\mathcal{P}})$ is factorial. Since $\mathcal{F}(G_{\mathcal{P}}) \ne B(G_{\mathcal{P}})$, the inclusion $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow F(G_{\mathcal{P}})$ is not a divisor theory. Conversely, suppose that there exist $k, l \in \mathbb{N}_{\ge 2}$ such that $-ke, le \in G_{\mathcal{P}}$. Let $m \in \mathbb{N}$. If $me \in G_{\mathcal{P}}$, then $me = \gcd\big((me)(-e)^m, (me)^k(-ke)^m\big)$, and if $-me \in G_{\mathcal{P}}$, then $-me = \gcd\big((-me)e^m, (-me)^l(le)^m\big)$. Thus every element of $G_{\mathcal{P}}$ is a greatest common divisor of a finite set of elements from $\mathcal{B}(G_{\mathcal{P}})$ and hence $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow F(G_{\mathcal{P}})$ is a divisor theory.

We now suppose that G and $G_{\mathcal{P}}$ are as in (4). To prove (a) it is sufficient to show that $\mathcal{B}(G_{\mathcal{P}}) \hookrightarrow \mathcal{F}(G_{\mathcal{P}})$ is a divisor theory. By [Geroldinger and Halter-Koch 2006, Proposition 2.5.6] we need only verify that $\langle G_{\mathcal{P}} \rangle = [G_{\mathcal{P}} \setminus \{g\}]$ for every $g \in G_{\mathcal{P}}$. Let $g \in G$. We will show that

$$[G_{\mathcal{P}} \setminus \{g\}] = [G_{\mathcal{P}}] = [e_0, \dots, e_r, -e_0, \dots, -e_r].$$

If $g \notin \{e_0, \dots, e_r, -e_0, \dots, -e_r\}$, the assertion is clear. By symmetry it suffices to consider the case where $g \in \{e_0, \dots, e_r\}$. If $g = e_i$ for some $i \in [1, r]$, then

$$e_i = \alpha e_0 + (-e_1) + \dots + (-e_{i-1}) + (-e_{i+1}) + \dots + (-e_r) \in [\{\pm e_\nu : \nu \in [0, r] \setminus \{i\}],$$

and hence

$$[G_{\mathcal{P}} \setminus \{g\}] \supset [\{\pm e_\nu : \nu \in [0, r] \setminus \{i\}\}] = [\pm e_0, \dots, \pm e_r] = G\,.$$

If $g = e_0$, then

$$e_0 = e_1 + \cdots + e_r + (\alpha - 1)(-e_0) \in [G_{\mathcal{P}} \setminus \{e_0\}],$$

and hence $[G_{\mathcal{P}} \setminus \{g\}] = [G_{\mathcal{P}}] = G$.

To prove (b) we use Lemma 3.2 and suppose that $0 \notin G_{\mathcal{P}}$ with $G_{\mathcal{P}} = G_1 \uplus G_2$ such that $\mathcal{A}(G_{\mathcal{P}}) = \mathcal{A}(G_1) \uplus \mathcal{A}(G_2)$. We must show that either G_1 or G_2 is empty. Suppose that $V = e_1 \cdot \ldots \cdot e_r (-e_0)^{\alpha} \in \mathcal{A}(G_1)$. Since $(-e_0)e_0, \ldots, (-e_r)e_r \in \mathcal{A}(G_{\mathcal{P}})$, it follows that $\{\pm e_0, \ldots, \pm e_r\} \subset G_1$. Let $g \in G_{\mathcal{P}}$. Since $G_{\mathcal{P}} \subset G = [\pm e_0, \ldots, \pm e_r]$, there exists $U \in \mathcal{A}(G_{\mathcal{P}})$ such that $g \in \operatorname{supp}(U) \subset \{g, \pm e_0, \ldots, \pm e_r\}$, and hence $g \in G_1$. Thus $G_1 = G_{\mathcal{P}}$ and $G_2 = \varnothing$. $\square$

For our characterization results, we need to recall the concept of an absolutely irreducible element, a classical notion in algebraic number theory. An element u in an atomic monoid H is called *absolutely irreducible* if $u \in \mathcal{A}(H)$ and $|\mathsf{Z}(u^n)| = 1$ for all $n \in \mathbb{N}$; equivalently, the divisor-closed submonoid of H generated by u is factorial. Suppose that $H \hookrightarrow \mathcal{F}(\mathcal{P})$ is a divisor theory with class group G and that $u = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m}$ where $m, k_1, \ldots, k_m \in \mathbb{N}$ and where $p_1, \ldots, p_m \in \mathcal{P}$ are pairwise distinct. Then u is absolutely irreducible. if and only if $(k_1, \ldots, k_m)$ is a minimal element of the set

$$\Gamma = \{(s_1, \ldots, s_m) \in \mathbb{N}_0^l : p_1^{s_1} \cdot \ldots \cdot p_m^{s_m} \in H\} \setminus \{\mathbf{0}\}$$

relative to the usual product ordering, and the torsion-free rank of $\langle [p_1], \ldots, [p_m] \rangle$ in G is $m-1$ (see [Geroldinger and Halter-Koch 2006, Proposition 7.1.4]). In particular, if $[p_1] \in G$ has finite order, then $p_1^{\operatorname{ord}([p_1])}$ is absolutely irreducible, and if $[p_1] \in G$ has infinite order, then $p_1 q_1$ is absolutely irreducible for all $q_1 \in \mathcal{P} \cap (-[p_1])$.

Proposition 6.2. *Let H be a Krull monoid, $\varphi : H \to \mathcal{F}(\mathcal{P})$ a cofinal divisor homomorphism with class group G, and let $G_{\mathcal{P}} \subset G$ denote the set of classes containing prime divisors.*

(1) *Suppose that $G_{\mathcal{P}}$ is infinite.*

  (a) *If $G_{\mathcal{P}}$ has an infinite subset G_0 such that $G_0 \cup (-G_0) \subset G_{\mathcal{P}}$ and $\langle G_0 \rangle$ has finite total rank, then $\mathcal{U}_k(H)$ is infinite for each $k \geq 2$. Moreover, $\mathsf{D}(G_{\mathcal{P}}) = \rho_k(H) = \omega(H) = \mathsf{t}(H) = \infty$.*

  (b) *If there exists $e \in G$ such that $G_{\mathcal{P}} \cap \{ke : k \in \mathbb{N}\}$ and $G_{\mathcal{P}} \cap \{-ke : k \in \mathbb{N}\}$ are both infinite, then $\Delta(H)$ is infinite and $\mathsf{c}(H) = \mathsf{c}_{\mathrm{mon}}(H) = \infty$.*

  (c) *If $G_{\mathcal{P}}$ contains an infinite group, then every finite subset $L \subset \mathbb{N}_{\geq 2}$ occurs as a set of lengths.*

(2) *Suppose that $G_{\mathcal{P}}$ is finite and that H is not factorial.*

  (a) *The set $\Delta(H)$ is finite and there is a constant $M_1 \in \mathbb{N}$ such that every set of lengths is an AAMP with difference $d \in \Delta^*(H)$ and bound M_1.*

(b) *There is a constant $M_2 \in \mathbb{N}$ such that, for every $k \geq 2$, the set $\mathcal{U}_k(H)$ is an AAMP with period $\{0, \min \Delta(H)\}$ and bound M_2.*

(c) $\mathsf{c}(H) \leq \omega(H) \leq \mathsf{t}(H) \leq 1 + \frac{1}{2}\mathsf{D}(G_{\mathcal{P}})(\mathsf{D}(G_{\mathcal{P}}) - 1)$ *and*

$$\mathsf{c}_{\mathrm{mon}}(H) < \frac{|G_{\mathcal{P}}^{\bullet}| + 2}{2}\big((2|G_{\mathcal{P}}^{\bullet}| + 2)(|G_{\mathcal{P}}^{\bullet}| + 2)(\mathsf{D}(G_{\mathcal{P}}) + 1)\big)^{|G_{\mathcal{P}}^{\bullet}|+1}.$$

(d) *Suppose that $G_{\mathcal{P}} = -G_{\mathcal{P}}$. Then $\omega(H) = \mathsf{D}(G_{\mathcal{P}})$, $\rho(H) = \mathsf{D}(G_{\mathcal{P}})/2$, $\rho_{2k}(H) = k\mathsf{D}(G_{\mathcal{P}}) < \infty$, and $\lambda_{k\mathsf{D}(G_{\mathcal{P}})+j}(H) = 2k + j$ for all $k \in \mathbb{N}$ and $j \in [0, 1]$. If G is torsion-free, then $\mathsf{D}(G_{\mathcal{P}})$ is the maximal number s of absolutely irreducible atoms $u_1, \ldots, u_s$ such that $2 \in \mathsf{L}(u_1 \cdot \ldots \cdot u_s)$.*

(e) *If, in particular, $G_{\mathcal{P}} = -G_{\mathcal{P}}$ and $\mathsf{D}(G_{\mathcal{P}}) = 2$, then $\mathcal{U}_k(H) = \{k\}$ for all $k \in \mathbb{N}$ and $\mathsf{c}_{\mathrm{mon}}(H) = \mathsf{c}(H) = \omega(H) = \mathsf{t}(H) = 2$.*

Proof. Throughout the proof we implicitly assume the results of Lemma 2.2 and of Lemma 3.4. In particular, we have $\rho(H) \leq \omega(H)$ and $\mathsf{c}(H) \leq \omega(H) \leq \mathsf{D}(G_{\mathcal{P}})$.

For (1), suppose that $G_{\mathcal{P}}$ is infinite. We first prove (a). Theorem 3.4.2 in [Geroldinger and Halter-Koch 2006] implies that $\mathcal{A}(G_0)$ and $\mathsf{D}(G_0)$ are infinite. Thus, for every $k \in \mathbb{N}$, there is $U_k \in \mathcal{A}(G_0)$ with $|U_k| \geq k$ and hence $\mathsf{L}\big(U_k(-U_k)\big) \supset \{2, |U_k|\}$. This implies that $\mathcal{U}_2(G_0)$ is infinite and thus $\mathcal{U}_k(G_0)$ is infinite for all $k \geq 2$. Therefore $\rho_k(H) = \rho_k(G_{\mathcal{P}}) = \infty$ for all $k \geq 2$ and, since $\rho(H) \leq \omega(H) \leq \mathsf{t}(H)$, each of these invariants is infinite.

Item 1(b) follows from [Geroldinger et al. 2010, Theorem 4.2].

Item 1(c) is a realization result is due to Kainrath. See [Kainrath 1999] or [Geroldinger and Halter-Koch 2006, Theorem 7.4.1].

Now, in order to prove (2), we suppose that $G_{\mathcal{P}}$ is finite and that H is not factorial. Then $\mathsf{D}(G_{\mathcal{P}}) > 1$ and $2 \leq \mathsf{c}(H) \leq \omega(H) \leq \mathsf{t}(H)$. By Proposition 6.1, $\mathcal{B}(G_{\mathcal{P}})$ is finitely generated and $\mathsf{D}(G_{\mathcal{P}}) < \infty$. The respective upper bounds given in (c) for $\mathsf{c}_{\mathrm{mon}}(H)$ and $\mathsf{t}(H)$ can be found in [Geroldinger and Yuan 2013, Theorem 3.4] and [Geroldinger and Halter-Koch 2006, Theorem 3.4.10].

We now consider (a). Since $2 + \sup \Delta(H) \leq \mathsf{c}(H) < \infty$, (c) implies that $\Delta(H)$ is finite. Since $\mathcal{B}(G_{\mathcal{P}})$ is finitely generated, the assertion on the structure of sets of lengths follows from [Geroldinger and Halter-Koch 2006, Theorem 4.4.11].

Since $\mathsf{t}(H) < \infty$ and $\Delta(H)$ is finite, [Gao and Geroldinger 2009, Theorems 3.5 and 4.2] imply the assertion in (b) on the structure of the unions of sets of lengths.

In order to prove (d), we suppose that $G_{\mathcal{P}} = -G_{\mathcal{P}}$. The statements about $\rho_{2k}(H)$, $\rho(H)$, and $\lambda_{k\mathsf{D}(G_{\mathcal{P}})+j}(H)$ follow from Lemma 3.3, and it remains to show that $\omega(H) = \mathsf{D}(G_{\mathcal{P}})$. We have $\omega(H) \leq \mathsf{D}(G_{\mathcal{P}}) < \infty$. If $\mathsf{D}(G_{\mathcal{P}}) = 2$, then $\omega(H) = \mathsf{D}(G_{\mathcal{P}})$. Suppose that $\mathsf{D}(G_{\mathcal{P}}) \geq 3$. If $V = g_1 \cdot \ldots \cdot g_l \in \mathcal{A}(G_{\mathcal{P}})$ with $|V| = l = \mathsf{D}(G_{\mathcal{P}})$ and $U_i = (-g_i)g_i$ for all $i \in [1, l]$, then $V \mid U_1 \cdot \ldots \cdot U_l$ but yet V divides no proper subproduct of $U_1 \cdot \ldots \cdot U_l$. Thus $\mathsf{D}(G_{\mathcal{P}}) \leq \omega(G_{\mathcal{P}}) \leq \omega(H)$.

Let t denote the maximal number of absolutely irreducible atoms with the required property. Since $\rho(H) = \mathsf{D}(G_{\mathcal{P}})/2$, it follows that $t \le \mathsf{D}(G_{\mathcal{P}})$. Let $V = g_1 \cdot \ldots \cdot g_l \in \mathcal{A}(G_{\mathcal{P}})$ with $|V| = l = \mathsf{D}(G_{\mathcal{P}})$. For $i \in [1, l]$ choose an element $p_i \in \mathcal{P} \cap g_i$ and an element $q_i \in \mathcal{P} \cap (-g_i)$. Since G is torsion-free, the element $u_i = p_i q_i \in H$ is absolutely irreducible for each $i \in [1, l]$ and, by construction, we have $2 \in \mathsf{L}(u_1 \cdot \ldots \cdot u_l)$.

The statement in (e) follows immediately from (c) and (d). □

Let all notation be as in Proposition 6.2. We note that if $G_{\mathcal{P}}$ is infinite but without a subset G_0 as in (1a), then none of the conclusions of (1a) need hold. A careful analysis of the case where G is an infinite cyclic groups is handled in [Geroldinger et al. 2010]. We also note that the description of the structure of sets of lengths given in (2a) is best possible (see [Schmid 2009c]).

By Lemma 3.4, many arithmetical phenomena of a Krull monoid H are determined by the tuple $(G, G_{\mathcal{P}})$. We now provide a first result indicating that conversely arithmetical phenomena give us back information on the class group. Indeed, our next corollary characterizes arithmetically whether the class group of a Krull monoid is torsion-free or not. To do so we must study the arithmetical behavior of elements similar to absolutely irreducible elements. Note that such a result cannot be accomplished via sets of lengths alone (see Propositions 6.12 and 6.13 and (1c) of Proposition 6.2; in fact, there is an open conjecture that every abelian group is the class group of a half-factorial Krull monoid [Geroldinger and Göbel 2003]).

Proposition 6.3. *Let H be a Krull monoid with class group G. Then G has an element of infinite order if and only if there exists an irreducible element $u \in H$ having the following two arithmetical properties.*

(a) *Whenever there are $v \in H \setminus H^\times$ and $m \in \mathbb{N}$ with $v \mid u^m$, then $u \mid v^n$ for some $n \in \mathbb{N}$.*

(b) *There exist $l \ge 2$ and $a_1, \ldots, a_l \in H$ such that $u \mid a_1 \cdot \ldots \cdot a_l$ but yet*

$$u \nmid a_\nu^{-1} (a_1 \cdot \ldots \cdot a_l)^N$$

for each $\nu \in [1, l]$ and for every $N \in \mathbb{N}$.

Proof. We may assume that H is reduced. Consider a divisor theory $H \hookrightarrow \mathcal{F}(\mathcal{P})$ and denote by $G_{\mathcal{P}} \subset G$ the set of classes containing prime divisors.

First suppose that G is a torsion group and let $u \in \mathcal{A}(H)$ have Property (a). Then $u = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m}$ for some $m, k_1, \ldots, k_m \in \mathbb{N}$ and pairwise distinct elements $p_1, \ldots, p_m \in \mathcal{P}$. Then (a) implies that $k = 1$ and hence u is absolutely irreducible. Thus Property (b) cannot hold for any $l \ge 2$.

Conversely, suppose that G is not a torsion group. Since $[G_{\mathcal{P}}] = G$ there exists a $p \in \mathcal{P}$ such that $[p] \in G$ has infinite order, and there is an element $u' \in \mathcal{A}(H)$ with

$p \mid u'$. Suppose that $u' = p_1 \cdot \ldots \cdot p_n \cdot q_1 \cdot \ldots \cdot q_r$, where

$$p = p_1, p_2, \ldots, p_n, q_1, \ldots, q_r \in \mathcal{P},$$

$[p_1], \ldots, [p_n]$ have infinite order, and $[q_1], \ldots, [q_r]$ have finite order, each of which divides some integer N. Then $(q_1 \cdot \ldots \cdot q_r)^N \in H$, whence $(p_1 \cdot \ldots \cdot p_n)^N \in H$. After a possible reordering there is an atom $u = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m} \in \mathcal{A}(H)$ dividing a power of $(p_1 \cdot \ldots \cdot p_n)^N$ such that there is no atom $v \in \mathcal{A}(H)$ with $\operatorname{supp}_{\mathcal{P}}(v) \subsetneq \{p_1, \ldots, p_m\}$. Thus u satisfies Property (a). Since $H \hookrightarrow \mathcal{F}(\mathcal{P})$ is a divisor theory, there exist $b_1, \ldots, b_s \in H$ such that

$$p_1^{k_1} \cdot \ldots \cdot p_{m-1}^{k_{m-1}} = \gcd(b_1, \ldots, b_s).$$

Hence there is an $i \in [1, s]$, say $i = 1$, such that $p_m \nmid b_1$. Similarly, there are $c_1, \ldots, c_t \in H$ such that $p_m^{k_m} = \gcd(c_1, \ldots, c_t)$. Without loss of generality, there exists $i \in [1, m-1]$ such that $p_i \nmid c_1$. Therefore $u \mid b_1 c_1$, but yet $u \nmid b_1^N$ and $u \nmid c_1^N$ for any $N \in \mathbb{N}$, and so Property (b) is satisfied. ☐

Propositions 6.1, 6.2, and 6.3 provide abstract finiteness and nonfiniteness results. To obtain more precise information on the arithmetical invariants, we require specific information on $G_{\mathcal{P}}$. In the next subsection we will use such specific information to give more concrete results.

6B. ***Specific sets $G_{\mathcal{P}}$ of classes containing prime divisors and arithmetical characterizations.*** We now provide an in-depth study of the arithmetic of three classes of Krull monoids studied in Sections 4 and 5. Theorem 6.4 describes the arithmetic of the monoids discussed in Examples 4.12, 4.20, 4.25 and in Theorem 4.12. Its arithmetic is simple enough that we can more or less give a complete description.

Theorem 6.4. *Let H be a Krull monoid with class group G and suppose that*

$$G_{\mathcal{P}} = \{e_0, \ldots, e_r, -e_0, \ldots, -e_r\} \subset G$$

is the set of classes containing prime divisors, where $r, \alpha \in \mathbb{N}$ with $r + \alpha > 2$ and $(e_1, \ldots, e_r)$ is an independent family of elements each having infinite order such that $e_1 + \cdots + e_r = \alpha e_0$. Then:

(1) *$\mathcal{A}(G_{\mathcal{P}}) = \{V, -V, U_\nu : \nu \in [0, r]\}$, where $V = (-e_0)^\alpha e_1 \cdot \ldots \cdot e_r$ and $U_\nu = (-e_\nu) e_\nu$ for all $\nu \in [0, r]$. In particular, $\mathsf{D}(G_{\mathcal{P}}) = r + \alpha$.*

(2) *Suppose that*

$$S = \prod_{i=0}^{r} e_i^{k_i} (-e_i)^{l_i} \in \mathcal{F}(G_{\mathcal{P}}),$$

where $k_0, l_0, \ldots, k_r, l_r \in \mathbb{N}_0$. Then $S \in \mathcal{B}(G_{\mathcal{P}})$ if and only if

$$l_i = \alpha^{-1}(k_0 - l_0) + k_i$$

for all $i \in [1, r]$. *If* $S \in \mathcal{B}(G_{\mathcal{P}})$ *with* $k_0 \ge l_0$ *and* $k^* = \min\{k_1, \ldots, k_r\}$, *then*

$$\mathsf{Z}(S) = \left\{ V^{\nu}(-V)^{\alpha^{-1}(k_0-l_0)+\nu} U_0^{l_0-\alpha\nu} \prod_{i=1}^{r} U_i^{k_i-\nu} : \nu \in [0, \min\{\alpha^{-1} l_0, k^*\}] \right\}$$

and

$$\mathsf{L}(S) = \left\{ \alpha^{-1}(k_0-l_0)+l_0+k_1+\cdots+k_r-(r+\alpha-2)\nu : \nu \in [0, \min\{\alpha^{-1} l_0, k^*\}] \right\}.$$

(3) *The system of sets of lengths of* H *can be described as follows.*

(a) $\Delta(H) = \{r+\alpha-2\}$.

(b) $\rho(H) = \mathsf{D}(G_{\mathcal{P}})/2$.

(c) *For each* $k \in \mathbb{N}$, *the set* $\mathcal{U}_k(H)$ *is an arithmetical progression with difference* $r+\alpha-2$.

(d) *For each* $k \in \mathbb{N}$ *and each* $j \in [0,1]$, $\rho_{2k+j}(H) = k\mathsf{D}(G_{\mathcal{P}}) + j$.

(e) *For each* $l \in \mathbb{N}_0$, $\lambda_{l\mathsf{D}(G_{\mathcal{P}})+j}(H) = 2l + j$ *whenever* $j \in [0, \mathsf{D}(G_{\mathcal{P}})-1]$ *and* $l\mathsf{D}(G_{\mathcal{P}}) + j \ge 1$.

(f) *Finally*,

$$\mathcal{L}(H) = \left\{ m + \{2k^* + (r+\alpha-2)\lambda : \lambda \in [0, k^*]\} : m, k^* \in \mathbb{N}_0 \right\}.$$

(4) $\mathsf{c}(H) = \mathsf{c}_{\mathrm{mon}}(H) = \omega(H) = \mathsf{t}(H) = \mathsf{D}(G_{\mathcal{P}}) = r+\alpha$.

Proof. By Lemma 3.4, all assertions on lengths of factorizations and on catenary degrees can be proved working in $\mathcal{B}(G_{\mathcal{P}})$ instead of H.

Obviously, $\{U_\nu : \nu \in [0, r]\} \subset \mathcal{A}(G_{\mathcal{P}})$ and to prove (1) it remains to verify that if $W \in \mathcal{A}(G_{\mathcal{P}})$ with $W \ne U_\nu$, then $W = V$. Note that $e_0 \in \langle e_1, \ldots, e_r \rangle$ but that $\langle e_0 \rangle \cap \langle e_i : i \in I \rangle = \{0\}$ for any proper subset $I \subsetneq [1, r]$. Thus, if

$$W = \prod_{i \in I} e_i^{k_i} (-e_0)^{k_0} \in \mathcal{A}(G_{\mathcal{P}}) \setminus \{U_\nu : \nu \in [0, r]\},$$

where $\emptyset \ne I \subset [1, r]$ and $k_0, k_i \in \mathbb{N}$ for all $i \in I$, then $k_0 e_0 = \sum_{i \in I} k_i e_i \in \langle e_i : i \in I \rangle$ and hence $I = [1, r]$. Assume to the contrary that there is $i \in [1, r]$ such that $k_i > 1$. Since $V \in \mathcal{B}(G_{\mathcal{P}})$ and $W \in \mathcal{A}(G_{\mathcal{P}})$, it follows that $k_0 \in [1, \alpha-1]$. Then

$$0 \ne (k_1-1)e_1+\cdots+(k_r-1)e_r = (k_0-\alpha)e_0 \in [e_1, \ldots, e_r] \cap [-e_1, \ldots, -e_r] = \{0\},$$

a contradiction. Thus $k_1 = \cdots = k_r = 1$ and we obtain that $k_0 = \alpha$, whence $W = V \in \mathcal{A}(G_{\mathcal{P}})$.

To prove (2), suppose that $S \in \mathcal{B}(G_{\mathcal{P}})$ and that $l_0 \ge k_0$. Then

$$S' = (-e_0)^{l_0-k_0} \prod_{i=1}^{r} e_i^{k_i} (-e_i)^{l_i} \in \mathcal{B}(G_{\mathcal{P}}),$$

whence $l_0 - k_0 = \alpha m_0 \in \alpha\mathbb{N}_0$, $S'' = \prod_{i=1}^{r} e_i^{k_i-m_0} (-e_i)^{l_i} \in \mathcal{B}(G_{\mathcal{P}})$, and

$$l_i = k_i - m_0 = k_i - \frac{l_0 - k_0}{\alpha} = \frac{k_0 - l_0}{\alpha} + k_i \quad \text{for each } i \in [1, r].$$

The same holds true if $l_0 \le k_0$. Conversely, if $l_1, \dots, l_r$ satisfy the asserted equations, then obviously $\sigma(S) = 0$.

Suppose that $S \in \mathcal{B}(G_{\mathcal{P}})$ and that $k_0 \ge l_0$. Then

$$\begin{aligned}
S &= e_0^{k_0}(-e_0)^{l_0} \prod_{i=1}^{r} e_i^{k_i}(-e_i)^{k_i+\alpha^{-1}(k_0-l_0)} \\
&= \big((-e_0)e_0\big)^{l_0}\big(e_0{}^{\alpha}(-e_1)\cdot\ldots\cdot(-e_r)\big)^{\alpha^{-1}(k_0-l_0)} \prod_{i=1}^{r} \big((-e_i)e_i\big)^{k_i} \\
&= \big((-e_0)e_0\big)^{l_0-\alpha\nu}\big(e_0{}^{\alpha}(-e_1)\cdot\ldots\cdot(-e_r)\big)^{\alpha^{-1}(k_0-l_0)+\nu}\big((-e_0)^{\alpha}e_1\cdot\ldots\cdot e_r\big)^{\nu} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times \prod_{i=1}^{r} \big((-e_i)e_i\big)^{k_i-\nu} \\
&= U_0^{l_0-\alpha\nu}(-V)^{\alpha^{-1}(k_0-l_0)+\nu} V^{\nu} \prod_{i=1}^{r} U_i^{k_i-\nu}
\end{aligned}$$

for each $\nu \in [0, \min\{\alpha^{-1} l_0, k^*\}]$. Therefore $\mathsf{Z}(S)$ and hence $\mathsf{L}(S)$ have the given forms.

We now consider the statements of (3). The assertion on $\Delta(G_{\mathcal{P}})$ follows immediately from (2). Since $\Delta(G_{\mathcal{P}}) = \{r + \alpha - 2\}$, all sets $\mathcal{U}_k(G_{\mathcal{P}})$ are arithmetical progressions with difference $r + \alpha - 2$. The assertion on each $\rho_{2k}(G_{\mathcal{P}})$ and each $\rho(G_{\mathcal{P}})$ follow from Proposition 6.2.

In order to determine $\mathcal{L}(G_{\mathcal{P}})$, let $S \in \mathcal{B}(G_{\mathcal{P}})$ be given with all parameters as in (2). First suppose that $l_0 \ge \alpha k^*$. Then

$$\begin{aligned}
\mathsf{L}(S) &= (\alpha^{-1}(k_0 - l_0) + l_0 + k_1 + \cdots + k_r) \\
&\qquad\qquad -(r+\alpha)k^* + \{(r+\alpha)k^* - (r+\alpha-2)\nu : \nu \in [0, k^*]\} \\
&= (\alpha^{-1}(k_0 - l_0) + l_0 + k_1 + \cdots + k_r) \\
&\qquad\qquad -(r+\alpha)k^* + \{2k^* + (r+\alpha-2)\lambda : \lambda \in [0, k^*]\}.
\end{aligned}$$

Thus $\mathsf{L}(S)$ has the form

$$\mathsf{L}(S) = m + \{2k^* + (r+\alpha-2)\lambda : \lambda \in [0, k^*]\}$$

for some $m, k^* \in \mathbb{N}_0$. Conversely, for every choice of $m, k^* \in \mathbb{N}_0$, there is an $S \in \mathcal{B}(G_{\mathcal{P}})$ such that $\mathsf{L}(S)$ has the given form.

Now suppose that $l_0 \le \alpha k^* - 1$ and set $m_0 = \lfloor l_0/\alpha \rfloor$. Then

$$\begin{aligned}
\mathsf{L}(S) &= (\alpha^{-1}(k_0 - l_0) + l_0 + k_1 + \cdots + k_r) \\
&\qquad\qquad -(r+\alpha)m_0 + \{(r+\alpha)m_0 - (r+\alpha-2)\nu : \nu \in [0, m_0]\}
\end{aligned}$$

$$= (\alpha^{-1}(k_0 - l_0) + l_0 + k_1 + \cdots + k_r) \\ -(r+\alpha)m_0 + \{2m_0 + (r+\alpha-2)\lambda : \lambda \in [0, m_0]\},$$

and hence $\mathsf{L}(S)$ has the form

$$\mathsf{L}(S) = m + \{2m_0 + (r+\alpha-2)\lambda : \lambda \in [0, m_0]\}$$

for some $m \in \mathbb{N}$ and $m_0 \in \mathbb{N}_0$.

Next we verify that, for every $k \in \mathbb{N}$, $\rho_{2k+1}(G_{\mathcal{P}}) \le k\mathsf{D}(G_{\mathcal{P}})+1$. By [Geroldinger and Halter-Koch 2006, Proposition 1.4.2], for all $k \in \mathbb{N}$,

$$\rho_k(G_{\mathcal{P}}) = \sup\{\sup L : L \in \mathcal{L}(G_{\mathcal{P}}), k = \min L\}.$$

Thus we may choose $k \in \mathbb{N}$ and $L \in \mathcal{L}(G_{\mathcal{P}})$ with $\min L = 2k+1$. Then, there exist $l, m, k^* \in \mathbb{N}_0$ such that

$$L = m + \{2k^* + (r+\alpha-2)\lambda : \lambda \in [0, k^*]\}$$

with $m = 2l+1$ and $2k+1 = \min L = 2(k^*+l)+1$. Now

$$\max L = m + (r+\alpha)k^* = 2l + 1 + (r+\alpha)(k-l) \\ = (r+\alpha)k + 1 - (r+\alpha-2)l \le k\mathsf{D}(G_{\mathcal{P}}) + 1,$$

and thus $\rho_{2k+1}(G_{\mathcal{P}}) \le k\mathsf{D}(G_{\mathcal{P}}) + 1$.

It remains to verify the assertions on the $\lambda_{l\mathsf{D}(G_{\mathcal{P}})+j}(G_{\mathcal{P}})$. Let $l \in \mathbb{N}_0$ and $j \in [0, \mathsf{D}(G_{\mathcal{P}})-1]$. Then Lemma 3.3 implies $\lambda_{l\mathsf{D}(G_p)+j}(G_{\mathcal{P}}) \le 2l + j$, and that equality holds if $j \in [0,1]$. It remains to verify that $\lambda_{l\mathsf{D}(G_p)+j}(G_{\mathcal{P}}) \ge 2l + j$ when $j \in [2, \mathsf{D}(G_{\mathcal{P}})-1]$. Let $L \in \mathcal{L}(G_{\mathcal{P}})$ with $l\mathsf{D}(G_{\mathcal{P}}) + j \in L$. Then there exist $m, k^* \in \mathbb{N}_0$ such that

$$L = m + \{2k^* + (r+\alpha-2)\lambda : \lambda \in [0, k^*]\} \\ = m + k^*\mathsf{D}(G_{\mathcal{P}}) - \{(\mathsf{D}(G_{\mathcal{P}})-2)\nu : \nu \in [0, k^*]\}.$$

Suppose $l\mathsf{D}(G_{\mathcal{P}}) + j = \max L - \nu(\mathsf{D}(G_{\mathcal{P}})-2) = m + k^*\mathsf{D}(G_{\mathcal{P}}) - \nu(\mathsf{D}(G_{\mathcal{P}})-2)$ for some $\nu \in [0, k^*]$. Then $j \equiv m + 2\nu \mod \mathsf{D}(G_{\mathcal{P}})$ and hence $m + 2\nu \ge j$. This implies

$$(k^* - \nu)\mathsf{D}(G_{\mathcal{P}}) + j \le m + k^*\mathsf{D}(G_{\mathcal{P}}) - \nu(\mathsf{D}(G_{\mathcal{P}})-2) = l\mathsf{D}(G_{\mathcal{P}}) + j,$$

and hence $l \ge k^* - \nu$. Therefore we obtain

$$\min L = l\mathsf{D}(G_{\mathcal{P}}) + j - (k^*-\nu)(\mathsf{D}(G_{\mathcal{P}})-2) \\ = (l - k^* + \nu)\mathsf{D}(G_{\mathcal{P}}) + j + 2(k^*-\nu) \ge 2l + j,$$

and thus $\lambda_{l\mathsf{D}(G_p)+j}(G_{\mathcal{P}}) \ge 2l + j$.

Finally we consider the catenary degrees of H and prove the statements given in (4). Using Proposition 6.2 we infer

$$\mathsf{D}(G_{\mathcal{P}}) = r + \alpha = 2 + \max \Delta(G_{\mathcal{P}}) \le \mathsf{c}(G_{\mathcal{P}}) = \mathsf{c}(H) \le \omega(H) \le \mathsf{t}(H).$$

Since $\mathsf{c}(G_{\mathcal{P}}) \le \mathsf{c}_{\mathrm{mon}}(G_{\mathcal{P}})$, it remains to show that $\mathsf{c}_{\mathrm{mon}}(G_{\mathcal{P}}) \le \mathsf{D}(G_{\mathcal{P}})$ and that $\mathsf{t}(H) \le \mathsf{D}(G_{\mathcal{P}})$.

We proceed in two steps. First we verify that

$$\mathsf{c}_{\mathrm{mon}}(G_{\mathcal{P}}) = \max\{\mathsf{c}_{\mathrm{eq}}(G_{\mathcal{P}}), \quad \mathsf{c}_{\mathrm{adj}}(G_{\mathcal{P}})\} \le r + \alpha.$$

Since

$$\mathcal{A}(\sim_{\mathcal{B}(G_{\mathcal{P}}),\mathrm{eq}}) = \{(-V, -V), (V, V), (U_\nu, U_\nu), (-U_\nu, -U_\nu) : \nu \in [0, r]\},$$

it follows that $\mathsf{c}_{\mathrm{eq}}(G_{\mathcal{P}}) = 0$. If $A_{r+\alpha-2} = \{x \in \mathsf{Z}(G_{\mathcal{P}}) : |x| - (r+\alpha-2) \in \mathsf{L}(\pi(x))\}$, then $\mathrm{Min}(A_{r+\alpha-2}) = \{U_0^\alpha U_1 \cdot \ldots \cdot U_r\}$, and hence $\mathsf{c}_{\mathrm{adj}}(G_{\mathcal{P}}) \le r + \alpha$ by Lemma 2.1.

In order to show that $\mathsf{t}(H) \le \mathsf{D}(G_{\mathcal{P}})$, we must verify the following assertion (see [Geroldinger and Hassler 2008, Theorem 3.6]).

(A) *Let $j \in \mathbb{N}$ and $w, w_1, \ldots, w_j \in \mathcal{A}(H)$ be such that w divides the product $w_1 \cdot \ldots \cdot w_j$ yet w divides no proper subproduct of $w_1 \cdot \ldots \cdot w_j$. Then*

$$\min \mathsf{L}(w^{-1} w_1 \cdot \ldots \cdot w_j) \le \mathsf{D}(G_{\mathcal{P}}) - 1.$$

Proof of (A). We use the transfer homomorphism $\boldsymbol{\beta} : H \to \mathcal{B}(G_{\mathcal{P}})$ as defined in Lemma 3.4. Set $W = \boldsymbol{\beta}(w)$ and $W_i = \boldsymbol{\beta}(w_i)$ for each $i \in [1, j]$. Then $j \le |W|$ and $W, W_1, \ldots, W_j \in \mathcal{A}(G_{\mathcal{P}})$. Clearly

$$\min \mathsf{L}(w^{-1} w_1 \cdot \ldots \cdot w_j) \le \max \mathsf{L}(W^{-1} W_1 \cdot \ldots \cdot W_j) \le \frac{|W_1 \cdot \ldots \cdot W_j| - |W|}{2}.$$

Thus, if $|W| = 2$, then

$$\min \mathsf{L}(w^{-1} w_1 \cdot \ldots \cdot w_j) \le \frac{|W_1| + |W_2| - |W|}{2} \le \mathsf{D}(G_{\mathcal{P}}) - 1.$$

It remains to consider the case $W \in \{-V, V\}$, and by symmetry we may suppose that $W = V$. If $|W_1| = \cdots = |W_j| = 2$, then $j = |V|$ and $w^{-1} w_1 \cdot \ldots \cdot w_j \in \mathcal{A}(H)$. Suppose there is $\nu \in [1, j]$, say $\nu = 1$, such that $\boldsymbol{\beta}(w_1) \in \{-V, V\}$. Since w does not divide a subproduct of $w_1 \cdot \ldots \cdot w_j$ and $\gcd(V, -V) = 1$, it follows that $\boldsymbol{\beta}(w_1) = V$. Then $\mathsf{L}(w^{-1} w_1 \cdot \ldots \cdot w_j) = \mathsf{L}(W^{-1} W_1 \cdot \ldots \cdot W_j) = \mathsf{L}(W_2 \cdot \ldots \cdot W_j)$ and hence

$$\min \mathsf{L}(w^{-1} w_1 \cdot \ldots \cdot w_j) \le j - 1 \le |V| - 1 = \mathsf{D}(G_{\mathcal{P}}) - 1. \qquad \square$$

The next corollary again reveals that certain arithmetical phenomena characterize certain algebraic properties of the class group.

Corollary 6.5. *Let H be a Krull monoid as in Theorem 6.4 with class group G and set $G_{\mathcal{P}}$ of classes containing prime divisors. Then $r+1$ is the minimum of all $s \in \mathbb{N}$ having the following property*:

(P) *There are absolutely irreducible elements $w_1, \ldots, w_s \in \mathcal{A}(H)$ such that $2, \mathsf{D}(G_{\mathcal{P}}) \in \mathsf{L}(w_1^{k_1} \cdot \ldots \cdot w_s^{k_s})$ for some $(k_1, \ldots, k_s) \in \mathbb{N}_0^s$.*

Proof. First we verify that $r+1$ satisfies property (P). For $i \in [0, s]$, let $p_i \in \mathcal{P} \cap e_i$ and $q_i \in \mathcal{P} \cap (-e_i)$ and set $w_i = p_i q_i$. Then $w_0, \ldots, w_s$ are absolutely irreducible elements and, by Theorem 6.4, it follows that $2, \mathsf{D}(G_{\mathcal{P}}) \in \mathsf{L}(w_0^{\alpha} w_1 \cdot \ldots \cdot w_r)$.

Conversely, let $s \in \mathbb{N}$, and let $w_1, \ldots, w_s$ and $k_1, \ldots, k_s$ be as above. For $i \in [1, s]$, we set $W_i = \boldsymbol{\beta}(w_i)$. Since $\rho(G_{\mathcal{P}}) = \mathsf{D}(G_{\mathcal{P}})/2$ and $2, \mathsf{D}(G_{\mathcal{P}}) \in \mathsf{L}(W_1^{k_1} \cdot \ldots \cdot W_s^{k_s})$, it follows that $\sum_{i=1}^{s} k_i |W_i| = \mathsf{D}(G_{\mathcal{P}})$, $|W_1| = \cdots = |W_s| = 2$ and $W_i = (-g_i) g_i$ for $i \in [1, s]$, and that $S = g_1^{k_1} \cdot \ldots \cdot g_s^{k_s} \in \mathcal{A}(G_{\mathcal{P}})$. Now Theorem 6.4 implies $S = (-e_0)^{\alpha} e_1 \cdot \ldots \cdot e_r$, whence $\{W_1, \ldots, W_s\} = \{(-e_0)e_0, \ldots, (-e_r)e_r\}$. Thus

$$|\{w_1, \ldots, w_s\}| \geq |\{W_1, \ldots, W_s\}| = r+1,$$

and so $r+1$ is minimal with Property (P). □

We now begin collecting information in order to study the arithmetic of the Krull monoid presented in Example 4.21. In spite of the simple geometric structure of $G_{\mathcal{P}}$ (the set consists of the vertices of the unit cube and their negatives), the arithmetic of this Krull monoid is highly complex. We get only very limited information. Nevertheless, this will be sufficient to give an arithmetical characterization.

Lemma 6.6. *Let G be an abelian group and let $(e_n)_{n \geq 1}$ be a family of independent elements each having infinite order. For $r \in \mathbb{N}$, set*

$$G_r^+ = \{a_1 e_1 + \cdots + a_r e_r : a_1, \ldots, a_r \in [0, 1]\},$$
$$G_r^- = -G_r^+, \quad G_r = G_r^+ \cup G_r^-.$$

(1) *Let $s \in [2, r]$, $f_0 = e_1 + \cdots + e_s$, and $f_i = f_0 - e_i$ for all $i \in [1, s]$. Then $(f_1, \ldots, f_s)$ is independent, $f_1 + \cdots + f_s = (s-1) f_0$, and*

$$\Delta(\{f_0, \ldots, f_s, -f_0, \ldots, -f_s\}) = \{2s-3\}.$$

(2) *Let $s \in [3, r]$, $f_0 = e_1 + \cdots + e_s$, $f_i = f_0 - e_i$ for each $i \in [1, s-1]$, and set $f_s' = -e_s$. Then $(f_1, \ldots, f_{s-1}, f_s')$ is independent,*

$$f_1 + \cdots + f_{s-1} + f_s' = (s-2) f_0,$$

and

$$\Delta(\{f_0, \ldots, f_{s-1}, f_s', -f_0, \ldots, -f_{s-1}, -f_s'\}) = \{2s-4\}.$$

(3) *If $s \leq [1, r-1]$, then $\mathsf{D}(G_r) \geq \mathsf{D}(G_s) + \mathsf{D}(G_{r-s}) - 1$. In particular, $\mathsf{D}(G_1) = 2$ and $\mathsf{D}(G_r) > \mathsf{D}(G_{r-1})$ for $r \geq 2$.*

Proof. Since $(e_1,\dots,e_s)$ is a basis, there is a matrix A_s with $(f_1,\dots,f_s)=(e_1,\dots,e_s)A_s$. Since $\det(A_s)\neq 0$, it follows that $(f_1,\dots,f_s)$ is independent. By definition, we have $f_1+\dots+f_s=(s-1)f_0$. The assertion on the set of distances then follows from Theorem 6.4 and we have proved (1).

We now consider (2). Note that $f_s'=f_s-f_0$. Using (1) we infer that

$$0=(f_1+\dots+f_s)-(s-1)f_0=(f_1+\dots+f_{s-1})+(f_s-f_0)-(s-2)f_0,$$

and hence $f_1+\dots+f_{s-1}+f_s'=(s-2)f_0$. Since $(f_1,\dots,f_{s-1},-e_s)=(e_1,\dots,e_s)B_s$ for some matrix B_s with $\det(B_s)=(-1)^{2s}\det(A_{s-1})\neq 0$, it follows that $(f_1,\dots,f_{s-1},f_s')$ is independent. The assertion on the set of distances follows from Theorem 6.4.

It is clear that $\mathsf{D}(G_1)=2$ and that $\mathsf{D}(G_r)>\mathsf{D}(G_{r-1})$ whenever $r\geq 2$. To prove the remaining statements of (3), suppose that $s\in[1,r-1]$. After a change of notation, we may suppose that $G_{r-s}\subset\langle e_{s+1},\dots,e_r\rangle$ such that $\langle G_r\rangle=\langle G_s\rangle\oplus\langle G_{r-s}\rangle$. If $U=a_1\cdot\ldots\cdot a_k\in\mathcal{A}(G_s)$ with $k=\mathsf{D}(G_s)$ and $V=b_1\cdot\ldots\cdot b_l\in\mathcal{A}(G_{r-s})$ with $l=\mathsf{D}(G_{r-s})$, then $W=(a_1+b_1)\cdot a_2\cdot\ldots\cdot a_k b_2\cdot\ldots\cdot b_l\in\mathcal{A}(G_r)$, and hence $\mathsf{D}(G_r)\geq |W|=k+l-1=\mathsf{D}(G_s)+\mathsf{D}(G_{r-s})-1$. □

In Theorem 6.7 we restrict to class groups of rank $r\geq 3$ because when $r\leq 2$ we are in the setting of Theorem 6.4 where we have precise information about arithmetical invariants. For $r\in\mathbb{N}_0$, we denote by F_r the r-th *Fibonacci number*. That is, $\mathsf{F}_0=0$, $\mathsf{F}_1=1$, and $\mathsf{F}_r=\mathsf{F}_{r-1}+\mathsf{F}_{r-2}$ for all $r\geq 2$.

Theorem 6.7. *Let H be a Krull monoid with free abelian class group G of rank $r\geq 3$ and let $G_{\mathcal{P}}\subset G$ denote the set of classes containing prime divisors. Suppose that there is a basis $(e_1,\dots,e_r)$ of G such that $G_{\mathcal{P}}^{\bullet}=G_{\mathcal{P}}^{+}\cup G_{\mathcal{P}}^{-}$, where*

$$G_{\mathcal{P}}^{+}=\{\epsilon_1e_1+\dots+\epsilon_re_r:\epsilon_1,\dots,\epsilon_r\in[0,1]\}\quad\textit{and}\quad G_{\mathcal{P}}^{-}=-G_{\mathcal{P}}^{+}.$$

(1) $\mathsf{F}_{r+2}\leq\mathsf{D}(G_{\mathcal{P}})$.

(2) $\mathsf{c}(H)\leq\omega(H)=\mathsf{D}(G_{\mathcal{P}})$, $\rho(H)=\mathsf{D}(G_{\mathcal{P}})/2$, *and* $\rho_{2k}(H)=k\mathsf{D}(G_{\mathcal{P}})$ *for each* $k\in\mathbb{N}$.

(3) $[1,2r-3]\subset\Delta^*(H)\subset\Delta(H)\subset[1,\mathsf{c}(H)-2]$.

Proof. See [Baeth et al. 2014] for the proof of assertion (1). Assertion (2) follows from Proposition 6.2.

Note that for every $s\in[2,r]$ we have $2s-3\in\Delta^*(H)$ and, by Lemma 6.6, for all $s\in[3,r]$ we have $2s-4\in\Delta^*(H)$. This implies that the interval $[1,2r-3]$ is contained in $\Delta^*(H)$, giving (3). □

The third class of Krull monoids studied in this subsection are Krull monoids with finite cyclic class group having prime divisors in each class. Thus Theorem 6.8 describes the arithmetic of the monoids constructed in Theorem 5.5. Holomorphy rings in global fields are Krull monoids with finite class group and prime divisors in all classes. For this reason this class of Krull monoids has received a great deal of attention.

Theorem 6.8. *Let H be a Krull monoid with finite cyclic class group G of order $|G| = n \ge 3$, and suppose that every class contains a prime divisor. Then*:

(1) $\mathsf{c}(H) = \omega(H) = \mathsf{D}(G) = n$ *and* $\Delta(H) = [1, n-2]$.

(2) *For every $k \in \mathbb{N}$ the set $\mathcal{U}_k(H)$ is a finite interval*, *whence*

$$\mathcal{U}_k(H) = [\lambda_k(H), \rho_k(H)].$$

Moreover, *for all $l \in \mathbb{N}_0$ with $ln + j \ge 1$*,

$$\rho_{2k+j}(H) = kn + j \quad \text{for } j \in [0,1],$$

$$\lambda_{ln+j}(H) = \begin{cases} 2l + j & \text{for } j \in [0,1], \\ 2l + 2 & \text{for } j \in [2, n-1]. \end{cases}$$

(3) $\max \Delta^*(H) = n - 2 \quad$ *and* $\quad \max(\Delta^*(H) \setminus \{n-2\}) = \lfloor \frac{n}{2} \rfloor - 1$.

Proof. The proof of (1) can be found in [Geroldinger and Halter-Koch 2006, Theorem 6.7.1] and the proof of (3) can be found in [Geroldinger and Halter-Koch 2006, Theorem 6.8.12]. For (2) see [Geroldinger 2009, Corollary 5.3.2]. ☐

Much recent research is devoted to the arithmetic of Krull monoids discussed in Theorem 6.8. We briefly address some open questions. Let H be as above and suppose that $n \ge 5$. The precise values of $\mathsf{t}(H)$ and of $\mathsf{c}_{\text{mon}}(H)$ are unknown. It is easy to check that $\mathsf{D}(G) = n < \mathsf{t}(H)$ (in contrast to what we have in Theorem 6.4). For recent results on lower and upper bounds of the tame degree, see [Gao et al. 2015]. We remark that there is a standing conjecture that the monotone catenary degree is that $n = \mathsf{c}(H) = \mathsf{c}_{\text{mon}}(H)$ (this coincides what we have in Theorem 6.4; see [Geroldinger and Yuan 2013]). For recent progress on $\Delta^*(H)$ we refer to [Plagne and Schmid 2013].

Having at least a partial description of the arithmetic of the three monoids described in Theorems 6.4, 6.7, and 6.8, we now work to show that except for in a small number of exceptions, these monoids have vastly different arithmetic. After some preliminary work this distinction is made clear in Corollary 6.10.

Lemma 6.9. *Let G be an abelian group with finite total rank and let $G_0 \subset G$ be a subset with $G_0 = -G_0$. Suppose that $\mathcal{L}(G_0) = \mathcal{L}(C_n)$ for some $n \ge 5$. Then there exists an absolutely irreducible element $U \in \mathcal{A}(G_0)$ with $|U| = \mathsf{D}(G_0)$.*

Proof. First observe that $\mathsf{D}(G_0) = \rho_2(G_0) = \rho_2(C_n) = \mathsf{D}(C_n) = n$ and, by [Geroldinger and Halter-Koch 2006, Theorem 3.4.2], $\mathcal{A}(G_0)$ is finite, say $\mathcal{A}(G_0) =$

$\{U_1, -U_1, \dots, U_q, -U_q\}$. If $g \in C_n$ with $\mathrm{ord}(g) = n$, then for all $k \in \mathbb{N}$ we have

$$L_k = \{2k + \nu(n-2) : \nu \in [0, k]\} = \mathsf{L}\big(g^{nk}(-g)^{nk}\big) \in \mathcal{L}(C_n) = \mathcal{L}(G_0).$$

Since $\rho(L_k) = \rho(G_0) = \mathsf{D}(G_0)/2$, there exists, for every $k \in \mathbb{N}$, a tuple

$$(k_1, \dots, k_q) \in \mathbb{N}_0^{(q)}$$

such that $k_1 + \cdots + k_q = k$ and

$$L_k = \mathsf{L}\big((-U_1)^{k_1} U_1^{k_1} \cdot \ldots \cdot (-U_q)^{k_q} U_q^{k_q}\big).$$

Therefore there exists $\lambda \in [1, q]$ such that $\mathsf{L}\big((-U_\lambda)^k U_\lambda^k\big) = L_k$ for every $k \in \mathbb{N}$. Set $U = U_\lambda$ and note that for every $V \in \mathcal{A}(G_0)$ with $V \mid (-U)^k U^k$ for some $k \in \mathbb{N}$, it follows that $|V| \in \{2, n\}$. After changing notation if necessary, we may suppose that there is no $V \in \mathcal{A}(G_0)$ such that $|V| = n$, $\mathrm{supp}(V) \subsetneq \mathrm{supp}(U)$, and $V \mid U^k(-U)^k$ for some $k \in \mathbb{N}$.

In order to show that U is absolutely irreducible, it remains to verify that the torsion-free rank of $\langle \mathrm{supp}(U) \rangle$ is $|\mathrm{supp}(U)| - 1$. Assume to the contrary that there exist $t \in [2, |\mathrm{supp}(U)| - 1]$ and $g_1, \dots, g_t \in \mathrm{supp}(U)$ which are linearly dependent. Then there are $s \in [1, t]$, $m_1, \dots, m_s \in \mathbb{N}$, and $m_{s+1}, \dots, m_t \in -\mathbb{N}$ such that

$$m_1 g_1 + \cdots + m_s g_s + (-m_{s+1})(-g_{s+1}) + \cdots + (-m_t)(-g_t) = 0.$$

Then

$$V = g_1^{m_1} \cdot \ldots \cdot g_s^{m_s} (-g_{s+1})^{-m_{s+1}} \cdot \ldots \cdot (-g_t)^{-m_t} \in \mathcal{B}(G_0).$$

Without restriction we may suppose that the above equation is minimal and that $V \in \mathcal{A}(G_0)$. Since $V \mid U^k(-U)^k$ for some $k \in \mathbb{N}$ and $|V| > 2$, we obtain a contradiction to the minimality of $\mathrm{supp}(U)$. □

The following corollary highlights that the observed arithmetical phenomena in our case studies — Theorems 6.4, 6.7, and 6.8 — are characteristic for the respective Krull monoids. In particular, this illustrates that the structure of direct-sum decompositions over the one-dimensional Noetherian local rings with finite representation type studied in Section 4 can be quite different from the structure of direct-sum decompositions over the two-dimensional Noetherian local Krull domains with finite representation type studied in Section 5. As characterizing tools we use the system of sets of lengths along with the behavior of absolutely irreducible elements.

Corollary 6.10. *For $i \in [1, 3]$, let H_i and H_i' be Krull monoids with class groups G_i and G_i'. Further suppose that*

- *G_1 and G_1' are finitely generated and torsion-free of rank r_1 and r_1' with sets of classes containing prime divisors as in Theorem 6.4 (with parameters $\alpha, \alpha' \in \mathbb{N}$ such that $\alpha + r_1 \geq \alpha' + r_1' > 2$).*

- *G_2 and G_2' are finitely generated and torsion-free of rank $r_2 \ge r_2' \ge 3$ with sets of classes containing prime divisors as in Theorem 6.7.*
- *G_3 and G_3' are finite cyclic of order $|G_3| \ge |G_3'| \ge 5$ such that every class contains a prime divisor.*

Then:

(1) *$\mathcal{L}(H_1) = \mathcal{L}(H_1')$ if and only if $r_1 + \alpha = r_1' + \alpha'$. If this holds, then the arithmetic behavior of the absolutely irreducible elements of H_1 and H_1' coincide in the sense of Corollary 6.5 if and only if $r_1 = r_1'$.*

(2) *$\mathcal{L}(H_2) = \mathcal{L}(H_2')$ if and only if $r_2 = r_2'$.*

(3) *$\mathcal{L}(H_3) = \mathcal{L}(H_3')$ if and only if $|G_3| = |G_3'|$.*

(4) *$\mathcal{L}(H_1) \ne \mathcal{L}(H_2)$ and $\mathcal{L}(H_1) \ne \mathcal{L}(H_3)$.*

(5) *For $i \in [2, 3]$, let s_i denote the maximal number of absolutely irreducible elements $u_1, \ldots, u_{s_i} \in H_i$ such that $2 \in \mathsf{L}(u_1 \cdot \ldots \cdot u_{s_i})$. Then either $\mathcal{L}(H_2) \ne \mathcal{L}(H_3)$ or $s_2 \ne s_3$.*

Proof. The if and only if statement in (1) follows immediately from Theorem 6.4. Suppose that $\mathcal{L}(H_1) = \mathcal{L}(H_1')$. Then the assertion in (1) on the arithmetic behavior of absolutely irreducible elements follows from Corollary 6.5.

To prove (2), first note that one implication is clear, both for H_2 and H_3. Suppose that $\mathcal{L}(H_2) = \mathcal{L}(H_2')$, and let $G_{\mathcal{P}} \subset G_2$ and $G_{\mathcal{P}}' \subset G_2'$ denote the set of classes containing prime divisors. Theorem 6.7 implies that

$$\mathsf{D}(G_{\mathcal{P}}) = \rho_2(H) = \rho_2(H') = \mathsf{D}(G_{\mathcal{P}}'),$$

and thus Lemma 6.6 implies $r_2 = r_2'$. Now consider (3). If $\mathcal{L}(H_3) = \mathcal{L}(H_3')$, then Theorem 6.8 implies that

$$|G_3| - 2 = \max \Delta(H_3) = \max \Delta(H_3') = |G_3'| - 2.$$

For (4), note that $\mathcal{L}(H_1)$ is distinct from both $\mathcal{L}(H_2)$ and $\mathcal{L}(H_3)$ since

$$|\Delta(H_1)| = 1, \quad |\Delta(H_2)| > 1, \quad |\Delta(H_3)| > 1.$$

For (5) we assume that $\mathcal{L}(H_2) = \mathcal{L}(H_3)$ and let $G_{\mathcal{P}} \subset G_2$ denote the set of classes containing prime divisors. Theorems 6.7 and 6.8 imply that

$$\mathsf{D}(G_{\mathcal{P}}) = \rho_2(H_2) = \rho_2(H_3) = |G_3|.$$

By Proposition 6.3 we obtain that $\mathsf{D}(G_{\mathcal{P}}) = s_2$. Now assume to the contrary that $s_2 = s_3$. If $|G_3| = n$, then there are absolutely irreducible elements $u_1, \ldots, u_n$ and atoms $v_1, v_2 \in \mathcal{A}(H_3)$ such that $v_1 v_2 = u_1 \cdot \ldots \cdot u_n$. Without restriction, we suppose H_3 is reduced and we consider a divisor theory $H \hookrightarrow \mathcal{F}(\mathcal{P})$. Since a minimal zero-sum sequence of length n over G_3 consists of one element of order n repeated

n times, the factorization of the atoms $v_1, v_2, u_1, \dots, u_n$ in $\mathcal{F}(\mathcal{P})$ must have the following form: $v_1 = p_1 \cdot \ldots \cdot p_n$, $v_2 = q_1 \cdot \ldots \cdot q_n$, and $u_i = p_i q_i$ for all $i \in [1, n]$, where $p_1, \dots, p_n, q_1, \dots, q_n \in \mathcal{P}$, $[p_1] = \dots = [p_n] \in G_3$, and $[q_1] = \dots = [q_n] = [-p_1]$. But [Geroldinger and Halter-Koch 2006, Proposition 7.1.5] implies that the elements $u_1, \dots, u_n$ are not absolutely irreducible, a contradiction. □

Remark 6.11. Let H_2 and H_3 be as in Corollary 6.10. We set $n = |G_3|$, $r = r_2$, and let $G_{\mathcal{P},r} \subset G_2$ denote the set of classes containing prime divisors. Assume that $\mathcal{L}(H_2) = \mathcal{L}(H_3)$. Then

$$\mathsf{F}_{r+2} \le \mathsf{D}(G_{\mathcal{P},r}) = \rho_2(H_2) = \rho_2(H_3) = n.$$

That is, the orders of the cyclic groups for which $\mathcal{L}(H_2) = \mathcal{L}(H_3)$ grow faster than the sequence of Fibonacci numbers. We conjecture that $\mathcal{L}(H_2)$ and $\mathcal{L}(H_3)$ are always distinct but have not further investigated this (rather delicate combinatorial) problem which would require a more detailed investigation of $\mathsf{D}(G_{\mathcal{P},r})$.

Now suppose that H is a Krull monoid with class group G such that every class contains a prime divisor. If $\mathcal{L}(H) = \mathcal{L}(H_3)$, then following Theorem 6.8, one can show that G is isomorphic to the finite cyclic group G_3 (see [Geroldinger 2009, Corollary 5.3.3]). Therefore sets of lengths characterize Krull monoids with finite cyclic class group having the property that every class contains a prime divisor.

6C. ***Small sets $G_{\mathcal{P}}$ of classes containing prime divisors and limits of arithmetical characterizations.*** In this final subsection we study the arithmetic of Krull monoids having small sets of classes containing prime divisors. This study pertains to the monoids of Theorem 4.12, Example 4.19, Example 4.20, and Theorem 5.5. The most striking phenomenon here is that these systems of sets of lengths are additively closed (see Proposition 6.14). As a consequence, if $\mathcal{L}(H)$ is such a system and H' is a monoid with $\mathcal{L}(H') \subset \mathcal{L}(H)$, then $\mathcal{L}(H \times H') = \mathcal{L}(H)$ (see Example 4.23, Example 4.24, and Corollary 6.15). These phenomena are in strong contrast to the results in the previous subsection, and they show up natural limits for obtaining arithmetical characterization results. Recall that, for $l \in \mathbb{N}_0$ and $d \in \mathbb{N}$, $P_l(d) = \{0, d, \dots, ld\}$.

Proposition 6.12. *Let H be a Krull monoid with infinite cyclic class group G and suppose that*

$$G_{\mathcal{P}} = \{-2e, -e, 0, e, 2e\} \subset G = \langle e \rangle$$

is the set of classes containing prime divisors. Then there is a transfer homomorphism $\theta : H \to \mathcal{B}(C_3)$, and hence

$$\mathcal{L}(H) = \mathcal{L}(C_3) = \mathcal{L}(C_2 \oplus C_2) = \{y + 2k + P_k(1) : y,\ k \in \mathbb{N}_0\}.$$

Moreover, $\mathcal{L}(H)$ coincides with the system of sets of lengths of the Krull monoid studied in Theorem 6.4 with parameters $r = 2$ and $\alpha = 1$.

Proof. By Lemma 3.4 there is a transfer homomorphism $\boldsymbol{\beta} : H \to \mathcal{B}(G_{\mathcal{P}})$. Since the composition of two transfer homomorphisms is a transfer homomorphism, it is sufficient to show that there is a transfer homomorphism $\theta' : \mathcal{B}(G_{\mathcal{P}}) \to \mathcal{B}(C_3)$. Write $C_3 = \{0, g, -g\}$. Since

$$\mathcal{B}(G_{\mathcal{P}}) = \mathcal{F}(\{0\}) \times \mathcal{B}(G_{\mathcal{P}}^{\bullet}) \quad \text{and} \quad \mathcal{B}(C_3) = \mathcal{F}(\{0\}) \times \mathcal{B}(\{-g, g\}),$$

it suffices to show that there is a transfer homomorphism $\theta : \mathcal{B}(G_{\mathcal{P}}^{\bullet}) \to \mathcal{B}(\{-g, g\})$. In this case, $\mathcal{L}(H) = \mathcal{L}(G_{\mathcal{P}}) = \mathcal{L}(C_3)$. Moreover, $\mathcal{L}(C_3) = \mathcal{L}(C_2 \oplus C_2)$ has the form given in [Geroldinger and Halter-Koch 2006, Theorem 7.3.2] and this coincides with the system of sets of lengths in Theorem 6.4, provided $(r, \alpha) = (2, 1)$.

Note that $\mathcal{A}(G_{\mathcal{P}}^{\bullet}) = \{V, -V, U_1, U_2\}$, where $V = e^2(-2e)$, $U_1 = (-e)e$, and $U_2 = (-2e)(2e)$, and $\mathcal{A}(\{-g, g\}) = \{\overline{V}, -\overline{V}, \overline{U}\}$, where $\overline{V} = g^3$ and $\overline{U} = (-g)g$. Then there is a monoid epimorphism

$$\widetilde{\theta} : \mathcal{F}(G_{\mathcal{P}}^{\bullet}) \to \mathcal{F}(\{-g, g\})$$

satisfying $\widetilde{\theta}(e) = \widetilde{\theta}(-2e) = g$ and $\widetilde{\theta}(-e) = \widetilde{\theta}(2e) = -g$. If

$$A = e^{k_1}(-e)^{k_1'}(2e)^{k_2}(-2e)^{k_2'} \in \mathcal{F}(G_{\mathcal{P}}^{\bullet}) \quad \text{with} \quad k_1, k_1', k_2, k_2' \in \mathbb{N}_0,$$

then $A \in \mathcal{B}(G_{\mathcal{P}}^{\bullet})$ if and only if $k_1 - k_1' + 2(k_2 - k_2') = 0$. If this holds, then $k_1 + k_2' - (k_1' + k_2) \equiv 0 \mod 3$ and hence

$$\widetilde{\theta}(A) = g^{k_1 + k_2'}(-g)^{k_1' + k_2} \in \mathcal{B}(\{-g, g\}).$$

Thus $\theta = \widetilde{\theta}\,|_{\mathcal{B}(G_{\mathcal{P}}^{\bullet})} : \mathcal{B}(G_{\mathcal{P}}^{\bullet}) \to \mathcal{B}(\{-g, g\})$ is a monoid epimorphism satisfying $\theta(V) = \overline{V}$, $\theta(-V) = -\overline{V}$, $\theta(U_1) = \theta(U_2) = \overline{U}$ and $\theta^{-1}(1) = \{1\} = \mathcal{B}(G_{\mathcal{P}}^{\bullet})^{\times}$.

Thus in order to show that θ is a transfer homomorphism, it remains to verify Property (T2). Let $A \in \mathcal{B}(G_{\mathcal{P}}^{\bullet})$ be as above and suppose that

$$\theta(A) = \widetilde{B}\widetilde{C}$$

with $\widetilde{B}, \widetilde{C} \in \mathcal{B}(\{-g, g\})$ and $\widetilde{B} = g^m(-g)^{m'}$ such that $m \in [0, k_1 + k_2']$, $m' \in [0, k_1' + k_2]$ and $m \equiv m' \mod 3$. Our goal is to find $B, C \in \mathcal{B}(G_{\mathcal{P}}^{\bullet})$ such that $A = BC$, $\theta(B) = \widetilde{B}$, and $\theta(C) = \widetilde{C}$. Clearly it is sufficient to find $B \in \mathcal{B}(G_{\mathcal{P}}^{\bullet})$ with $B \mid A$ and $\theta(B) = \widetilde{B}$, that is, to find parameters

$$m_1 \in [0, k_1], \quad m_1' \in [0, k_1'], \quad m_2 \in [0, k_2], \quad m_2' \in [0, k_2'],$$

such that

(C1) $\qquad m_1 + m_2' = m, \quad m_1' + m_2 = m', \quad m_1 - m_1' + 2(m_2 - m_2') = 0.$

To do so we proceed by induction on $|\widetilde{B}|$. If $|\widetilde{B}| = |A|$, then

$$k_1 + k_1' + k_2 + k_2' = |A| = |\widetilde{B}| = m + m',$$

and hence $m = k_1 + k_1'$ and $m' = k_1' + k_2$. Thus we set

$$m_1 = k_1, \quad m_1' = k_1', \quad m_2 = k_2, \quad m_2' = k_2',$$

and the assertion is satisfied with $B = A$. Suppose now that the quadruple (m_1, m_1', m_2, m_2') satisfies (C1) with respect to the pair (m, m'). Dividing $\widetilde{B}$ by an atom of $\mathcal{B}(\{-g, g\})$ (if possible) shows that we must verify that there are solutions to (C1) with respect to each of the pairs $(m-1, m'-1)$, $(m-3, m')$, and $(m, m'-3)$ in $N_0^{(2)}$. One checks respectively that at least one of the following quadruples satisfy (C1).

- $(m_1-1, m_1'-1, m_2, m_2')$ or $(m_1, m_1', m_2-1, m_2'-1)$;
- $(m_1-2, m_1', m_2, m_2'-1)$ or $(m_1-3, m_1'-1, m_2+1, m_2')$;
- $(m_1, m_1'-2, m_2-1, m_2')$ or $(m_1-1, m_1'-3, m_2, m_2'+1)$.

Now the assertion follows by the induction hypothesis. □

Proposition 6.13. *Let H be a Krull monoid with free abelian class group G of rank 2. Let (e_1, e_2) be a basis of G and suppose that*

$$G_{\mathcal{P}} = \{0, e_1, e_2, 2e_2, e_1+2e_2, -e_1, -e_2, -2e_2, -e_1-2e_2\}$$

is the set of classes containing prime divisors. Then there is a transfer homomorphism $\theta : H \to \mathcal{B}(C_4)$ and hence

$$\begin{aligned}\mathcal{L}(H) &= \mathcal{L}(C_4)\\ &= \{y+k+1+P_k(1) : y, k \in \mathbb{N}_0\} \cup \{y+2k+P_k(2) : y, k \in \mathbb{N}_0\}\\ &\supset \mathcal{L}(C_3).\end{aligned}$$

Proof. As in Proposition 6.12, it suffices to show that there is a transfer homomorphism $\theta : \mathcal{B}(G_{\mathcal{P}}^\bullet) \to \mathcal{B}(C_4^\bullet)$. Then $\mathcal{L}(H) = \mathcal{L}(G_{\mathcal{P}}) = \mathcal{L}(C_4)$ and $\mathcal{L}(C_4)$ has the form given in [Geroldinger and Halter-Koch 2006, Theorem 7.3.2]. Proposition 6.12 shows that $\mathcal{L}(C_3) \subset \mathcal{L}(C_4)$.

We note that $\mathcal{A}(G_{\mathcal{P}}^\bullet) = \{W, -W, V_1, -V_1, V_2, -V_2, U_1, U_2, U_3, U_4\}$, where

$$\begin{aligned}W &= e_1e_2e_2(-e_1-2e_2), & U_3 &= (-e_1-2e_2)(e_1+2e_2),\\ U_1 &= (-e_1)e_1, & V_1 &= e_1(2e_2)(-e_1-2e_2),\\ U_2 &= (-e_2)e_2, & V_2 &= e_2e_2(-2e_2),\end{aligned}$$
$$U_4 = (-2e_2)(2e_2).$$

We set $C_4 = \{0, g, 2g, -g\}$ and observe that $\mathcal{A}(C_4^\bullet) = \{\overline{W}, -\overline{W}, \overline{V}, -\overline{V}, \overline{U}_1, \overline{U}_2\}$, where

$$\overline{W} = g^4, \quad \overline{V} = g^2(2g), \quad \overline{U}_1 = (-g)g \quad \text{and} \quad \overline{U}_2 = (2g)(2g).$$

There is a monoid epimorphism $\widetilde{\theta} : \mathcal{F}(G_{\mathcal{P}}^\bullet) \to \mathcal{F}(C_4^\bullet)$ satisfying

$$\begin{aligned} \widetilde{\theta}(e_1) = \widetilde{\theta}(e_2) = \widetilde{\theta}(-e_1 - 2e_2) = g, \\ \widetilde{\theta}(-e_1) = \widetilde{\theta}(-e_2) = \widetilde{\theta}(e_1 + 2e_2) = -g, \\ \widetilde{\theta}(2e_2) = \widetilde{\theta}(-2e_2) = 2g. \end{aligned}$$

If

$$A = e_1^{k_1}(-e_1)^{k_1'} e_2^{k_2}(-e_2)^{k_2'}(2e_2)^{k_3}(-2e_2)^{k_3'}(e_1+2e_2)^{k_4}(-e_1-2e_2)^{k_4'} \in \mathcal{F}(G_{\mathcal{P}}^\bullet),$$

with $k_1, k_1', \ldots, k_4, k_4' \in \mathbb{N}_0$, then $A \in \mathcal{B}(G_{\mathcal{P}}^\bullet)$ if and only if

$$k_1 - k_1' + k_4 - k_4' = 0 \quad \text{and} \quad k_2 - k_2' + 2k_3 - 2k_3' + 2k_4 - 2k_4' = 0.$$

If this holds, then

$$k_1 - k_1' + k_2 - k_2' - (k_4 - k_4') + 2k_3 + 2k_3' \equiv 0 \mod 4,$$

and hence

$$\widetilde{\theta}(A) = g^{k_1+k_2+k_4'}(-g)^{k_1'+k_2'+k_4}(2g)^{k_3+k_3'} \in \mathcal{B}(C_4^\bullet).$$

Thus $\theta = \widetilde{\theta}\,|_{\mathcal{B}(G_{\mathcal{P}}^\bullet)} : \mathcal{B}(G_{\mathcal{P}}^\bullet) \to \mathcal{B}(C_4^\bullet)$ is a monoid epimorphism satisfying

$$\begin{aligned} &\theta(W) = \overline{W}, && \theta(-V_1) = \theta(-V_2) = -\overline{V}, \\ &\theta(-W) = -\overline{W}, && \theta(U_1) = \theta(U_2) = \theta(U_3) = \overline{U}_1, \\ &\theta(V_1) = \theta(V_2) = \overline{V}, && \theta(U_4) = \overline{U}_2, \end{aligned}$$
$$\theta^{-1}(1) = \{1\} = \mathcal{B}(G_{\mathcal{P}}^\bullet)^\times.$$

Thus in order to show that θ is a transfer homomorphism, it remains to verify Property (T2). Let $A \in \mathcal{B}(G_{\mathcal{P}}^\bullet)$ be as above and suppose that

$$\theta(A) = \widetilde{B}\widetilde{C}$$

with $\widetilde{B}, \widetilde{C} \in \mathcal{B}(C_4^\bullet)$ and $\widetilde{B} = g^m(-g)^{m'}(2g)^{m''}$ such that

$$m \in [0, k_1 + k_2 + k_4'], \quad m' \in [0, k_1' + k_2' + k_4], \quad m'' \in [0, k_3 + k_3'],$$

and

$$m - m' + 2m'' \equiv 0 \mod 4.$$

Our goal is to find $B, C \in \mathcal{B}(G_{\mathcal{P}}^{\bullet})$ such that $A = BC$, $\theta(B) = \widetilde{B}$, and $\theta(C) = \widetilde{C}$. It will suffice to find $B \in \mathcal{B}(G_{\mathcal{P}}^{\bullet})$ with $B \mid A$ and $\theta(B) = \widetilde{B}$. Thus we must find parameters

$$m_\nu \in [0, k_\nu] \quad \text{and} \quad m'_\nu \in [0, k'_\nu] \qquad \text{for } \nu \in [1,4],$$

such that

$$\text{(C2)} \qquad \begin{gathered} m_1 + m_2 + m'_4 = m, \quad m'_1 + m'_2 + m_4 = m', \quad m_3 + m'_3 = m'', \\ m_1 - m'_1 + m_4 - m'_4 = 0, \quad m_2 - m'_2 + 2m_3 - 2m'_3 + 2m_4 - 2m'_4 = 0. \end{gathered}$$

We proceed by induction on $|\widetilde{B}| = m + m' + m''$. If $|\widetilde{B}| = |A|$, then we set $m_\nu = k_\nu$ and $m'_\nu = k'_\nu$ for all $\nu \in [1,4]$, and the assertion is satisfied with $B = A$. Suppose now that the octuplet $(m_1, m'_1, \ldots, m_4, m'_4)$ satisfies (C2) with respect to the triple (m, m', m''). Dividing $\widetilde{B}$ by an element of $\mathcal{A}(C_4^{\bullet})$ (if possible) shows that we must verify that there are solutions to (C2) with respect to each of the triples

$$\begin{gathered} (m-1, m'-1, m''), \quad (m-2, m', m''-1), \quad (m, m'-2, m''-1), \\ (m, m', m''-2), \quad (m-4, m', m''), \quad (m, m'-4, m''), \end{gathered}$$

provided that they lie in $\mathbb{N}_0^{(8)}$. As in proof of the previous proposition, one finds the required solutions and hence the assertion follows by the induction hypothesis. □

Let $\mathcal{L}$ be a family of subsets of $\mathbb{Z}$. We say that $\mathcal{L}$ is *additively closed* if the sumset $L + L' \in \mathcal{L}$ for all $L, L' \in \mathcal{L}$.

Proposition 6.14. *Let G be a finite cyclic group. Then $\mathcal{L}(G)$ is additively closed if and only if $|G| \le 4$.*

Proof. We suppose that $|G| = n$ and distinguish four cases.

First assume that $n \le 2$. Since $\mathcal{B}(G)$ is factorial, it follows that

$$\mathcal{L}(G) = \{\{m\} : m \in \mathbb{N}_0\},$$

which is obviously additively closed.

Next assume that $n = 3$. By Proposition 6.12 we have

$$\mathcal{L}(C_3) = \{y + 2k + P_k(1) : y, k \in \mathbb{N}_0\}.$$

If $y_1, y_2, k_1, k_2 \in \mathbb{N}_0$. Then

$$\begin{aligned} &(y_1 + 2k_1 + P_{k_1}(1)) + (y_2 + 2k_2 + P_{k_2}(1)) \\ &\qquad = (y_1 + y_2) + 2(k_1 + k_2) + P_{k_1+k_2}(1) \in \mathcal{L}(C_3), \end{aligned}$$

and hence $\mathcal{L}(C_3)$ is additively closed.

Now assume that $n = 4$. By Proposition 6.13 we have

$$\mathcal{L}(C_4) = \{y + k + 1 + P_k(1) : y,\, k \in \mathbb{N}_0\} \cup \{y + 2k + P_k(2) : y, k \in \mathbb{N}_0\}.$$

Clearly, the sumset of two sets of the first form is of the first form again, and the sumset of two sets of the second form again the second form. Thus it remains to consider the sumset $L_1 + L_2$ where L_1 has the first form, L_2 has the second form, and both L_1 and L_2 have more than one element. If $y_1, y_2 \in \mathbb{N}_0$ and $k_1, k_2 \in \mathbb{N}$, then

$$(y_1 + k_1 + 1 + P_{k_1}(1)) + (y_2 + 2k_2 + P_{k_2}(2)) \\ = (y_1 + y_2) + (k_1 + 2k_2) + 1 + P_{k_1+2k_2}(1) \in \mathcal{L}(C_4).$$

Finally, assume that $n \geq 5$ and assume to the contrary that $\mathcal{L}(G)$ is additively closed. Let $d \in [1, n-2]$. Then $\{2, d+2\} \in \mathcal{L}(G)$ by [Geroldinger and Halter-Koch 2006, Theorem 6.6.2], and hence the k-fold sumset

$$\{2, d+2\} + \cdots + \{2, d+2\} = 2k + P_k(d)$$

lies in $\mathcal{L}(G)$ for all $k \in \mathbb{N}$. Then [Geroldinger and Halter-Koch 2006, Corollary 4.3.16] implies that $n-3$ divides some $d \in \Delta^*(G)$. By Theorem 6.8 we have

$$\max \Delta^*(G) = n-2 \quad \text{and} \quad \max(\Delta^*(G) \setminus \{n-2\}) = \left\lfloor \frac{n}{2} \right\rfloor - 1,$$

a contradiction to $n \geq 5$. □

Corollary 6.15. (1) *Let H be an atomic monoid such that $\mathcal{L}(H)$ is additively closed, and let H' be an atomic monoid with $\mathcal{L}(H') \subset \mathcal{L}(H)$. Then*

$$\mathcal{L}(H \times H') = \mathcal{L}(H).$$

(2) *Let H be an atomic monoid with $\mathcal{L}(H) = \mathcal{L}(C_n)$ for $n \in [3, 4]$. For $k \in \mathbb{N}$ and $i \in [1, k]$, let H_i be an atomic monoid with $\mathcal{L}(H_i) \subset \mathcal{L}(C_n)$. Then $\mathcal{L}(H \times H_1 \times \cdots \times H_k) = \mathcal{L}(C_n)$.*

Proof. Since $\mathcal{L}(H \times H') = \{L + L' : L \in \mathcal{L}(H), L' \in \mathcal{L}(H')\}$, (1) follows.

For (2), we set $H' = H_1 \times \cdots \times H_k$. Since $\mathcal{L}(C_n)$ is additively closed by Proposition 6.14, it follows that

$$\mathcal{L}(H') = \{L_1 + \cdots + L_k : L_i \in \mathcal{L}(H_i), i \in [1, k]\} \subset \mathcal{L}(C_n).$$

Finally (1) implies that $\mathcal{L}(H \times H') = \mathcal{L}(H)$. □

We conclude this manuscript by suggesting a rich program for further study. Any progress in these directions will lead to a better understanding of direct-sum decompositions of classes of modules where each module has a semilocal endomorphism ring. Moreover, this program could stimulate new studies in combinatorial factorization theory where much of the focus has been on Krull monoids having finite class group.

Program for further study

A. Module-theoretic aspect. Let R be a ring and let $\mathcal{C}$ be a class of right R-modules which is closed under finite direct sums, direct summands, and isomorphisms, and such that the endomorphism ring $\mathrm{End}_R(M)$ is semilocal for each module M in $\mathcal{C}$ (such classes of modules are presented in a systematic way in [Facchini 2004]). Then $\mathcal{V}(\mathcal{C})$, the monoid of isomorphism classes of modules in $\mathcal{C}$ is a reduced Krull monoid with class group G and set $G_{\mathcal{P}} \subset G$ of classes containing prime divisors.

Since the long-term goal — to determine the characteristic of $\mathcal{V}(\mathcal{C})$ — is out of reach in most cases, the focus of study should be on those properties of $G_{\mathcal{P}}$ which have most crucial influence on the arithmetic of direct-sum decompositions. In particular,

- Is $G_{\mathcal{P}}$ finite or infinite?
- Is $G_{\mathcal{P}}$ well-structured in the sense of Proposition 6.2?

B. Arithmetical aspect of direct-sum decompositions. Let H be a Krull monoid with finitely generated class group G and let $G_{\mathcal{P}} \subset G$ denote the set of classes containing prime divisors.

1. *Finiteness conditions.*

   (a) Characterize the finiteness of arithmetical invariants (introduced in Section 2) and the validity of structural finiteness results (as given in Proposition 6.2, items (2a) and (2b)).

   For infinite cyclic groups much work in this direction can be found done in [Geroldinger et al. 2010].

   (b) If $G_{\mathcal{P}}$ contains an infinite group, then every finite subset $L \subset \mathbb{N}_{\geq 2}$ occurs as a set of lengths in H (see Proposition 6.2) and hence $\Delta(H) = \mathbb{N}$, and $\mathcal{U}_k(H) = \mathbb{N}_{\geq 2}$ for all $k \geq 2$. Weaken the assumption on $G_{\mathcal{P}}$ to obtain similar results.

   A weak condition on $G_{\mathcal{P}}$ enforcing that $\Delta(H) = \mathbb{N}$ can be found in [Hassler 2002].

2. *Upper bounds and precise formulas.* Suppose that G is torsion-free, say $G_{\mathcal{P}} \subset G = \mathbb{Z}^{(q)} \subset (\mathbb{R}^{(q)}, |\cdot|)$, where $|\cdot| : \mathbb{R}^{(q)} \to \mathbb{R}_{\geq 0}$ is a Euclidean norm.

   (a) If $G_{\mathcal{P}} \subset \{x \in \mathbb{R} : |x| \leq M\}$ for some $M \in \mathbb{N}$, then derive upper bounds for the arithmetical invariants in terms of M.

   (b) If $G_{\mathcal{P}}$ has a simple geometric structure (e.g., the set of vertices in a cube; see Examples 4.21 and 4.22), derive precise formulas for the arithmetical invariants, starting with the Davenport constant.

   A first result in this direction can be found in [Baeth et al. 2014].

(c) Determine the extent to which the arithmetic of a Krull monoid with $G_{\mathcal{P}}$ as in (b) is characteristic for $G_{\mathcal{P}}$. In particular, determine how this compares with the arithmetic of a Krull monoid H' where $G'_{\mathcal{P}}$ has the same geometric structure as $G_{\mathcal{P}}$ with different parameters and how this compares with the arithmetic of a Krull monoid having finite class group and prime divisors in all classes.

References

[Arnavut et al. 2007] M. Arnavut, M. Luckas, and S. Wiegand, "Indecomposable modules over one-dimensional Noetherian rings", *J. Pure Appl. Algebra* **208**:2 (2007), 739–760. MR 2007j:13010 Zbl 1111.13007

[Azumaya 1950] G. Azumaya, "Corrections and supplementaries to my paper concerning Krull–Remak–Schmidt's theorem", *Nagoya Math. J.* **1** (1950), 117–124. MR 12,314e Zbl 0040.01201

[Baeth 2007] N. R. Baeth, "A Krull–Schmidt theorem for one-dimensional rings of finite Cohen–Macaulay type", *J. Pure Appl. Algebra* **208**:3 (2007), 923–940. MR 2007h:13017 Zbl 1106.13010

[Baeth 2009] N. R. Baeth, "Direct sum decompositions over two-dimensional local domains", *Comm. Algebra* **37**:5 (2009), 1469–1480. MR 2010d:13011 Zbl 1181.13007

[Baeth and Luckas 2009] N. R. Baeth and M. R. Luckas, "Bounds for indecomposable torsion-free modules", *J. Pure Appl. Algebra* **213**:7 (2009), 1254–1263. MR 2010h:13015 Zbl 1168.13006

[Baeth and Luckas 2011] N. R. Baeth and M. R. Luckas, "Monoids of torsion-free modules over rings with finite representation type", *J. Commut. Algebra* **3**:4 (2011), 439–458. MR 2012m:13014 Zbl 1244.13009

[Baeth and Saccon 2012] N. R. Baeth and S. Saccon, "Monoids of modules over rings of infinite Cohen–Macaulay type", *J. Commut. Algebra* **4**:3 (2012), 297–326. MR 3024258 Zbl 1260.13018

[Baeth and Wiegand 2013] N. R. Baeth and R. Wiegand, "Factorization theory and decompositions of modules", *Amer. Math. Monthly* **120**:1 (2013), 3–34. MR 3007364 Zbl 1271.13023

[Baeth et al. 2014] N. R. Baeth, A. Geroldinger, D. J. Grynkiewicz, and D. Smertnig, "A semigroup-theoretical view of direct-sum decompositions and associated combinatorial problems", *J. Algebra Appl.* (2014). To appear.

[Baginski et al. 2013] P. Baginski, A. Geroldinger, D. J. Grynkiewicz, and A. Philipp, "Products of two atoms in Krull monoids and arithmetical characterizations of class groups", *European J. Combin.* **34**:8 (2013), 1244–1268. MR 3082196

[Bass 1962] H. Bass, "Torsion free and projective modules", *Trans. Amer. Math. Soc.* **102** (1962), 319–327. MR 25 #3960 Zbl 0103.02304

[Blanco et al. 2011] V. Blanco, P. A. García-Sánchez, and A. Geroldinger, "Semigroup-theoretical characterizations of arithmetical invariants with applications to numerical monoids and Krull monoids", *Illinois J. Math.* **55**:4 (2011), 1385–1414. MR 3082874 Zbl 06197210

[Bourbaki 1988] N. Bourbaki, *Commutative algebra: chapters 1–7*, Springer, New York, 1988. MR 90a:13001 Zbl 0666.13001

[Brieskorn 1967–1968] E. Brieskorn, "Rationale Singularitäten komplexer Flächen", *Invent. Math.* **4** (1967–1968), 336–358. MR 36 #5136 Zbl 0219.14003

[Buchweitz et al. 1987] R.-O. Buchweitz, G.-M. Greuel, and F.-O. Schreyer, "Cohen–Macaulay modules on hypersurface singularities, II", *Invent. Math.* **88**:1 (1987), 165–182. MR 88d:14005 Zbl 0617.14034

[Cimen 1998] N. Cimen, "One-dimensional rings of finite Cohen–Macaulay type", *J. Pure Appl. Algebra* **132**:3 (1998), 275–308. MR 99g:13007 Zbl 0971.13018

[Cimen et al. 1995] N. Cimen, R. Wiegand, and S. Wiegand, "One-dimensional rings of finite representation type", pp. 95–121 in *Abelian groups and modules* (Padova, 1994), edited by A. Facchini and C. Menini, Math. Appl. **343**, Kluwer, Dordrecht, 1995. MR 97a:13014 Zbl 0856.13006

[Coykendall and Smith 2011] J. Coykendall and W. W. Smith, "On unique factorization domains", *J. Algebra* **332** (2011), 62–70. MR 2012c:13046 Zbl 1235.13014

[Diracca 2007] L. Diracca, "On a generalization of the exchange property to modules with semilocal endomorphism rings", *J. Algebra* **313**:2 (2007), 972–987. MR 2008m:16013 Zbl 1124.16003

[Evans 1973] E. G. Evans, "Krull–Schmidt and cancellation over local rings", *Pacific J. Math.* **46** (1973), 115–121. MR 48 #2170 Zbl 0272.13006

[Evans and Griffith 1985] E. G. Evans and P. Griffith, *Syzygies*, London Mathematical Society Lecture Note Series **106**, Cambridge University Press, 1985. MR 87b:13001 Zbl 0569.13005

[Facchini 2002] A. Facchini, "Direct sum decompositions of modules, semilocal endomorphism rings, and Krull monoids", *J. Algebra* **256**:1 (2002), 280–307. MR 2003k:16009 Zbl 1016.16002

[Facchini 2003] A. Facchini, "The Krull–Schmidt theorem", pp. 357–397 in *Handbook of algebra*, vol. 3, edited by M. Hazewinkel, North-Holland, Amsterdam, 2003. MR 2005c:16009 Zbl 1079.16002

[Facchini 2004] A. Facchini, "Geometric regularity of direct-sum decompositions in some classes of modules", *Fundam. Prikl. Mat.* **10**:3 (2004), 231–244. In Russian; translated in *J. Math. Sci.* (*New York*) **139**:4 (2006), 6814–6822. MR 2005m:16007 Zbl 1073.16004

[Facchini 2006] A. Facchini, "Krull monoids and their application in module theory", pp. 53–71 in *Algebras, rings and their representations* (Lisbon, 2003), edited by A. Facchini et al., World Scientific, Hackensack, NJ, 2006. MR 2007b:16010 Zbl 1113.20049

[Facchini 2012] A. Facchini, "Direct-sum decompositions of modules with semilocal endomorphism rings", *Bull. Math. Sci.* **2**:2 (2012), 225–279. MR 2994204 Zbl 1277.16001

[Facchini and Halter-Koch 2003] A. Facchini and F. Halter-Koch, "Projective modules and divisor homomorphisms", *J. Algebra Appl.* **2**:4 (2003), 435–449. MR 2004j:16007 Zbl 1058.16002

[Facchini and Wiegand 2004] A. Facchini and R. Wiegand, "Direct-sum decompositions of modules with semilocal endomorphism rings", *J. Algebra* **274**:2 (2004), 689–707. MR 2005d:20106 Zbl 1094.20036

[Facchini et al. 2006] A. Facchini, W. Hassler, L. Klingler, and R. Wiegand, "Direct-sum decompositions over one-dimensional Cohen–Macaulay local rings", pp. 153–168 in *Multiplicative ideal theory in commutative algebra*, edited by J. W. Brewer et al., Springer, New York, 2006. MR 2007k:13045 Zbl 1121.13024

[Frankild et al. 2008] A. J. Frankild, S. Sather-Wagstaff, and R. Wiegand, "Ascent of module structures, vanishing of Ext, and extended modules", *Michigan Math. J.* **57** (2008), 321–337. MR 2010k:13025 Zbl 1174.13013

[Gao and Geroldinger 2006] W. Gao and A. Geroldinger, "Zero-sum problems in finite abelian groups: a survey", *Expo. Math.* **24**:4 (2006), 337–369. MR 2008d:11014 Zbl 1122.11013

[Gao and Geroldinger 2009] W. Gao and A. Geroldinger, "On products of k atoms", *Monatsh. Math.* **156**:2 (2009), 141–157. MR 2010g:20111 Zbl 1184.20051

[Gao et al. 2015] W. Gao, A. Geroldinger, and W. A. Schmid, "Local and global tameness in Krull monoids", *Commun. Algebra* **43**:1 (2015), 262–296.

[Geroldinger 2009] A. Geroldinger, "Additive group theory and non-unique factorizations", pp. 1–86 in *Combinatorial number theory and additive group theory*, edited by A. Geroldinger and I. Z. Ruzsa, Birkhäuser, Basel, 2009. MR 2011a:20153 Zbl 1221.20045

[Geroldinger and Göbel 2003] A. Geroldinger and R. Göbel, "Half-factorial subsets in infinite abelian groups", *Houston J. Math.* **29**:4 (2003), 841–858. MR 2004m:13050 Zbl 1095.13531

[Geroldinger and Halter-Koch 2006] A. Geroldinger and F. Halter-Koch, *Non-unique factorizations: algebraic, combinatorial and analytic theory*, Pure and Applied Mathematics (Boca Raton) **278**, Chapman & Hall/CRC, Boca Raton, FL, 2006. MR 2006k:20001 Zbl 1113.11002

[Geroldinger and Hassler 2008] A. Geroldinger and W. Hassler, "Local tameness of v-Noetherian monoids", *J. Pure Appl. Algebra* **212**:6 (2008), 1509–1524. MR 2009b:20114 Zbl 1133.20047

[Geroldinger and Kainrath 2010] A. Geroldinger and F. Kainrath, "On the arithmetic of tame monoids with applications to Krull monoids and Mori domains", *J. Pure Appl. Algebra* **214**:12 (2010), 2199–2218. MR 2011i:20085 Zbl 1207.20055

[Geroldinger and Ruzsa 2009] A. Geroldinger and I. Z. Ruzsa, *Combinatorial number theory and additive group theory*, Birkhäuser, Basel, 2009. MR 2010f:11005 Zbl 1177.11005

[Geroldinger and Yuan 2013] A. Geroldinger and P. Yuan, "The monotone catenary degree of Krull monoids", *Results Math.* **63**:3-4 (2013), 999–1031. MR 3057352 Zbl 06186656

[Geroldinger et al. 2010] A. Geroldinger, D. J. Grynkiewicz, G. J. Schaeffer, and W. A. Schmid, "On the arithmetic of Krull monoids with infinite cyclic class group", *J. Pure Appl. Algebra* **214**:12 (2010), 2219–2250. MR 2011h:20120 Zbl 1208.13003

[Green and Reiner 1978] E. L. Green and I. Reiner, "Integral representations and diagrams", *Michigan Math. J.* **25**:1 (1978), 53–84. MR 80g:16032 Zbl 0365.16015

[Grynkiewicz 2013] D. J. Grynkiewicz, *Structural additive theory*, Developments in Mathematics **30**, Springer, Cham, 2013. MR 3097619 Zbl 06162097

[Halter-Koch 1998] F. Halter-Koch, *Ideal systems: an introduction to multiplicative ideal theory*, Monographs and Textbooks in Pure and Applied Mathematics **211**, Marcel Dekker, New York, 1998. MR 2001m:13005 Zbl 0953.13001

[Hassler 2002] W. Hassler, "Factorization properties of Krull monoids with infinite class group", *Colloq. Math.* **92**:2 (2002), 229–242. MR 2003c:20072 Zbl 0997.20056

[Hassler and Wiegand 2009] W. Hassler and R. Wiegand, "Extended modules", *J. Commut. Algebra* **1**:3 (2009), 481–506. MR 2011f:13013 Zbl 1189.13007

[Hassler et al. 2007] W. Hassler, R. Karr, L. Klingler, and R. Wiegand, "Big indecomposable modules and direct-sum relations", *Illinois J. Math.* **51**:1 (2007), 99–122. MR 2008g:13013 Zbl 1129.13010

[Heitmann 1993] R. C. Heitmann, "Characterization of completions of unique factorization domains", *Trans. Amer. Math. Soc.* **337**:1 (1993), 379–387. MR 93g:13006 Zbl 0792.13011

[Herbera and Příhoda 2010] D. Herbera and P. Příhoda, "Big projective modules over Noetherian semilocal rings", *J. Reine Angew. Math.* **648** (2010), 111–148. MR 2012b:16008 Zbl 1215.16002

[Kainrath 1999] F. Kainrath, "Factorization in Krull monoids with infinite class group", *Colloq. Math.* **80**:1 (1999), 23–30. MR 2000c:20088 Zbl 0936.20050

[Karr and Wiegand 2011] R. Karr and R. Wiegand, "Direct-sum behavior of modules over one-dimensional rings", pp. 251–275 in *Commutative algebra: Noetherian and non-Noetherian perspectives*, edited by M. Fontana et al., Springer, New York, 2011. MR 2012c:13021 Zbl 1235.13008

[Knörrer 1987] H. Knörrer, "Cohen–Macaulay modules on hypersurface singularities, I", *Invent. Math.* **88**:1 (1987), 153–164. MR 88d:14004 Zbl 0617.14033

[Krull 1925] W. Krull, "Über verallgemeinerte endliche abelsche Gruppen", *Math. Z.* **23**:1 (1925), 161–196. MR 1544736 JFM 51.0116.03

[Lech 1986] C. Lech, "A method for constructing bad Noetherian local rings", pp. 241–247 in *Algebra, algebraic topology and their interactions* (Stockholm, 1983), edited by J.-E. Roos, Lecture Notes in Math. **1183**, Springer, Berlin, 1986. MR 87m:13010a Zbl 0589.13006

[Leedham-Green 1972] C. R. Leedham-Green, "The class group of Dedekind domains", *Trans. Amer. Math. Soc.* **163** (1972), 493–500. MR 45 #1888 Zbl 0231.13008

[Leuschke and Wiegand 2012] G. J. Leuschke and R. Wiegand, *Cohen–Macaulay representations*, Mathematical Surveys and Monographs **181**, American Mathematical Society, Providence, RI, 2012. MR 2919145 Zbl 1252.13001

[Levy and Odenthal 1996] L. S. Levy and C. J. Odenthal, "Krull–Schmidt theorems in dimension 1", *Trans. Amer. Math. Soc.* **348**:9 (1996), 3391–3455. MR 96m:16006a Zbl 0858.16016

[Maclagan-Wedderburn 1909] J. H. Maclagan-Wedderburn, "On the direct product in the theory of finite groups", *Ann. of Math.* (2) **10**:4 (1909), 173–176. MR 1502387 JFM 40.0192.02

[Plagne and Schmid 2013] A. Plagne and W. A. Schmid, "On congruence half-factorial Krull monoids with cyclic class group", 2013. Submitted.

[Remak 1911] R. Remak, "Über die Zerlegung der endlichen Gruppen in direkte unzerlegbare Faktoren", *J. Reine Angew. Math.* **139** (1911), 293–308. JFM 42.0156.01

[Rotthaus et al. 1999] C. Rotthaus, D. Weston, and R. Wiegand, "Indecomposable Gorenstein modules of odd rank", *J. Algebra* **214**:1 (1999), 122–127. MR 2000c:13017 Zbl 0936.13007

[Saccon 2010] S. Saccon, *One-dimensional local rings of infinite Cohen–Macaulay type*, thesis, University of Nebraska, Lincoln, NE, 2010, Available at http://search.proquest.com/docview/748220858. MR 2941519

[Schmid 2009a] W. A. Schmid, "Arithmetical characterization of class groups of the form $\mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}$ via the system of sets of lengths", *Abh. Math. Semin. Univ. Hambg.* **79**:1 (2009), 25–35. MR 2010h:20140 Zbl 1191.20069

[Schmid 2009b] W. A. Schmid, "Characterization of class groups of Krull monoids via their systems of sets of lengths: a status report", pp. 189–212 in *Number theory and applications* (Allahabad, 2006–2007), edited by S. D. Adhikari and B. Ramakrishnan, Hindustan Book Agency, New Delhi, 2009. MR 2010k:20101 Zbl 1244.20053

[Schmid 2009c] W. A. Schmid, "A realization theorem for sets of lengths", *J. Number Theory* **129**:5 (2009), 990–999. MR 2010h:11182 Zbl 1191.11031

[Schmidt 1929] O. Schmidt, "Über unendliche Gruppen mit endlicher Kette", *Math. Z.* **29**:1 (1929), 34–41. MR 1544991 JFM 54.0148.03

[Wiegand 2001] R. Wiegand, "Direct-sum decompositions over local rings", *J. Algebra* **240**:1 (2001), 83–97. MR 2002b:13017 Zbl 0986.13010

[Wiegand and Wiegand 1994] R. Wiegand and S. Wiegand, "Bounds for one-dimensional rings of finite Cohen–Macaulay type", *J. Pure Appl. Algebra* **93**:3 (1994), 311–342. MR 95c:13004 Zbl 0813.13013

Received June 27, 2013.

NICHOLAS R. BAETH
MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF CENTRAL MISSOURI
WARRENSBURG, MO 64093
UNITED STATES
baeth@ucmo.edu

ALFRED GEROLDINGER
INSTITUT FÜR MATHEMATIK UND WISSENSCHAFTLICHES RECHNEN
KARL-FRANZENS-UNIVERSITÄT, NAWI GRAZ
8010 GRAZ
AUSTRIA
alfred.geroldinger@uni-graz.at

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.321

ON THE TORSION ANOMALOUS CONJECTURE IN CM ABELIAN VARIETIES

SARA CHECCOLI AND EVELINA VIADA

The torsion anomalous conjecture (TAC) states that a subvariety V of an abelian variety A has only finitely many maximal torsion anomalous subvarieties. In this work we prove, with an effective method, some cases of the TAC when the ambient variety A has CM, generalising our previous results in products of CM elliptic curves. When V is a curve, we give new results and we deduce some implications on the effective Mordell–Lang conjecture.

1. Introduction

Let A be an abelian variety embedded in the projective space and let V be a proper subvariety of A. Assume that both A and V are defined over the algebraic numbers.

Definition 1.1. The variety V is a *translate* (resp. a *torsion variety*) if it is a finite union of translates of proper algebraic subgroups by points (resp. by torsion points).

V is *transverse* (resp. *weak-transverse*) in A if V is irreducible and V is not contained in any translate (resp. in any torsion subvariety) of A.

It is a classical problem in diophantine geometry to investigate the relationship between the above geometrical definitions and the arithmetical properties of the variety V. In this direction, there are several celebrated theorems, such as the Manin–Mumford, Mordell–Lang and Bogomolov conjectures.

Recently E. Bombieri, D. Masser and U. Zannier [Bombieri et al. 2007] suggested a new approach to this kind of investigation, introducing in particular the notion of torsion anomalous intersections.

Definition 1.2. An irreducible subvariety Y of V is *V-torsion anomalous* if:

- Y is an irreducible component of $V \cap (B+\zeta)$, with $B+\zeta$ an irreducible torsion variety.
- The dimension of Y is larger than expected; i.e.,

$$\operatorname{codim} Y < \operatorname{codim} V + \operatorname{codim} B.$$

MSC2010: primary 11G50; secondary 14G40.
Keywords: diophantine approximation, heights, abelian varieties, intersections with torsion varieties.

The variety $B+\zeta$ is *minimal* for Y if, in addition, it has minimal dimension. The *relative codimension* of Y is the codimension of Y in such a minimal $B+\zeta$.

We say that Y is *maximal* if it is not contained in any V-torsion anomalous variety of strictly larger dimension.

In [Bombieri et al. 2007], the authors formulate several conjectures. We state here one natural variant.

Conjecture 1.3 (TAC, torsion anomalous conjecture). *For any algebraic subvariety V of a* (*semi*)*abelian variety*, *there are only finitely many maximal V-torsion anomalous varieties.*

The TAC is well known to have several important consequences. It implies, for instance, the Manin–Mumford and the Mordell–Lang conjectures; it is also related to model theory by the work of B. Zilber and to algebraic dynamics by the recent work of J. H. Silverman and P. Morton. In addition, R. Pink generalised it to mixed Shimura varieties.

Only a few cases of the TAC are known: Viada [2008] proved it for curves in a product of elliptic curves, Maurin [2008] for curves in a torus, Bombieri et al. [2007] for varieties of codimension 2 in a torus. Habegger [2008] gave related results under some stronger assumptions on V.

In [Checcoli et al. 2014], we prove an effective TAC for maximal V-torsion anomalous varieties of relative codimension 1 in a product of CM elliptic curves. Our bounds are explicit and uniform in their dependence on V. As an immediate corollary, we prove the TAC for varieties of codimension 2, obtaining an elliptic analogue of the toric result in [Bombieri et al. 2007]. In the present work, we generalise our results to CM abelian varieties. In [Checcoli et al. 2014], we also point out interesting relations between this kind of theorem and other relevant conjectures, such as the Zilber–Pink conjecture and the above-mentioned ones.

Let $A \subseteq \mathbb{P}^m$ be an abelian variety with CM defined over a number field k and let k_{tor} be the field of definition of the torsion points of A. Let A be isogenous to a product of simple abelian varieties of dimension at most g. For a point $x \in A$, we denote by $\hat{h}(x)$ its canonical Néron–Tate height. For a subvariety $V \subseteq A$, we denote by $h(V)$ its normalised height and by $k_{\mathrm{tor}}(V)$ its field of definition over k_{tor} (see Section 2). By $\ll$ we denote an inequality up to a multiplicative constant depending on A. Our main result is the following:

Theorem 1.4. *Let $V \subseteq A$ be a weak-transverse subvariety of codimension $> g$. Then there are only finitely many maximal V-torsion anomalous subvarieties Y of relative codimension* 1.

Effective version: *More precisely*, *if $B+\zeta$ is minimal for Y*, *then for any positive real η*, *there exist constants depending only on A and η such that*:

(1) *If Y is not a translate*, *then*

$$\deg B \ll_\eta (h(V)+\deg V)^{\operatorname{codim} B+\eta},$$
$$h(Y) \ll_\eta (h(V)+\deg V)^{\operatorname{codim} B+\eta},$$
$$\deg Y \ll_\eta \deg V(h(V)+\deg V)^{\operatorname{codim} B-1+\eta}.$$

(2) *If Y is a point*, *then*

$$\deg B \ll_\eta \big((h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]\big)^{\operatorname{codim} B+\eta},$$
$$\hat{h}(Y) \ll_\eta (h(V)+\deg V)^{\operatorname{codim} B+\eta}[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]^{\operatorname{codim} B-1+\eta},$$
$$[k_{\mathrm{tor}}(Y):k_{\mathrm{tor}}] \ll_\eta \deg V[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]^{\operatorname{codim} B+\eta}(h(V)+\deg V)^{\operatorname{codim} B-1+\eta}.$$

(3) *If Y is a translate of positive dimension*, *then*

$$\deg B \ll_\eta \big((h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]\big)^{\operatorname{codim} B+\eta},$$
$$h(Y) \ll_\eta (h(V)+\deg V)^{\operatorname{codim} B+\eta}[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]^{\operatorname{codim} B-1+\eta},$$
$$\deg Y \ll_\eta \deg V\big((h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]\big)^{\operatorname{codim} B-1+\eta}.$$

In addition, *the torsion points ζ belong to a finite set of cardinality effectively bounded in terms of* $\deg V$, $\deg B$ *and constants depending only on A*.

This theorem can be reformulated in the context of several other well-known conjectures, as explained in the introduction of [Checcoli et al. 2014].

The proof of Theorem 1.4 is split into two sections, depending on whether Y is a translate or not: in Section 4 we prove part (1) and in Section 5 we prove parts (2) and (3).

The main ingredients (see Section 2.3) needed for the proof of Theorem 1.4 are Zhang's inequality, the arithmetic Bézout theorem by P. Philippon, our sharp Bogomolov-type bound proved in [Checcoli et al. 2012], and the relative Lehmer estimate by M. Carrizosa. As usual, the CM hypothesis is due to the use of a Lehmer bound, known only for CM varieties. This result is only needed when Y is a translate, while case (1) of Theorem 1.4 holds with the weaker assumption that A has a positive density of ordinary primes, as required to apply a Bogomolov-type bound (see [Galateau 2010, p. 779]). In particular, our method could treat the case of general abelian varieties, if the Lehmer- and Bogomolov-type bounds were known in such generality.

In Theorem 1.5, proved in Section 6, we expand our method in order to get some new effective results for curves in abelian varieties. This is particularly relevant, as bounds for the height in weak-transverse curves are hard to obtain. For instance, such bounds allow us to deduce some cases of the effective Mordell–Lang conjecture, stated in Corollary 1.6. The two classical approaches to the effective

Mordell–Lang conjecture in abelian varieties are the Chabauty–Coleman and the Manin–Demjanenko methods. These methods require hypotheses that are similar to ours, but our result is of easier application and more explicit. Finally in Section 7, we give some generalisations to varieties in abelian varieties.

In particular we prove the following results. We fix an isogeny of the CM abelian variety A to the product $\prod_{i=1}^{\ell} A_i^{e_i}$ of nonisogenous simple factors A_i of dimension g_i. Since isogenies preserve finiteness, without loss of generality, we identify A with $\prod_{i=1}^{\ell} A_i^{e_i}$. If $H \subsetneq A$ is a subgroup, then $H = H_1 \times \cdots \times H_\ell$, where $H_i \subseteq A_i^{e_i}$ is isogenous to $A_i^{f_i}$ for some $f_i \le e_i$; therefore the matrix of the coefficients of the forms defining H has the structure of a block diagonal matrix with entries in the endomorphism ring of the corresponding varieties. We can now state our effective result for weak-transverse curves, which is an example for the effective Zilber–Pink conjecture.

Theorem 1.5. *Let $C \subseteq A = \prod_{i=1}^{\ell} A_i^{e_i}$ be a weak-transverse curve. Then the set*

$$\mathcal{S}(C) = C \cap \left(\bigcup_{H \in \mathcal{F}} H \right)$$

is a set of bounded Néron–Tate height, where $\mathcal{F}$ is the family of all subgroups $H = \prod_{i=1}^{\ell} H_i \subseteq A$ such that

$$\operatorname{codim} H_j > g_j \dim H_j$$

for at least one index j (here $\operatorname{codim} H_j$ *is the codimension of H_j in $A_j^{e_j}$).*

More precisely, if $Y \in C \cap H$, then for any real $\eta > 0$, there exists a constant, depending only on A and η, such that

$$\hat{h}(Y) \ll_\eta (h(C) + \deg C)^{\frac{\operatorname{codim} H_j}{\operatorname{codim} H_j - g_j \dim H_j} + \eta} [k_{\mathrm{tor}}(C) : k_{\mathrm{tor}}]^{\frac{g_j \dim H_j}{\operatorname{codim} H_j - g_j \dim H_j} + \eta}.$$

To prove Theorem 1.5 we first work in the projection on the j-th factor of A, and then we lift the construction to the variety A.

As an immediate consequence, we deduce the following corollary (proved in Section 6.1).

Let Γ be a subgroup of $A = \prod_{i=1}^{\ell} A_i^{e_i}$. Assume that the group $\overline{\Gamma}_i < A_i$ generated by the coordinates of the projections of Γ on the factors $A_i^{e_i}$ is an $\mathrm{End}(A_i)$-module of rank t_i.

Corollary 1.6. *Let A be a CM abelian variety and let C be a weak-transverse curve in A. Let Γ be a subgroup as above, and suppose that $t_j < e_j/(g_j + 1)$ for some index j. Then, for any positive η, there exists a constant depending only on A*

and η, such that the set $C \cap \Gamma$ has Néron–Tate height bounded as

$$\hat{h}(C \cap \Gamma) \ll_\eta (h(C) + \deg C)^{\frac{e_j - t_j}{e_j - (g_j+1)t_j} + \eta} [k_{\mathrm{tor}}(C) : k_{\mathrm{tor}}]^{\frac{g_j t_j}{e_j - (g_j+1)t_j} + \eta}.$$

We remark that the corollary applies also to some Γ of infinite rank; indeed, we only assume that the rank on one projection is small (see Remark 6.1).

2. Preliminaries

2.1. ***Height and subgroups.*** We assume that all varieties are defined over the field of algebraic numbers.

Let A be an abelian variety with CM. We fix, up to an isogeny, a decomposition of $A = \prod_{i=1}^{\ell} A_i^{e_i}$ in simple factors of dimension $\dim A_i = g_i$. We consider an embedding $i_{\mathscr{L}}$ of A in $\mathbb{P}^m$ given by a symmetric ample line bundle $\mathscr{L}$ on A. Heights and degrees corresponding to $\mathscr{L}$ are computed via $i_{\mathscr{L}}$. More precisely, the degree of a subvariety of A is the degree of its image under $i_{\mathscr{L}}$; $\hat{h} = \hat{h}_{\mathscr{L}}$ is the $\mathscr{L}$-canonical Néron–Tate height of a point in A, and h is the normalised height of a subvariety of A as defined, for instance, in [Philippon 1991]. Notice that if $x \in A$ is a point, then $\hat{h}(x) = h(x)$.

By Lemma 2.2 in [Masser and Wüstholz 1993], if A is an abelian variety defined over a number field k, then every abelian subvariety of A is defined over a finite extension of k of degree bounded by $3^{16(\dim A)^4}$; thus, without loss of generality, we assume that all abelian subvarieties of A are defined over k.

Let $B + \zeta$ be an irreducible torsion variety of A. Then $B = B_1 \times \cdots \times B_l$, where $B_i \subseteq A_i^{e_i}$ is isogenous to $A_i^{f_i}$ for some integer $0 \le f_i \le e_i$.

There exists a natural correspondence between abelian subvarieties B of A, morphisms from A to $\prod_{i=1}^{\ell} A_i^{e_i - f_i}$, and matrices made of ℓ blocks where the i-th block is an $(e_i - f_i) \times e_i$-matrix with entries in the endomorphism ring of A_i. For details on such a correspondence see, for instance, [Checcoli et al. 2012, Section 2.5]. In short, the abelian subvariety B defines the projection morphism $\pi_B : A \to A/B$. The successive minima of B give a matrix $\mathscr{H}_B$ of the above type. By multiplication on the left, the matrix $\mathscr{H}_B$ gives a morphism Φ_B from A to $\prod_{i=1}^{\ell} A_i^{e_i - f_i}$, where B is the zero component of $\ker \Phi_B$.

By Minkowski's theorem, $\deg B$ is (up to constants depending only on A) the product of the squares of the norms of the rows of $\mathscr{H}_B$. In addition, B is the zero component of the zero set of the forms $h_1, \ldots, h_r$ corresponding to the rows of $\mathscr{H}_B$. We order the h_i by increasing degrees d_i so that

$$d_1 \cdots d_r \ll \deg(B + \zeta) \ll d_1 \cdots d_r.$$

We also recall that, from [Masser and Wüstholz 1993, Lemmas 1.3 and 1.4], if B is an abelian subvariety of A and $B^\perp$ is its orthogonal complement, then $\deg B^\perp \ll \deg B$, and therefore $\#(B \cap B^\perp) \ll (\deg B)^2$.

2.2. ***Torsion anomalous varieties.*** We recall some preliminary lemmas on torsion anomalous varieties used for our geometric constructions in the following sections.

Lemma 2.1 [Checcoli et al. 2014, Lemma 3.5]. *Let Y be a maximal V-torsion anomalous variety and let $B+\zeta$ be minimal for Y. Then Y is weak-transverse in $B+\zeta$ (i.e., Y is not contained in any proper torsion subvariety of $B+\zeta$).*

Lemma 2.2 [Checcoli et al. 2014, Lemma 3.6]. *Let Y be a maximal V-torsion anomalous variety, and let $B+\zeta$ be minimal for Y. Then Y is a component of $V \cap (B'+\zeta)$ for every algebraic subgroup $B' \supseteq B$ with* $\operatorname{codim} B' \geq \dim V - \dim Y$.

The following lemma is due to Philippon [2012] and to certain properties of orthogonality in the Mordell–Weil groups studied by D. Bertrand [1986].

We recall that the *essential minimum* of a subvariety $X \subseteq A$ is defined as

$$\mu(X) = \sup\bigl\{\theta \in \mathbb{R} \bigm| \{x \in X(\overline{\mathbb{Q}}) \mid \hat{h}(x) \leq \theta\} \text{ is nondense in } X\bigr\}.$$

Lemma 2.3. *Let $H+Y_0$ be a weak-transverse translate in A, with Y_0 a point in the orthogonal complement $H^\perp$ of H. Then $\mu(Y_0) = \mu(H+Y_0)$.*

We conclude with a remark on translations by torsion points.

Remark 2.4. Notice that, for any subvariety X of A, translations by a torsion point ζ leave invariant the degree, the field of definition over k_{tor} and the normalised height of X (see also [Philippon 1991, Proposition 9]). In addition, if $Y \subseteq V \cap (B+\zeta)$ is V-torsion anomalous, then $Y-\zeta \subseteq (V-\zeta) \cap B$ is $(V-\zeta)$-torsion anomalous. Therefore, without loss of generality, we can work in V or in $V-\zeta$ with the advantage, in the latter case, that B is an abelian subvariety.

2.3. ***Main ingredients.*** We recall here the main ingredients used in the proof of Theorem 1.4.

2.3.1. *The Zhang estimate.* The theorem below follows from the crucial result in Zhang's proof [1998] of the Bogomolov conjecture and from the definition of normalised height.

Theorem 2.5. *Let $X \subseteq A$ be an irreducible subvariety.*

Then

$$\mu(X) \leq \frac{h(X)}{\deg X} \leq (1+\dim X)\mu(X).$$

2.3.2. *The arithmetic Bézout theorem.* The following version of the arithmetic Bézout theorem is due to Philippon [1995].

Theorem 2.6. *Let X and Y be irreducible subvarieties of the projective space $\mathbb{P}^n$ defined over $\overline{\mathbb{Q}}$; let $Z_1, \dots, Z_g$ be the irreducible components of $X \cap Y$. Then*

$$\sum_{i=1}^{g} h(Z_i) \le \deg X\, h(Y) + \deg Y\, h(X) + c(n) \deg X \deg Y,$$

where $c(n)$ is a constant depending only on n.

2.3.3. *An effective Bogomolov estimate.* The following sharp Bogomolov bound is proved by [Checcoli et al. 2012] and generalises a result of A. Galateau [2010].

Theorem 2.7 (Checcoli, Veneziano, Viada). *Let A be an abelian variety with a positive density of ordinary primes, and let Y be an irreducible subvariety of A transverse in a translate $B + p$. Then, for any $\eta > 0$, there exists a positive constant c_1 depending on A and η, such that*

$$\mu(Y) \ge c_1 \frac{(\deg B)^{1/(\dim B - \dim Y) - \eta}}{(\deg Y)^{1/(\dim B - \dim Y) + \eta}}.$$

2.3.4. *A relative Lehmer estimate.* The following Lehmer bound is proved in [Carrizosa 2009].

Theorem 2.8. *Let A be an abelian variety with CM defined over a number field k, and let k_{tor} be the field of definition of all torsion points of A. Let P be a point of infinite order in A, and let $B + \zeta$ be the torsion variety of minimal dimension containing P, with B an abelian subvariety and ζ a torsion point. Then for every $\eta > 0$, there exists a positive constant c_2 depending on A and η, such that*

$$\hat{h}(P) \ge c_2 \frac{(\deg B)^{1/\dim B - \eta}}{[k_{\mathrm{tor}}(P) : k_{\mathrm{tor}}]^{1/\dim B + \eta}}.$$

3. Finitely many maximal V-torsion anomalous varieties in $V \cap (B + \mathrm{Tor}_A)$

Let V be a weak-transverse variety in an abelian variety A. Let us fix an abelian subvariety B of A. In [Checcoli et al. 2014, Lemma 3.9] we proved that there are only finitely many $\zeta \in \mathrm{Tor}_{B^\perp}$ such that $V \cap (B + \zeta)$ has a maximal V-torsion anomalous component. In this section we prove that the number of such ζ is effectively bounded in terms of $\deg V$, $\deg B$ and some constants depending on A (Proposition 3.5). We thank the referee for pointing out the effectivity question and for his useful comments.

The proof of such an effective result is based on an induction on the dimension of V, on Rémond's quantitative version of the Manin–Mumford conjecture [2000] and on the effective bound for the degree of the maximal translates in a variety implied, for instance, by a result of Bombieri and Zannier [1996]. We first recall these results and some other well-known bounds.

Recall that A is an abelian variety and $\mathscr{L}$ is a symmetric ample line bundle on A. We denote by $h_1(A)$ the projective height of the zero of A in the embedding associated with $\mathscr{L}^{\otimes 16}$ (as defined in [David and Philippon 2002, Notation 3.2]) and by d_A the degree of the field of definition of A. If G is an abelian subvariety of A or a quotient of A, then $h_1(G)$ is bounded in terms of $h_1(A)$, $\deg A$, $\dim A$ and $\deg G$ (see [ibid., Proposition 3.9]).

Moreover, in several works, Masser and Wüstholz and then other authors proved that for any abelian subvariety G of A, the degree of the field of definition of G is at most $3^{16(\dim A)^4} d_A$ (see [Masser and Wüstholz 1993, Lemma 2.2]). Below, we sum up these bounds.

Estimate 3.1. *If G is an abelian subvariety of A or a quotient of A, then*

- d_G *is bounded in terms of* d_A *and* $\dim A$;
- $h_1(G)$ *is bounded in terms of* $h_1(A)$, $\deg A$, $\dim A$ *and* $\deg G$.

For simplicity, in what follows we shall denote by $c(A)$ any constant depending on $\dim A, d_A, h_1(A)$ and $\deg A$.

We recall the following consequence of Rémond's result [2000, Theorem 1.2].

Estimate 3.2. *The number of irreducible components of the closure of the torsion of a weak-transverse variety V in an abelian variety A is effectively bounded as*

$$c(A)(\deg V)^{(\dim A)^{5(\dim V+1)^2}}.$$

Following the work of Rémond [2000, Theorem 2.1] and the results in [David and Philippon 2002] one sees that if G is an abelian subvariety of A or a quotient of A, then the corresponding constant $c(G)$ appearing in Estimate 3.2 is bounded only in terms of $\dim A, d_A, h_1(A)$, $\deg A$ and $\deg G$.

In our previous joint work with F. Veneziano [Checcoli et al. 2014, Lemma 7.4], we gave an explicit version of a corollary of Lemma 2 in [Bombieri and Zannier 1996]. This is a bound for the degree of the maximal translates contained in a variety, and so in particular for the degree of each component of the closure of the torsion. More precisely:

Estimate 3.3. *If V is weak-transverse in an abelian variety A, then the maximal translates contained in V have degree bounded by* $c(A)(\deg V)^{2^{\dim V}}$.

Notice that if ζ is a torsion point such that $V \cap (B+\zeta)$ has a V-torsion anomalous component, then all the points in $\zeta + (B \cap B^\perp)$ share the same property. Indeed $B+\zeta = B+\zeta+(B\cap B^\perp)$. Clearly, we shall avoid such a redundancy and work up to points in $B \cap B^{\perp}$. Nevertheless, $|B \cap B^{\perp}| \ll (\deg B)^2$. This makes the following definition consistent.

Definition 3.4. Let B be an abelian subvariety of an abelian variety A. Let V be a weak-transverse subvariety of A. We denote by $Z_{V,A}$ the set of torsion points $\zeta \in B^{\perp}/B \cap B^{\perp}$ such that $V \cap (B+\zeta)$ has a maximal V-torsion anomalous component Y_ζ.

We point out that the set $Z_{V,A}$ also depends on B. However, in our proof B is fixed, while V and A vary. To simplify the notation we only indicate the dependence on V and A.

In the following proposition we estimate the number of points in $Z_{V,A}$. The number of maximal V-torsion anomalous components in $V \cap (B + \mathrm{Tor}_A)$ is clearly estimated by $|Z_{V,A}| \deg V \deg B$.

Proposition 3.5. *Let B be an abelian subvariety of an abelian variety A. Let V be weak-transverse in A. Then the cardinality of $Z_{V,A}$ is effectively bounded in terms of* $\deg V, \deg B$ *and constants depending only on* $\dim A, h_1(A), d_A$ *and* $\deg A$.

Proof. Consider the projection

$$\pi_B : A \to A/B.$$

We recall that the degree of the image via π_B of a variety $X \subseteq A$ and the degree of the preimage via π_B of a variety $X \subseteq A/B$ only depend on $\deg X$, $\deg B$ and $\deg A$. In particular, $\deg \pi_B(V)$ is bounded in terms of $\deg V$, $\deg B$, and $\deg A$ and $\deg A/B$ is bounded in terms of $\deg B$ and $\deg A$.

The proof of our proposition is done by induction on the dimension of V.

The base of our induction is the case of a curve, i.e., $\dim V = 1$. Then $\pi_B(V)$ is a weak-transverse curve in A/B because V is weak-transverse in A. Moreover the points of $Z_{V,A}$ map to torsion points of $\pi_B(V)$. The number of torsion points of $\pi_B(V)$ is estimated using Estimate 3.2. Their preimage, which contains $Z_{V,A}$, then has cardinality effectively bounded in terms of $\deg V$, $\deg B$ and $c(A)$.

Suppose by inductive hypothesis that the proposition holds for every variety V with $\dim V < n$. We then show that it holds for V of dimension n.

To prove our result, we are going to partition $Z_{V,A}$ into a finite union of subsets Z_X associated with irreducible subvarieties X of V of dimension $< n$. We then verify that such varieties X satisfy the assumption of the proposition; by the inductive hypothesis, we deduce that the cardinalities $|Z_X|$ are effectively bounded in terms of $\deg V$, $\deg B$ and $c(A)$.

Denote by $f : V \to A/B$ the restriction of π_B to V.

If f is dominant, then the generic fibre $F_p = V \cap (B + \tilde{p})$ has dimension

$$\dim F_p = \dim V - \operatorname{codim} B, \tag{1}$$

where p belongs to an open subset of A/B and $f(\tilde{p}) = p$. The dimensional equation (1) shows that the generic fibre is not anomalous. Consider the subset V_π of A/B given by all points that do not have generic fibre. By the fibre dimension theorem (see, for instance, [Shafarevich 1972, Section 6.3, Theorem 7]), this is a proper closed subset of $\pi_B(V)$ and its degree is effectively bounded in terms of $\deg V$, $\deg B$ and $c(A)$. Note that the image of $Z_{V,A}$ via π_B is a subset of the torsion of V_π; indeed the fibre of a point in $\pi_B(Z_{V,A})$ is torsion anomalous and therefore does not satisfy the equality (1).

If f is not dominant, then set $V_\pi = \pi_B(V)$. Clearly, $Z_{V,A}$ is a subset of the torsion of V_π.

Note that in both cases

(a) $\deg V_\pi$ is bounded in terms of $\deg V$, $\deg B$ and $c(A)$.

Let $T_1, \dots, T_r$ be the isolated components of the closure of the torsion of V_π intersecting $\pi_B(Z_{V,A})$. Clearly

$$Z_{V,A} = \bigcup_{i=1}^{r} (\pi_B^{-1}(T_i) \cap Z_{V,A})$$

and

$$|Z_{V,A}| = \sum_{i=1}^{r} \left|\pi_B^{-1}(T_i) \cap Z_{V,A}\right|.$$

From Estimate 3.2 and (a), the number r is effectively bounded in terms of $\deg V$, $\deg B$, $\deg A$ and $c(A)$. Thus we shall prove that, for every $1 \le i \le r$, the cardinality $|(\pi_B^{-1}(T_i) \cap Z_{V,A})|$ is effectively bounded in terms of $\deg V$, $\deg B$ and $c(A)$.

Let T be one of the above components. Define

$$W = \pi_B^{-1}(T) \cap V.$$

We have that:

(i) $\deg W$ is bounded in terms of $\deg V$, $\deg B$ and $\deg A$. Indeed, by Bézout's theorem, $\deg W \le \deg \pi_B^{-1}(T) \deg V$. By Estimate 3.3, $\deg T$ is bounded in terms of the degree and the dimension of V_π and thus, by (a), in terms of $\deg B$, $\deg V$ and $c(A)$.

(ii) $\dim W < n$ because V is weak-transverse in A and so it is not contained in $\pi_B^{-1}(T)$.

(iii) For $\zeta \in \pi_B^{-1}(T) \cap Z_{V,A}$, each maximal V-torsion anomalous component Y_ζ of $V \cap (B + \zeta)$ is contained in W; indeed, $\pi_B(Y_\zeta) = \pi_B(\zeta) \in T$.

By (iii), the variety W contains all the Y_ζ that we are counting; however, W is not necessarily irreducible. Therefore we cannot hope to use the inductive hypothesis on W and we have to consider its irreducible components.

Let $X_1, \dots, X_s$ be the irreducible components of W. For $\zeta \in \pi_B^{-1}(T) \cap Z_{V,A}$, we denote by Y_ζ any maximal V-torsion anomalous component of $V \cap (B + \zeta)$. By (iii), clearly each Y_ζ is contained in some X_i. We are going to count the number of Y_ζ contained in each X_i.

Denote

$$Z_{X_j} = \{\zeta \in \pi_B^{-1}(T) \cap Z_{V,A} \mid X_j \text{ contains some } Y_\zeta\}/(B \cap B^\perp).$$

Then

$$\pi_B^{-1}(T) \cap Z_{V,A} = \bigcup_{j=1}^{s} Z_{X_j}.$$

The number s of irreducible components of W is bounded by $\deg W$. Thus, by (i), s is effectively bounded only in terms of $\deg V$, $\deg B$ and $c(A)$.

To conclude our proof we are left to bound in an effective way the cardinality of each Z_X for X running over all irreducible components of W.

If X does not contain any Y_ζ, then $|Z_X| = 0$.

If $X = Y_{\zeta_0}$ for some $\zeta_0 \in Z_{V,A}$, then $|Z_X| = 1$.

Suppose that X strictly contains Y_{ζ_0} for some $\zeta_0 \in Z_{V,A}$. In this case we are going to show that $|Z_X| \le |Z_{X-\zeta_0, \pi_B^{-1}(T)-\zeta_0}|$. Applying the inductive hypothesis, we estimate $|Z_{X-\zeta_0, \pi_B^{-1}(T)-\zeta_0}|$ in terms of $\deg V$, $\deg B$ and $c(A)$.

We first verify that $X - \zeta_0$ in $\pi_B^{-1}(T) - \zeta_0$ satisfies the assumption of the inductive hypothesis, that is, the assumption of the proposition with $\dim X < n$. Observe that we need to translate by ζ_0 in order to obtain ambient varieties which are abelian varieties.

- The variety $\pi_B^{-1}(T) - \zeta_0$ is an abelian variety containing B. Indeed $B + \zeta_0$ is a subvariety of $\pi_B^{-1}(T)$, and $\zeta_0 \in \pi_B^{-1}(T)$.
- The variety $X - \zeta_0$ is weak-transverse in $\pi_B^{-1}(T) - \zeta_0$. Equivalently, by Remark 2.4, we show that X is weak-transverse in $\pi_B^{-1}(T)$. Since Y_{ζ_0} is a maximal V-torsion anomalous variety and X strictly contains Y_{ζ_0}, then X cannot be V-torsion anomalous. Recall that X is a component of $V \cap \pi_B^{-1}(T)$. Thus

$$\dim \pi_B^{-1}(T) - \dim X = \dim A - \dim V. \tag{2}$$

  If X was not weak-transverse in $\pi_B^{-1}(T)$, then $X \subseteq B_1 \cap V$ with $B_1 \subsetneq \pi_B^{-1}(T)$ a torsion variety. This contradicts relation (2).
- By (ii), $\dim X \le \dim W < n$.

Thus, by inductive hypothesis, we get that

$$|Z_{X-\zeta_0, \pi_B^{-1}(T)-\zeta_0}| \text{ is effectively bounded in terms of } \deg X, \deg B, c(\pi_B^{-1}(T)).$$

We now show that by our construction $\deg X$ and $c(\pi_B^{-1}(T))$ only depend on $\deg V$, $\deg B$ and A.

- By (i), $\deg X \leq \deg W$ is effectively bounded in terms of $\deg V$, $\deg B$ and $c(A)$.
- By Estimate 3.2, we know that $\deg T$ is effectively bounded in terms of $\deg V_\pi$ and $\dim V$. Moreover, by (a), $\deg V_\pi$ is bounded in terms of $\deg V$, $\deg B$ and $c(A)$. Finally, Estimate 3.1 ensures that $h_1(\pi_B^{-1}(T))$ and $d_{\pi_B^{-1}(T)}$ are effectively bounded in terms of $\deg V$, $\deg B$ and $c(A)$.

Therefore,

$$|Z_{X-\zeta_0,\pi_B^{-1}(T)-\zeta_0}| \text{ is effectively bounded in terms of } \deg V, \deg B \text{ and } c(A).$$

We finally prove that

$$|Z_X| \leq |Z_{X-\zeta_0,\pi_B^{-1}(T)-\zeta_0}|.$$

We shall show that for every maximal V-torsion anomalous variety $Y_\zeta \subseteq X$, the variety $Y_\zeta - \zeta_0$ is a maximal $(X-\zeta_0)$-torsion anomalous variety in $\pi^{-1}(T)-\zeta_0$.

Clearly $Y_\zeta - \zeta_0 \subseteq (X-\zeta_0) \cap (B+\zeta-\zeta_0)$. Since Y_ζ is V-torsion anomalous we have

$$\dim B - \dim Y_\zeta < \dim A - \dim V.$$

From this and (2) we obtain

$$\dim B - \dim Y_\zeta < \dim A - \dim V = \dim \pi_B^{-1}(T) - \dim X.$$

Thus $Y_\zeta - \zeta_0$ is a $(X-\zeta_0)$-torsion anomalous variety.

In addition, $Y_\zeta - \zeta_0$ is maximal: let $Y' \supset Y_\zeta - \zeta_0$ be a maximal $(X-\zeta_0)$-torsion anomalous variety and let $B' + \zeta'$ be minimal for Y'. From (2), we have

$$\dim B' - \dim Y' < \dim \pi_B^{-1}(T) - \dim X = \dim A - \dim V.$$

Thus $Y' + \zeta_0 \subseteq V \cap (B' + \zeta' + \zeta_0)$ is V-torsion anomalous and contains Y_ζ. The maximality of Y_ζ as V-torsion anomalous implies $Y' + \zeta_0 = Y_\zeta$.

In conclusion, collecting all our bounds, we have proven that $|Z_{V,A}|$ is effectively bounded in terms of $\deg V$, $\deg B$ and $c(A)$. □

4. Nontranslate torsion anomalous varieties

Proof of Theorem 1.4, part (1). Let Y be a maximal V-torsion anomalous variety which is not a translate, and so of positive dimension. Let $B+\zeta$ be minimal for Y. We use the arithmetic Bézout theorem and the Bogomolov bound to prove that $\deg B$ is bounded only in terms of V and A, then we deduce the bounds for $h(Y)$ and $\deg Y$.

By Lemma 2.1, Y is weak-transverse in $B+\zeta$, and by assumption $\dim B = \dim Y + 1$; therefore, Y is transverse in $B+\zeta$. Applying the Bogomolov estimate (Theorem 2.7) to Y in $B+\zeta$, we get

$$\frac{(\deg B)^{1-\eta}}{(\deg Y)^{1+\eta}} \ll_\eta \mu(Y). \tag{3}$$

Let $h_1,\dots,h_r$ be the forms of increasing degrees d_i such that $B+\zeta$ is a component of their zero set. We have that $r \le \operatorname{codim} B \le rg$ and

$$d_1\cdots d_r \ll \deg(B+\zeta) = \deg B \ll d_1\cdots d_r. \tag{4}$$

Consider the algebraic subgroup given by the first $h_1\cdots h_{r-1}$ forms, and let B' be one of its irreducible components containing $B+\zeta$. Then by (4) we have

$$\deg B' \ll d_1\cdots d_{r-1} \ll (\deg B)^{(r-1)/r}$$

and $\operatorname{codim} B' \ge \operatorname{codim} B - g$.

Since $\operatorname{codim} V \ge g+1 = g+\dim B - \dim Y$, this implies that $\operatorname{codim} B' \ge \dim V - \dim Y$, and thus, by Lemma 2.2, Y is a component of $V\cap B'$.

We apply the arithmetic Bézout theorem to $V \cap B'$ and recall that $h(B')=0$ because B' is a torsion variety; we get

$$h(Y) \ll (h(V)+\deg V)\deg B' \ll (h(V)+\deg V)(\deg B)^{(r-1)/r}. \tag{5}$$

Zhang's inequality, with (3) and (5), gives

$$\frac{(\deg B)^{1-\eta}}{(\deg Y)^{1+\eta}} \ll_\eta \mu(Y) \ll (h(V)+\deg V)\frac{(\deg B)^{(r-1)/r}}{\deg Y}.$$

Recall that Y is a component of $V\cap(B+\zeta)$. By Bézout's theorem, $\deg Y \le \deg B \deg V$, thus

$$(\deg B)^{1-\eta} \ll_\eta (h(V)+\deg V)(\deg B)^{(r-1)/r}(\deg B\deg V)^\eta,$$

and therefore

$$(\deg B)^{1/r-2\eta} \ll_\eta (h(V)+\deg V)(\deg V)^\eta.$$

For η small enough, we get

$$\deg B \ll_\eta (h(V)+\deg V)^{r+\eta}(\deg V)^\eta; \tag{6}$$

this proves that the degree of B is bounded only in terms of V and A. Since there are finitely many abelian subvarieties of bounded degree, applying Proposition 3.5, we conclude that ζ belongs to a finite set of cardinality effectively bounded.

The bound on the height of Y is now given by (5) and (6):

$$h(Y) \ll_\eta (h(V)+\deg V)^{r+\eta}(\deg V)^\eta.$$

Finally, the bound on the degree is obtained from (6) and Bézout's theorem for the component Y of $V \cap B'$:

$$\deg Y \ll_\eta (h(V)+\deg V)^{r-1+\eta}(\deg V). \qquad \square$$

5. Torsion anomalous translates

Proof of Theorem 1.4, parts (2) and (3). Let Y be a maximal V-torsion anomalous translate with $B+\zeta$ minimal for Y.

We proceed to bound $\deg B$ and, in turn, the height and the degree of Y, using the Lehmer estimate and the arithmetic Bézout theorem.

The variety $B+\zeta$ is a component of the torsion variety defined as the zero set of forms $h_1, \dots, h_r$ of increasing degrees d_i, and

$$d_1 \cdots d_r \ll \deg B = \deg(B+\zeta) \ll d_1 \cdots d_r.$$

We have that $r \le \operatorname{codim} B \le rg$.

Consider the torsion variety defined as the zero set of the first $r-1$ forms $h_1, \dots, h_{r-1}$, and take a connected component B' containing $B+\zeta$, so that $\deg B' \ll d_1 \cdots d_{r-1} \ll (\deg B)^{(r-1)/r}$ and $\operatorname{codim} B' \ge \operatorname{codim} B - g$.

By Lemma 2.2, Y is a component of $V \cap B'$; indeed

$$\operatorname{codim} B' \ge \operatorname{codim} B - g = \dim A - g - \dim Y - 1 > \dim V - \dim Y - 1.$$

The proof is now divided in two cases, depending on $\dim Y$. If Y has dimension zero we use the arithmetic Bézout theorem and the Lehmer estimate; if $Y = H + Y_0$ is a translate of positive dimension, we can reduce to the zero dimensional case using some properties of the essential minimum.

Proof of part (2). Consider first the case of a maximal torsion anomalous point Y.

All conjugates of Y over $k_{\mathrm{tor}}(V)$ are components of $V \cap (B+\zeta)$; they all have the same normalised height and their number is at least

$$[k_{\mathrm{tor}}(V,Y) : k_{\mathrm{tor}}(V)] \ge \frac{[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}]}{[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]}.$$

We then apply the arithmetic Bézout theorem in $V \cap B'$, obtaining

$$[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}]\,\hat{h}(Y) \ll (h(V)+\deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}](\deg B)^{(r-1)/r}. \tag{7}$$

Applying Theorem 2.8 to Y in $B+\zeta$, we obtain that, for every positive real η,

$$\hat{h}(Y) \gg_\eta \frac{(\deg B)^{1-\eta}}{[k_{\mathrm{tor}}(Y):k_{\mathrm{tor}}]^{1+\eta}}. \tag{8}$$

Combining (8) and (7), we have

$$\begin{aligned}\frac{(\deg B)^{1-\eta}}{[k_{\mathrm{tor}}(Y):k_{\mathrm{tor}}]^{\eta}} &\ll_\eta [k_{\mathrm{tor}}(Y):k_{\mathrm{tor}}]\,\hat{h}(Y)\\ &\ll (h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}](\deg B)^{(r-1)/r}.\end{aligned}$$

For η small enough, we obtain

$$\deg B \ll_\eta \big((h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]\big)^{r+\eta}[k_{\mathrm{tor}}(Y):k_{\mathrm{tor}}]^{\eta}. \tag{9}$$

Apply now Bézout's theorem to $V\cap B'$. All the conjugates of Y over $k_{\mathrm{tor}}(V)$ are components of this intersection, so

$$\frac{[k_{\mathrm{tor}}(Y):k_{\mathrm{tor}}]}{[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]} \ll_\eta (\deg B)^{(r-1)/r}(\deg V). \tag{10}$$

Substituting (9) into (10) we have the last bound of part (2) in the statement.

Finally, we apply the arithmetic Bézout theorem to $V\cap B'$ to get

$$\begin{aligned}\hat{h}(Y) &\ll (h(V)+\deg V)(\deg B)^{(r-1)/r}\\ &\ll_\eta (h(V)+\deg V)^{r+\eta}[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]^{r-1+\eta}.\end{aligned}$$

Having bounded $\deg B$, in view of Proposition 3.5 the points ζ belong to a finite set of cardinality effectively bounded.

Proof of part (3). Assume now that Y is a translate of positive dimension and write $Y=H+Y_0$, with H an abelian variety and Y_0 a point in $H^\perp$.

To bound $\deg B$ we can assume, without loss of generality, that $\zeta=0$ (see Remark 2.4). By Lemma 2.3,

$$\mu(Y_0)=\mu(H+Y_0). \tag{11}$$

Since the intersection $V\cap B'$ is defined over $k_{\mathrm{tor}}(V)$, every conjugate of $H+Y_0$ over $k_{\mathrm{tor}}(V)$ is a component of $V\cap B'$; as before, such components have the same normalised height and their number is at least

$$\frac{[k_{\mathrm{tor}}(H+Y_0):k_{\mathrm{tor}}]}{[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]}.$$

We apply the arithmetic Bézout theorem in $V\cap B'$ and we obtain

$$h(H+Y_0)\frac{[k_{\mathrm{tor}}(H+Y_0):k_{\mathrm{tor}}]}{[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]} \ll (h(V)+\deg V)(\deg B)^{(r-1)/r}. \tag{12}$$

By Zhang's inequality, (11) and (12), we deduce

$$\mu(Y_0) \ll \frac{(h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}](\deg B)^{(r-1)/r}}{[k_{\mathrm{tor}}(H+Y_0):k_{\mathrm{tor}}]\deg H}. \tag{13}$$

The lower bound for $\mu(Y_0)$ is derived as in the case of dimension zero.

Consider the smallest abelian subvariety H_0 of B containing Y_0. Clearly H_0 is the irreducible component of $H^{\perp} \cap B$ containing Y_0. Indeed, they are both one-dimensional abelian varieties containing the point Y_0 of infinite order.

By the definition of H_0, we have $B = H + H_0$, and from [Masser and Wüstholz 1993, Lemma 1.2], we obtain

$$\#(H \cap H_0)\deg B \leq \deg H \deg H_0. \tag{14}$$

Moreover, from $H \cap H_0 \subseteq H \cap H^{\perp}$, we get

$$\#(H \cap H_0) \leq \#(H \cap H^{\perp}) \ll (\deg H)^2. \tag{15}$$

Applying Theorem 2.8 to Y_0 in H_0 we get that, for every positive real η,

$$\mu(Y_0) = \hat{h}(Y_0) \gg_{\eta} \frac{(\deg H_0)^{1-\eta}}{[k_{\mathrm{tor}}(Y_0):k_{\mathrm{tor}}]^{1+\eta}}. \tag{16}$$

We remark that

$$[k_{\mathrm{tor}}(Y_0):k_{\mathrm{tor}}] \leq [k_{\mathrm{tor}}(H+Y_0):k_{\mathrm{tor}}]\cdot \#(H \cap H_0) \tag{17}$$

because if σ is in $\mathrm{Gal}(\bar{k}_{\mathrm{tor}}/k_{\mathrm{tor}}(H+Y_0))$, then $\sigma(Y_0) - Y_0$ is in $H \cap H_0$, so $[k_{\mathrm{tor}}(Y_0):k_{\mathrm{tor}}(H+Y_0)] \leq \#(H \cap H_0)$.

Combining the upper bound and the lower bound for $\mu(Y_0)$ in (13) and (16), and using also (14), (15) and (17), for η sufficiently small, we have

$$\deg B \ll_{\eta} \big((h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]\big)^{r+\eta}, \tag{18}$$

where the dependence on $\deg H[k_{\mathrm{tor}}(H+Y_0):k_{\mathrm{tor}}]$ has been removed by applying Bézout's theorem to the intersection $V \cap B'$.

This also gives

$$\deg(H+Y_0) \ll_{\eta} (\deg V)\big((h(V)+\deg V)[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]\big)^{r-1+\eta}.$$

Finally, from (12), (18) and the trivial bound $[k_{\mathrm{tor}}(H+Y_0):k_{\mathrm{tor}}] \geq 1$, we obtain

$$h(H+Y_0) \ll_{\eta} (h(V)+\deg V)^{r+\eta}[k_{\mathrm{tor}}(V):k_{\mathrm{tor}}]^{r-1+\eta}.$$

Since we have bounded $\deg B$, we can conclude from Proposition 3.5 that the points ζ belong to a finite set of cardinality effectively bounded. □

6. The case of a curve and applications to the effective Mordell–Lang conjecture

Recall that $A=\prod_{i=1}^{\ell} A_i^{e_i}$ with A_i nonisogenous simple CM factors of dimension g_i. To prove Theorem 1.5 we essentially follow the proof of Theorem 1.4, part (2), working first in the projection on one factor, and then lifting the construction to the abelian variety A.

Proof of Theorem 1.5. Clearly all the points in $\mathcal{S}(C)$ are C-torsion anomalous; in addition, since C is a weak-transverse curve, each torsion anomalous point is maximal.

If $Y \in \mathcal{S}(C)$, then $Y \in C \cap H$, with $H = \prod_i H_i$ the subgroup containing Y which is minimal with respect to the inclusion.

Denote by Y_i the projection of Y on H_i and by C_i the projection of C on $A_i^{e_i}$. Let j be one of the indices satisfying the hypothesis of the theorem. Assume first that Y_j is a torsion point, and define $H' = A_1^{e_1} \times \cdots \times \{Y_j\} \times \cdots \times A_\ell^{e_\ell}$. Clearly $\deg H' \ll 1$ and $h(H') = 0$. Then, applying the arithmetic Bézout theorem to Y in $C \cap H'$, we get $\hat{h}(Y) \ll (h(C) + \deg C)$.

Assume now that Y_j is not a torsion point. Let $B_j + \zeta_j$ be a component of H_j containing Y_j. Clearly $\dim B_j = \dim H_j$ and $Y_j \in C_j \cap (B_j + \zeta_j)$ with $B_j + \zeta_j$ minimal for Y_j. Furthermore, Y_j is a component of $C_j \cap (B_j + \zeta_j)$ because C_j is weak-transverse and, by assumption, $\operatorname{codim} H_j > g_j \dim H_j > 0$. This ensures that the matrix associated to $B_j + \zeta_j$ has at least two rows, which is necessary to apply the method.

We now sketch the proof, which follows the proof of Theorem 1.4, part (2), and we give the relevant bounds.

The variety $B_j + \zeta_j$ is a component of the zero set of forms $h_1, \dots, h_r$ of increasing degrees d_j with

$$d_1 \cdots d_r \ll \deg B_j = \deg(B_j + \zeta_j) \ll d_1 \cdots d_r,$$

and we have that $r = \operatorname{codim} B_j / g_j = \operatorname{codim} H_j / g_j$.

Consider the torsion variety defined as the zero set of h_1, and let B_j' be one of its connected components containing $B_j + \zeta_j$; then $\deg B_j' \ll d_1 \ll (\deg B_j)^{1/r} = (\deg B_j)^{g_j/\operatorname{codim} B_j}$.

From Theorem 2.8 applied to Y_j in $B_j + \zeta_j$, for every positive real η, we get

$$\hat{h}(Y_j) \gg_\eta \frac{(\deg B_j)^{1/\dim B_j - \eta}}{[k_{\mathrm{tor}}(Y_j) : k_{\mathrm{tor}}]^{1/\dim B_j + \eta}}. \tag{19}$$

Notice that all conjugates of Y_j over $k_{\mathrm{tor}}(C_j)$ are components of $C_j \cap B_j'$ and they all have the same height. Applying the arithmetic Bézout theorem to $C_j \cap B_j'$

and arguing as in the proof of Theorem 1.4, we have

$$\hat{h}(Y_j)\frac{[k_{\mathrm{tor}}(Y_j):k_{\mathrm{tor}}]}{[k_{\mathrm{tor}}(C_j):k_{\mathrm{tor}}]} \ll (h(C_j)+\deg C_j)(\deg B_j)^{g_j/\dim B_j}. \tag{20}$$

Recall that $[k_{\mathrm{tor}}(Y_j):k_{\mathrm{tor}}]\geq 1$. From (19) and (20) we get

$$(\deg B_j)^{\frac{\operatorname{codim} B_j - g_j \dim B_j}{\operatorname{codim} B_j \dim B_j}-\eta} \ll_\eta [k_{\mathrm{tor}}(C_j):k_{\mathrm{tor}}](h(C_j)+\deg C_j)[k_{\mathrm{tor}}(Y_j):k_{\mathrm{tor}}]^{1/\dim B_j - 1+\eta}.$$

Since $\operatorname{codim} B_j - g_j \dim B_j = \operatorname{codim} H_j - g_j \dim H_j \geq 1$, for η sufficiently small this yields

$$\deg B_j \ll_\eta \big([k_{\mathrm{tor}}(C_j):k_{\mathrm{tor}}](h(C_j)+\deg C_j)\big)^{\frac{\operatorname{codim} H_j \dim H_j}{\operatorname{codim} H_j - g_j \dim H_j}+\eta}, \tag{21}$$

where if $\dim B_j>1$, we use $[k_{\mathrm{tor}}(Y_j):k_{\mathrm{tor}}]\geq 1$, and if $\dim B_j=1$, we use $[k_{\mathrm{tor}}(Y_j):k_{\mathrm{tor}}]\leq [k_{\mathrm{tor}}(C_j):k_{\mathrm{tor}}]\deg B \deg C_j$.

We now lift the construction to A as follows. Define $H'=A_1^{e_1}\times\cdots\times B_j'\times\cdots\times A_\ell^{e_\ell}$. Clearly $\deg H'\leq \deg A \deg B_j'$ and Y is a component of $C\cap H'$. Applying the arithmetic Bézout theorem to $C\cap H'$ and using (21), we obtain

$$\begin{aligned}\hat{h}(Y) &\ll (h(C)+\deg C)\deg H' \ll (h(C)+\deg C)(\deg B_j)^{g_j/\dim H_j}\\ &\ll_\eta (h(C)+\deg C)^{\frac{\operatorname{codim} H_j}{\operatorname{codim} H_j - g_j\dim H_j}+\eta}[k_{\mathrm{tor}}(C):k_{\mathrm{tor}}]^{\frac{g_j \dim H_j}{\operatorname{codim} H_j - g_j \dim H_j}+\eta}.\end{aligned} \tag{22}$$

□

6.1. *An application to the effective Mordell–Lang conjecture.*

Proof of Corollary 1.6. Let $x\in C\cap\Gamma$. Let j be an index such that $e_j/(g_j+1)>t_j$ and denote by $(x_1,\dots,x_{e_j})$ the projection of x in Γ_j.

Let $\gamma_1,\dots,\gamma_{t_j}$ be generators of the free part of $\overline{\Gamma}_j$. Then there exist elements $0\neq a_k\in \operatorname{End}(A_j)$ for $k=1,\dots,e_j$, an $e_j\times t_j$-matrix M_j with coefficients in $\operatorname{End}(A_j)$ and a torsion point $\zeta\in A_j^{e_j}$ such that

$$(a_1x_1,\dots,a_{e_j}x_{e_j})^t = M_j(\gamma_1,\dots,\gamma_{t_j})^t+\zeta.$$

If the rank of M_j is zero, then $(x_1,\dots,x_{e_j})$ is a torsion point and so has height zero.

If M_j has positive rank r_j, we can choose r_j equations of the system corresponding to r_j linearly independent rows of M_j. We use these equations to write the γ_k in terms of the x_k and we substitute these expressions in the remaining equations. We obtain a system of maximal rank with $e_j-r_j\geq e_j-t_j$ linearly independent equations in the variables $x_1, \dots, x_{e_j}$:

$$\begin{cases} m_{11}x_1 + \cdots + m_{1e_j}x_{e_j} = \xi_1, \\ \quad \vdots \\ m_{e_j-r_j,1}x_1 + \cdots + m_{e_j-r_j,e_j}x_{e_j} = \xi_{e_j-r_j}, \end{cases}$$

where $\xi_k \in A_j$ are torsion points and $m_{k\ell} \in \mathrm{End}(A_j)$. These equations define a torsion variety $H_j \subseteq A_j^{e_j}$. Since $(g_j+1)t_j < e_j$ we have $\operatorname{codim} H_j > g_j \dim H_j$.

Then $x \in C \cap H$, where H satisfies the hypothesis of Theorem 1.5, which gives the bound for the height of x. □

Remark 6.1. Notice that it is possible to apply the corollary also to subgroups Γ whose rank is bounded only on one projection.

For example, let E_1, E_2 be two elliptic curves defined over $\mathbb{Q}$ and such that $E_1(\mathbb{Q})$ is an abelian group of rank 1, and consider the product $A = E_1^4 \times E_2$.

Let C be a weak-transverse curve in A. Consider the subgroup $\Gamma = E_1(\mathbb{Q})^4 \times E_2(\overline{\mathbb{Q}})$ of A. Then Γ is not of finite rank, but with the notation of the corollary, we have $g_1 = 1$, $e_1 = 4$, $t_1 = 1$ and $t_1 < e_1/(g_1+1) = \frac{4}{2}$.

The hypothesis of the corollary is therefore verified, and we have that

$$\hat{h}(C \cap \Gamma) \ll_\eta (h(C) + \deg C)^{3/2+\eta}[k_{\mathrm{tor}}(C) : k_{\mathrm{tor}}]^{1/2+\eta}.$$

7. From curves to varieties

We now adapt the proof strategy of Theorem 1.4 to obtain some new results for varieties V of dimension > 1 embedded in a power E^n of a CM elliptic curve. For subvarieties of general CM abelian varieties some technical conditions arise. This makes a straightforward generalisation of our method of little interest.

For torsion anomalous varieties which are translates, the proof can be easily adapted, while for nontranslates a new argument is needed. Indeed, in this last case, the torsion anomalous variety is not transverse, but only weak-transverse in its minimal variety, a condition which is not sufficient to use the sharp Bogomolov bound.

The torsion varieties contained in V are already covered by the Manin–Mumford conjecture, therefore we restrict ourselves to the V-torsion anomalous varieties which are not torsion.

Theorem 7.1. *Let E be a CM elliptic curve defined over a number field k and let $n > 1$ be an integer. Denote by k_{tor} a field of definition of all torsion points of E.*

Let $V \subseteq E^n$ be a weak-transverse variety. Let $Y \subseteq V \cap B + \zeta$ be a maximal V-torsion anomalous variety which is not a torsion variety, and let $B+\zeta$ be minimal for Y.

Set $b = \dim B$, $v = \dim V$ and $y = \dim Y$ and assume that $(n-b) > (v-y)(b-y)$.

Then for any $\eta > 0$, there exist constants depending only on E^n and η such that:

(1) *If Y is a point, then*

$$\deg B \ll_\eta \big((h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]\big)^{\frac{(n-b)b}{(n-b)-vb}+\eta},$$
$$\hat{h}(Y) \ll_\eta (h(V) + \deg V)^{\frac{n-b}{(n-b)-vb}+\eta}[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]^{\frac{vb}{(n-b)-vb}+\eta},$$
$$[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}] \ll_\eta \deg V(h(V) + \deg V)^{\frac{vb}{(n-b)-vb}+\eta}[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]^{\frac{n-b}{(n-b)-vb}+\eta}.$$

(2) *If Y is a translate of positive dimension, then*

$$\deg B \ll_\eta \big((h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]\big)^{\frac{(n-b)(b-y)}{(n-b)-(v-y)(b-y)}+\eta},$$
$$h(Y) \ll_\eta (h(V) + \deg V)^{\frac{n-b}{(n-b)-(v-y)(b-y)}+\eta}[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]^{\frac{(v-y)(b-y)}{(n-b)-(v-y)(b-y)}+\eta},$$
$$\deg Y \ll_\eta \deg V\big((h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]\big)^{\frac{(v-y)(b-y)}{(n-b)-(v-y)}+\eta}.$$

(3) *If Y is not a translate, then*

$$\deg B \ll_\eta \big((h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]\big)^{\frac{(b-y)(n-b)}{(n-b)-(v-y)(b-y)}+\eta},$$
$$h(Y) \ll_\eta (h(V) + \deg V)^{\frac{(n-b)}{(n-b)-(v-y)(b-y)}+\eta}[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]^{\frac{(v-y)(b-y)}{(n-b)-(v-y)(b-y)}+\eta},$$
$$\deg Y \ll_\eta \deg V((h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}])^{\frac{(v-y)(b-y)}{(n-b)-(v-y)(b-y)}+\eta}.$$

In addition the torsion points ζ belong to a finite set.

Proof of Theorem 7.1, part (1). Let Y be a maximal V-torsion anomalous point with $B+\zeta$ minimal for Y.

We proceed to bound $\deg B$ and, in turn, the height of Y and the degree of its field of definition. To this aim we use the Lehmer bound in Theorem 2.8 and the arithmetic Bézout theorem.

Let $v = \dim V$ and $b = \dim B$. By Lemma 2.2, Y is a component of $V \cap B'$ where B' is, like in the proof of Theorem 1.4, the zero component of the torsion variety defined by the first v rows $h_1, \dots, h_v$ of the matrix of B. Then $\operatorname{codim} B' = v$ and $\deg B' \ll (\deg B)^{v/(n-b)}$.

We apply the arithmetic Bézout theorem to $V \cap B'$ to obtain

$$\hat{h}(Y) \ll \frac{(h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]}{[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}]}(\deg B)^{v/(n-b)}. \tag{23}$$

Applying the Lehmer estimate in Theorem 2.8 to Y in $B+\zeta$, instead, we have that for every positive real η,

$$\hat{h}(Y) \gg_\eta \frac{(\deg B)^{1/b-\eta}}{[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}]^{1/b+\eta}}. \tag{24}$$

From (23) and (24), and for η small enough, we get the bound for $\deg B$ if $b > 1$. If $b = 1$ we use Bézout's theorem to bound the factor $[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}]^{\eta}$ as $\big((\deg B)(\deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]\big)^{\eta}$.

We then apply Bézout's theorem in $V \cap B'$ to bound $[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}]$ and the arithmetic Bézout theorem in $V \cap B'$ to prove the bound for $\hat{h}(Y)$. Finally, from Proposition 3.5 it follows that the points ζ belong to a finite set of cardinality effectively bounded. □

Proof of Theorem 7.1, part (2). Let $Y = H + Y_0$ be a maximal V-torsion anomalous translate of positive dimension with minimal $B + \zeta$; assume also that $Y_0 \in H^{\perp}$.

We use the Lehmer bound and the arithmetic Bézout theorem to bound $\deg B$ and, in turn, the height and the degree of $H + Y_0$. In view of Remark 2.4, without loss of generality, we can assume that $\zeta = 0$.

Let $b = \dim B$, $v = \dim V$ and $y = \dim Y = \dim H$. Clearly $v - y < n - b$ because Y is torsion anomalous.

As before, by Lemma 2.2 we have that Y is a component of $V \cap B'$, where B' is an irreducible torsion variety with $\operatorname{codim} B' = v - y$ and $\deg B' \ll (\deg B)^{(v-y)/(n-b)}$. Arguing as usual on the conjugates of $H + Y_0$ over $k_{\mathrm{tor}}(V)$, we see that the intersection $V \cap B'$ has at least $[k_{\mathrm{tor}}(H + Y_0) : k_{\mathrm{tor}}]/[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]$ components.

We apply the arithmetic Bézout theorem to the intersection $V \cap B'$, obtaining

$$h(H + Y_0) \ll (h(V) + \deg V)(\deg B)^{(v-y)/(n-b)} \frac{[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]}{[k_{\mathrm{tor}}(H + Y_0) : k_{\mathrm{tor}}]}. \tag{25}$$

By Zhang's inequality, Lemma 2.3 and (25), we deduce

$$\mu(H + Y_0) = \mu(Y_0) \ll \frac{(h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}](\deg B)^{(v-y)/(n-b)}}{[k_{\mathrm{tor}}(H + Y_0) : k_{\mathrm{tor}}] \deg H}. \tag{26}$$

For the lower bound for $\mu(Y_0)$, the proof follows the case of $\dim Y = 0$. Let $H_0 = H^{\perp} \cap B$. By minimality of B we have that H_0 is a torsion variety of minimal dimension containing Y_0, thus

$$\dim H_0 = \dim H^{\perp} + \dim B - n = (n - y) + b - n = b - y.$$

As in Theorem 1.4, part (3), one can easily see that

$$[k_{\mathrm{tor}}(Y_0) : k_{\mathrm{tor}}] \le [k_{\mathrm{tor}}(H + Y_0) : k_{\mathrm{tor}}] \cdot \#(H \cap H_0). \tag{27}$$

By the definition of H_0, we have $B = H + H_0$ and from [Masser and Wüstholz 1993, Lemma 1.2], we get

$$\#(H \cap H_0) \deg B \le \deg H \deg H_0. \tag{28}$$

In addition, $H \cap H_0 \subseteq H \cap H^{\perp}$, thus

$$\#(H \cap H_0) \ll (\deg H)^2. \tag{29}$$

Applying Theorem 2.8 to Y_0 in H_0 we get that, for every positive real η,

$$\mu(Y_0) = \hat{h}(Y_0) \gg_\eta \frac{(\deg H_0)^{1/(b-y)-\eta}}{[k_{\mathrm{tor}}(Y_0) : k_{\mathrm{tor}}]^{1/(b-y)+\eta}}. \tag{30}$$

Combining (26) and (30), we have

$$\frac{(\deg H_0)^{1/(b-y)-\eta}}{[k_{\mathrm{tor}}(Y_0) : k_{\mathrm{tor}}]^{1/(b-y)+\eta}} \ll_\eta \frac{(h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}](\deg B)^{(v-y)/(n-b)}}{[k_{\mathrm{tor}}(H + Y_0) : k_{\mathrm{tor}}] \deg H},$$

and hence, using (27)–(29) as in Theorem 1.4, part (3), we get the bound for $\deg B$; notice that if $b - y > 1$, the argument is in fact simpler, as we don't need to deal with the $[k_{\mathrm{tor}}(Y_0) : k_{\mathrm{tor}}]^\eta$ term.

Having obtained a bound for $\deg B$, the degree of $H + Y_0$ can be bounded by applying Bézout's theorem to the intersection $V \cap B'$ and using $\deg B' \ll \deg B^{(v-y)/(n-b)}$. The bound for $h(H + Y_0)$, instead, is derived from (25) and the bound for $\deg B$. Finally, from Proposition 3.5 we conclude that the points ζ belong to a finite set of cardinality effectively bounded. □

Proof of Theorem 7.1, part (3). Assume that $Y \subseteq V \cap (B + \zeta)$ is not a translate.

If Y is transverse in $B + \zeta$, the proof of Theorem 1.4, part (1) easily adapts to this case as well, yielding the desired bounds; let us then assume that Y is not transverse. Without loss of generality, we can assume $\zeta = 0$ (see Remark 2.4). Then Y is transverse in a translate $H_1 + Y_0 \subsetneq B$, with $Y_0 \in H_1^{\perp}$ and H_1 of minimal dimension.

We define $H_0 = B \cap H_1^{\perp}$ so that $B = H_1 + H_0$ and

$$\deg B = \deg(H_1 + H_0) \le \frac{\deg H_1 \deg H_0}{\#(H_1 \cap H_0)}. \tag{31}$$

We set $y = \dim Y$, $v = \dim V$, $b = \dim B$, $h_1 = \dim H_1$ and $h_0 = \dim H_0 = b - h_1$.

Writing $Y = Y_1 + Y_0$, we have that $Y_1 \subseteq H_1$ is transverse in H_1 because Y is transverse in $H_1 + Y_0$, and $Y_0 \subseteq H_0$ is transverse in H_0 because B is minimal for Y.

By definition $Y_1 \subseteq H_1$ and $Y_0 \in H_1^{\perp}$. From Lemma 2.3 and the definition of essential minimum, we get

$$\mu(Y) = \mu(Y_1) + \hat{h}(Y_0).$$

As usual, the upper bound for $\mu(Y)$ is obtained using the arithmetic Bézout theorem in $V \cap B'$ for some abelian variety B' constructed by deleting $v - y$

suitable rows from B. All conjugates of Y are components of same height in $V \cap B'$. This gives

$$\mu(Y) \ll (h(V) + \deg V)(\deg B)^{(v-y)/(n-b)} \frac{[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]}{\deg Y [k_{\mathrm{tor}}(Y_1 + Y_0) : k_{\mathrm{tor}}]}. \tag{32}$$

Moreover,

$$[k_{\mathrm{tor}}(Y_1 + Y_0) : k_{\mathrm{tor}}] \#(H_1 \cap H_0) \geq [k_{\mathrm{tor}}(Y_0) : k_{\mathrm{tor}}] \tag{33}$$

because for every $\sigma \in \mathrm{Gal}(\bar{k}_{\mathrm{tor}}/k_{\mathrm{tor}})$ which fixes $Y_1 + Y_0$, the difference $\sigma(Y_0) - Y_0$ lies in $H_1 \cap H_0$.

To obtain a lower bound for $\mu(Y)$ we either apply the Bogomolov bound to Y_1 in H_1 or the Lehmer estimate to Y_0 in H_0. These give

$$\frac{(\deg H_1)^{1/(h_1-y)-\eta}}{(\deg Y)^{1/(h_1-y)+\eta}} \ll_\eta \mu(Y_1) \leq \mu(Y) \tag{34}$$

and

$$\frac{(\deg H_0)^{1/h_0-\eta}}{[k_{\mathrm{tor}}(Y_0) : k_{\mathrm{tor}}]^{1/h_0+\eta}} \ll_\eta \hat{h}(Y_0) \leq \mu(Y). \tag{35}$$

We now relate the left-hand side to $\deg B$. Notice that either

$$(\deg B)^{(h_1-y)/(b-y)} < \deg H_1 \tag{i}$$

or

$$(\deg B)^{h_0/(b-y)} \leq \frac{\deg H_0}{\#(H_1 \cap H_0)}. \tag{ii}$$

Indeed if (i) and (ii) were both false, then

$$\deg B = (\deg B)^{\frac{h_1-y}{b-y} + \frac{h_0}{b-y}} > \frac{\deg H_1 \deg H_0}{\#(H_1 \cap H_0)},$$

which contradicts (31).

Assume that (i) holds. Then (32), (34), (i) and the fact that $n-b > (v-y) \cdot (b-y)$ give the bound

$$\deg B \ll_\eta \big((h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]\big)^{\frac{(b-y)(n-b)}{(n-b)-(v-y)(b-y)} + \eta},$$

where, if $h_1 - y = 1$, the factor $(\deg Y)^\eta$ has been removed by applying Bézout's theorem to Y in $V \cap B$ and changing η.

Assume that (ii) holds. Then (32), (35) (ii), the fact that $n - b > (v - y)(b - y)$ and (33) give the bound

$$\deg B \ll_\eta \big((h(V) + \deg V)[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}]\big)^{\frac{(b-y)(n-b)}{(n-b)-(v-y)(b-y)} + \eta},$$

where, if $h_0 = 1$ the dependence on $[k_{\mathrm{tor}}(Y_0) : k_{\mathrm{tor}}]$ can be removed using (33), bounding $[k_{\mathrm{tor}}(Y) : k_{\mathrm{tor}}]$ as $[k_{\mathrm{tor}}(V) : k_{\mathrm{tor}}] \deg V \deg B$ by the Bézout theorem applied to Y in $V \cap B$ and observing that

$$\#(H_1 \cap H_0) \leq \#(H_1 \cap H_1^{\perp}) \ll (\deg H_1)^2 \ll (\deg Y)^{2h_1} \leq (\deg V \deg B)^{2h_1}$$

because, since Y_1 is transverse in H_1, we have $H_1 = Y_1 + \cdots + Y_1$ (h_1 times), from which $\deg H_1 \ll (\deg Y)^{h_1}$.

So we have bounded $\deg B$. We obtain the bounds for $\deg Y$ and $h(Y)$ applying respectively the Bézout Theorem and the arithmetic Bézout theorem to the intersection $Y \subseteq V \cap B'$. Finally, Proposition 3.5 guarantees that the points ζ belong to a finite set of cardinality effectively bounded. □

Acknowledgements

We thank Francesco Veneziano for an accurate revision of an earlier version of this paper. We kindly thank the referee for his useful comments and corrections. Especially, we thank him for pointing out the effectivity question of Proposition 3.5.

References

[Bertrand 1986] D. Bertrand, "Relations d'orthogonalité sur les groupes de Mordell–Weil", pp. 33–39 in *Séminaire de théorie des nombres, Paris 1984–85*, edited by C. Goldstein, Progr. Math. **63**, Birkhäuser, Boston, 1986. MR 88h:11037 Zbl 0607.14014

[Bombieri and Zannier 1996] E. Bombieri and U. Zannier, "Heights of algebraic points on subvarieties of abelian varieties", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **23**:4 (1996), 779–792. MR 98j:11043 Zbl 0897.11020

[Bombieri et al. 2007] E. Bombieri, D. Masser, and U. Zannier, "Anomalous subvarieties: structure theorems and applications", *Int. Math. Res. Not.* **2007**:19 (2007), Art. ID #rnm057. MR 2008k:11060 Zbl 1145.11049

[Carrizosa 2009] M. Carrizosa, "Petits points et multiplication complexe", *Int. Math. Res. Not.* **2009**:16 (2009), 3016–3097. MR 2011c:11102 Zbl 1176.11025

[Checcoli et al. 2012] S. Checcoli, F. Veneziano, and E. Viada, "A sharp Bogomolov-type bound", *New York J. Math.* **18** (2012), 891–910. MR 2991428 Zbl 1276.11099

[Checcoli et al. 2014] S. Checcoli, F. Veneziano, and E. Viada, "On torsion anomalous intersections", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **25**:1 (2014), 1–36. MR 3180478 Zbl 06293297

[David and Philippon 2002] S. David and P. Philippon, "Minorations des hauteurs normalisées des sous-variétés de variétés abeliennes, II", *Comment. Math. Helv.* **77**:4 (2002), 639–700. MR 2004a:11055 Zbl 1030.11026

[Galateau 2010] A. Galateau, "Une minoration du minimum essentiel sur les variétés abéliennes", *Comment. Math. Helv.* **85**:4 (2010), 775–812. MR 2011i:11110 Zbl 1250.11071

[Habegger 2008] P. Habegger, "Intersecting subvarieties of $\mathbf{G}_m^n$ with algebraic subgroups", *Math. Ann.* **342**:2 (2008), 449–466. MR 2009f:14044 Zbl 1168.14019

[Masser and Wüstholz 1993] D. Masser and G. Wüstholz, "Periods and minimal abelian subvarieties", *Ann. of Math.* (2) **137**:2 (1993), 407–458. MR 94g:11040 Zbl 0796.11023

[Maurin 2008] G. Maurin, "Courbes algébriques et équations multiplicatives", *Math. Ann.* **341**:4 (2008), 789–824. MR 2009g:14026 Zbl 1154.14017

[Philippon 1991] P. Philippon, "Sur des hauteurs alternatives, I", *Math. Ann.* **289**:2 (1991), 255–283. MR 92m:11061 Zbl 0726.14017

[Philippon 1995] P. Philippon, "Sur des hauteurs alternatives, III", *J. Math. Pures Appl.* (9) **74**:4 (1995), 345–365. MR 97a:11098 Zbl 0878.11025

[Philippon 2012] P. Philippon, "Sur une question d'orthogonalité dans les puissances de courbes elliptiques", preprint, Institut de Mathématiques de Jussieu, 2012, http://hal.archives-ouvertes.fr/hal-00801376.

[Rémond 2000] G. Rémond, "Décompte dans une conjecture de Lang", *Invent. Math.* **142**:3 (2000), 513–545. MR 2002f:14058 Zbl 0972.11054

[Shafarevich 1972] I. R. Shafarevich, *Basic algebraic geometry, 1: Varieties in projective space*, Springer, Heidelberg, 1972. MR 3100243 Zbl 0284.14001

[Viada 2008] E. Viada, "The intersection of a curve with a union of translated codimension-two subgroups in a power of an elliptic curve", *Algebra Number Theory* **2**:3 (2008), 249–298. MR 2009f:11079 Zbl 1168.11024

[Zhang 1998] S.-W. Zhang, "Equidistribution of small points on abelian varieties", *Ann. of Math.* (2) **147**:1 (1998), 159–165. MR 99e:14032 Zbl 0991.11034

Received July 10, 2013. Revised May 6, 2014.

SARA CHECCOLI
INSTITUT FOURIER
UNIVERSITÉ JOSEPH FOURIER, GRENOBLE
100 RUE DES MATHS
38402 ST MARTIN D'HÈRES
FRANCE
sara.checcoli@ujf-grenoble.fr

EVELINA VIADA
MATHEMATISCHES INSTITUT
GEORG-AUGUST-UNIVERSITÄT
BUNSENSTRASSE 3-5
D-D-37073 GÖTTINGEN
GERMANY
evelina.viada@unibas.ch

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.347

EIGENVALUE ESTIMATE AND COMPACTNESS FOR CLOSED f-MINIMAL SURFACES

XU CHENG, TITO MEJIA AND DETANG ZHOU

Let Ω be a bounded domain with convex boundary in a complete noncompact Riemannian manifold with Bakry–Émery Ricci curvature bounded below by a positive constant. We prove a lower bound on the first eigenvalue of the weighted Laplacian for closed embedded f-minimal hypersurfaces contained in Ω. Using this estimate, we prove a compactness theorem for the space of closed embedded f-minimal surfaces with uniform upper bounds on genus and diameter in a complete 3-manifold with Bakry–Émery Ricci curvature bounded below by a positive constant and admitting an exhaustion by bounded domains with convex boundary.

1. Introduction

A hypersurface Σ immersed in a Riemannian manifold $(M, \bar{g})$ is said to be f-minimal if its mean curvature H satisfies, for any $p \in \Sigma$,

$$H = \langle \overline{\nabla} f, \nu \rangle,$$

where ν is the unit normal at $p \in \Sigma$, f is a smooth function defined on M, and $\overline{\nabla} f$ denotes the gradient of f on M. When f is a constant function, an f-minimal hypersurface is just a minimal hypersurface. One nontrivial class of f-minimal hypersurfaces is that of self-shrinkers. Recall that a self-shrinker (for the mean curvature flow in the Euclidean space $(\mathbb{R}^{n+1}, g_{\text{can}})$) is a hypersurface immersed in $(\mathbb{R}^{n+1}, g_{\text{can}})$ satisfying

$$H = \tfrac{1}{2}\langle x, \nu \rangle,$$

where x is the position vector in $\mathbb{R}^{n+1}$. Hence a self-shrinker is an f-minimal hypersurface Σ with $f = |x|^2/4$ (see more information on self-shrinkers in [Colding and Minicozzi 2012a] and references therein).

In the study of f-minimal hypersurfaces, it is convenient to consider the ambient space as a smooth metric measure space $(M, \bar{g}, e^{-f} d\mu)$, where $d\mu$ is the volume

Cheng and Zhou were partially supported by CNPq and FAPERJ of Brazil. Mejia was supported by CNPq of Brazil.
MSC2010: primary 58J50; secondary 58E30.
Keywords: Riemannian manifold, eigenvalue, drifted Laplacian, minimal surface.

form of $\bar{g}$. For $(M, \bar{g}, e^{-f}\, d\mu)$, an important and natural tensor is the Bakry–Émery Ricci curvature $\overline{\mathrm{Ric}}_f := \overline{\mathrm{Ric}} + \overline{\nabla}^2 f$. There are many interesting examples of smooth metric measure spaces $(M, \bar{g}, e^{-f}\, d\mu)$ with $\overline{\mathrm{Ric}}_f \geq k$, for a positive constant k. A nontrivial class of examples is the shrinking gradient Ricci solitons. It is known that, after a normalization, a shrinking gradient Ricci soliton $(M, \bar{g}, f)$ satisfies the equation $\overline{\mathrm{Ric}} + \overline{\nabla}^2 f = \bar{g}/2$ or, equivalently, $\overline{\mathrm{Ric}}_f = \frac{1}{2}$. We refer to [Cao 2010], a survey of this topic where some compact and noncompact examples are explained. Even though the asymptotic growth of the potential function f of a noncompact shrinking gradient Ricci soliton is the same as that of a Gaussian shrinking soliton [Cao and Zhou 2010], both the geometry and topology can be quite different from known examples. We may consider f-minimal hypersurfaces in a shrinking gradient Ricci soliton. For instance, a self-shrinker in $\mathbb{R}^{n+1}$ can be viewed as an f-minimal hypersurface in the Gaussian shrinking soliton $(\mathbb{R}^{n+1}, g_{\mathrm{can}}, |x|^2/4)$.

There are other examples of f-minimal hypersurfaces. Let M be the hyperbolic space $\mathbb{H}^{n+1}(-1)$. Let r denote the distance function from a fixed point $p \in M$ and $f(x) = nar^2(x)$, where $a > 0$ is a constant. Then $\overline{\mathrm{Ric}}_f \geq n(2a-1)$, and the geodesic sphere of radius r centered at p in $\mathbb{H}^{n+1}(-1)$ is an f-minimal hypersurface if it satisfies $2ar = \coth r$.

An f-minimal hypersurface Σ has two aspects to view. One is that Σ is f-minimal if and only if Σ is a critical point of the weighted volume functional $e^{-f}\, d\sigma$, where $d\sigma$ is the volume element of Σ. Another one is that Σ is f-minimal if and only if Σ is minimal in the new conformal metric $\tilde{g} = e^{-2f/n}\bar{g}$ (see Section 2). f-minimal hypersurfaces, even more general stationary hypersurfaces for parametric elliptic functionals, have been studied before. See, for instance, the work of White [1987] and Colding and Minicozzi [2002].

In this paper, we will first estimate the lower bound on the first eigenvalue of the weighted Laplacian $\Delta_f = \Delta - \langle \nabla f, \nabla \cdot \rangle$ for closed (i.e., compact and without boundary) embedded f-minimal hypersurfaces in a complete metric measure space $(M, \bar{g}, e^{-f}\, d\mu)$. Subsequently using the eigenvalue estimate, we study compactness for the space of closed embedded f-minimal surfaces in a complete noncompact 3-manifold. To explain our result, we give some background.

Choi and Wang [1983] estimated the lower bound for the first eigenvalue of closed minimal hypersurfaces in a complete Riemannian manifold with Ricci curvature bounded below by a positive constant and proved the following:

Theorem 1. *If M is a simply connected complete Riemannian manifold with Ricci curvature bounded below by a constant $k > 0$ and Σ is a closed embedded minimal hypersurface, then the first eigenvalue of the Laplacian Δ on Σ is at least $k/2$.*

Later, using a covering argument, Choi and Schoen [1985] proved that the assumption that M is simply connected is not needed. Recently Ma and Du [2010] extended Theorem 1 to the first eigenvalue of the weighted Laplacian Δ_f on a closed embedded f-minimal hypersurface in a simply connected compact manifold with positive Bakry–Émery Ricci curvature $\overline{\mathrm{Ric}}_f$. Very recently Li and Wei [2012] also used the covering argument to delete the assumption that the ambient space is simply connected in the result of Ma and Du.

The Bonnet–Myers theorem says that a complete manifold with Ricci curvature bounded below by a positive constant must be compact. But the corresponding result is not true for complete manifolds with Bakry–Émery Ricci curvature $\overline{\mathrm{Ric}}_f$ bounded below by a positive constant. One example is the Gaussian shrinking soliton $(\mathbb{R}^{n+1}, g_{\mathrm{can}}, e^{-|x|^2/4}d\mu)$, with $\overline{\mathrm{Ric}}_f = \frac{1}{2}$. Hence the theorems of Ma and Du and Li and Wei cannot be applied to self-shrinkers.

For self-shrinkers, Ding and Xin [2013] recently obtained a lower bound on the first eigenvalue $\lambda_1(\mathscr{L})$ of the weighted Laplacian $\mathscr{L} = \Delta - \langle x, \nabla\cdot\rangle/2$ (i.e., Δ_f) on a closed n-dimensional embedded self-shrinker in the Euclidean space $\mathbb{R}^{n+1}$, that is, $\lambda_1(\mathscr{L}) \geq \frac{1}{4}$.

We will discuss a lower bound for the first eigenvalue of Δ_f of a closed embedded f-minimal hypersurface in the case that the ambient space is complete and noncompact. Precisely, we prove the following:

Theorem 2. *Let $(M^{n+1}, \bar{g}, e^{-f}d\mu)$ be a complete noncompact smooth metric measure space with Bakry–Émery Ricci curvature $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Let Σ be a closed embedded f-minimal hypersurface in M. If there is a bounded domain D in M with convex boundary ∂D so that Σ is contained in D, then the first eigenvalue $\lambda_1(\Delta_f)$ of the weighted Laplacian Δ_f on Σ satisfies*

$$\lambda_1(\Delta_f) \geq \frac{k}{2}. \tag{1}$$

Here and below the boundary ∂D is called convex if, for any $p \in \partial D$, the second fundamental form A of ∂D at p is nonnegative with respect to the outer unit normal of ∂D.

A closed self-shrinker Σ^n in $\mathbb{R}^{n+1}$ satisfies the assumption of Theorem 2 since there always exists a ball D of $\mathbb{R}^{n+1}$containing Σ. Therefore Theorem 2 implies the result of Ding and Xin for self-shrinkers mentioned before. Also we give a different and hence alternative proof of their result.

Remark. If M is a Cartan–Hadamard manifold, all geodesic balls are convex. If M is a complete noncompact Riemannian manifold with nonnegative sectional curvature, the work of Cheeger and Gromoll [1972] asserts that M admits an exhaustion by convex domains.

Choi and Wang [1983] used the lower bound estimate of the first eigenvalue in Theorem 1 to obtain an upper bound on the area of a simply connected closed embedded minimal surface Σ in a 3-manifold, depending on the genus g of Σ and the positive lower bound k of Ricci curvature of M. Further the lower bound on the first eigenvalue and the upper bound on the area were used in [Choi and Schoen 1985] to prove a smooth compactness theorem for the space of closed embedded minimal surfaces of genus g in a closed 3-manifold M^3 with positive Ricci curvature. Very recently Li and Wei [2012] proved a compactness theorem for closed embedded f-minimal surfaces in a compact 3-manifold with Bakry–Émery Ricci curvature $\overline{\mathrm{Ric}}_f \geq k$, for a constant $k > 0$.

On the other hand, Ding and Xin [2013] recently applied the lower bound estimate of the first eigenvalue of the weighted Laplacian on a self-shrinker to prove a compactness theorem for closed self-shrinkers in $\mathbb{R}^3$ with uniform bounds on genus and diameter. As was mentioned before, a self-shrinker in $\mathbb{R}^3$ is an f-minimal surface in a complete noncompact $\mathbb{R}^3$ with $\overline{\mathrm{Ric}}_f \geq \frac{1}{2}$. Motivated by this example, we consider compactness for f-minimal surfaces in a complete noncompact manifold. We prove:

Theorem 3. *Let $(M^3, \bar{g}, e^{-f}\,d\mu)$ be a complete noncompact smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Assume that M admits an exhaustion by bounded domains with convex boundary. Then the space, denoted by $S_{D,g}$, of closed embedded f-minimal surfaces in M with genus at most g and diameter at most D is compact in the C^m topology, for any $m \geq 2$. Namely any sequence in $S_{D,g}$ has a subsequence that converges in the C^m topology on compact subsets to a surface in $S_{D,g}$, for any $m \geq 2$.*

Theorem 3 implies especially the compactness theorem of Ding and Xin for self-shrinkers. We also prove the following compactness theorem, which implies Theorem 3.

Theorem 4. *Let $(M^3, \bar{g}, e^{-f}\,d\mu)$ be a complete noncompact smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Given a bounded domain $\Omega \subset M$, let S be the space of closed embedded f-minimal surfaces in M with genus at most g and contained in the closure $\overline{\Omega}$. If there is a bounded domain $U \subset M$ with convex boundary so that $\overline{\Omega} \subset U$, then S is compact in the C^m topology, for any $m \geq 2$. Namely any sequence in S has a subsequence that converges in the C^m topology on compact subsets to a surface in S, for any $m \geq 2$.*

If M admits an exhaustion by bounded domains with convex boundary, such U as in Theorem 4 always exists. Also the assumption that f-minimal surfaces are contained in the closure of a bounded domain Ω in Theorem 4 is equivalent to there being a uniform upper bound on the extrinsic diameter of f-minimal surfaces (see remark on page 361).

We mention that, for self-shrinkers in $\mathbb{R}^3$, Colding and Minicozzi [2012b] proved a smooth compactness theorem for complete embedded self-shrinkers with uniform upper bound on genus and uniform scale-invariant area growth. In [Cheng et al. 2012], we generalized their result to the complete embedded f-minimal surfaces in a complete noncompact smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, for a constant $k > 0$.

Theorems 3 and 4 have some immediate corollaries. First they imply the corresponding compactness theorems for embedded closed f-minimal surfaces of fixed topological type and bounded diameter; see Theorems 7 and 8. Second, by using an argument as in [Choi and Schoen 1985], we have the following uniform curvature estimates:

Corollary of Theorem 3. *Let $(M^3, \bar{g}, e^{-f} d\mu)$ be a complete smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Assume that M admits an exhaustion by bounded domains with convex boundary. Then, for any integer g and a positive constant D, there exists a constant C depending only on M, g and D such that if Σ is a closed embedded f-minimal surface of genus g and diameter at most D in M, the norm $|A|$ of the second fundamental form of Σ satisfies*

$$\max_{x\in\Sigma} |A| \leq C.$$

Corollary of Theorem 4. *Let $(M^3, \bar{g}, e^{-f} d\mu)$ be a complete noncompact smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Let Ω be a bounded domain whose closure is contained in a bounded domain U with convex boundary. Then, for any integer g, there exists a constant C depending only on U, g such that if Σ is a closed embedded f-minimal surface of genus g contained in $\overline{\Omega}$, the norm $|A|$ of the second fundamental form of Σ satisfies*

$$\max_{x\in\Sigma} |A| \leq C.$$

An argument similar to the proof of Theorem 2 also works for the case where the ambient space is a compact manifold with convex boundary. Hence we have the following estimate:

Theorem 5. *Let $(M^{n+1}, \bar{g})$ be a simply connected compact manifold with convex boundary ∂M and f a nonconstant smooth function on M. Assume that $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. If Σ is a closed f-minimal hypersurface embedded in M and does not intersect the boundary ∂M, then the first eigenvalue of the weighted Laplacian on Σ satisfies*

$$\lambda_1(\Delta_f) \geq \frac{k}{2}. \tag{2}$$

Here we give a remark: the assumption in Theorem 5 that f is a nonconstant smooth function on M is necessary. The reason is that under the assumption

$\overline{\mathrm{Ric}} \geq k > 0$, any closed minimal hypersurface Σ must intersect the convex boundary ∂M by a standard argument similar to the one in Frankel's intersection theorem.

The rest of this paper is organized as follows: In Section 2, some definitions and notation are given. In Section 3, we give some facts which will be used later. In Section 4, we prove Theorems 2 and 5. In Section 6, we prove Theorems 3 and 4. For completeness, we give in an appendix the proof of the known Reilly formula for a weighted metric measure space.

2. Definitions and notation

In general, a smooth metric measure space, denoted by $(N, g, e^{-w} d\mathrm{vol})$, is a Riemannian manifold (N, g) together with a weighted volume form $e^{-w} d\mathrm{vol}$ on N, where w is a smooth function on N and $d\mathrm{vol}$ the volume element induced by the Riemannian metric g. The associated weighted Laplacian Δ_w is defined by

$$\Delta_w u := \Delta u - \langle \nabla w, \nabla u \rangle,$$

where Δ and ∇ are the Laplacian and gradient on (N, g), respectively.

The second-order operator Δ_w is a self-adjoint operator on the space of square integrable functions on N with respect to the measure $e^{-w} d\mathrm{vol}$. For a closed manifold N, the first eigenvalue of Δ_w, denoted by $\lambda_1(\Delta_w)$, is the lowest nonzero real number λ_1 satisfying

$$\Delta_w u = -\lambda_1 u, \quad \text{on } N.$$

It is well known that the definition of $\lambda_1(\Delta_w)$ is equivalent to

$$\lambda_1(\Delta_w) = \inf_{\substack{\int_N u e^{-w} \, d\mathrm{vol}=0 \\ u \not\equiv 0}} \frac{\int_N |\nabla u|^2 e^{-w} \, d\mathrm{vol}}{\int_N u^2 e^{-w} \, d\mathrm{vol}}.$$

The ∞-Bakry–Émery Ricci curvature tensor Ric_w (for simplicity, Bakry–Émery Ricci curvature) on $(N, g, e^{-w} d\mathrm{vol})$ is defined by

$$\mathrm{Ric}_w := \mathrm{Ric} + \nabla^2 w,$$

where Ric denotes the Ricci curvature of (N, g) and $\nabla^2 w$ is the Hessian of w on N. If w is constant, Δ_w and Ric_w are the Laplacian Δ and Ricci curvature Ric on N, respectively.

Now let $(M^{n+1}, \bar{g})$ be an $(n+1)$-dimensional Riemannian manifold. Assume that f is a smooth function on M so that $(M^{n+1}, \bar{g}, e^{-f} d\mu)$ is a smooth metric measure space, where $d\mu$ is the volume element induced by $\bar{g}$.

Let $i : \Sigma^n \to M^{n+1}$ be an n-dimensional smooth immersion. Then

$$i : (\Sigma^n, i^* \bar{g}) \to (M^{n+1}, \bar{g})$$

is an isometric immersion with the induced metric $i^*\bar{g}$. For simplicity, we still denote $i^*\bar{g}$ by $\bar{g}$ whenever there is no confusion. Let $d\sigma$ denote the volume element of $(\Sigma, \bar{g})$. Then the function f induces a weighted measure $e^{-f}\,d\sigma$ on Σ. Thus we have an induced smooth metric measure space $(\Sigma^n, \bar{g}, e^{-f}\,d\sigma)$.

In this paper, unless otherwise specified, we denote by a bar all quantities on $(M, \bar{g})$, for instance by $\overline{\nabla}$ and $\overline{\mathrm{Ric}}$, the Levi-Civita connection and the Ricci curvature tensor of $(M, \bar{g})$, respectively. Also we denote, for example, by ∇, Ric, Δ and Δ_f, the Levi-Civita connection, the Ricci curvature tensor, the Laplacian, and the weighted Laplacian on $(\Sigma, \bar{g})$, respectively. Let $p \in \Sigma$ and ν a unit normal at p. The second fundamental form A, the mean curvature H, and the mean curvature vector $\boldsymbol{H}$ of hypersurface $(\Sigma, \bar{g})$ are defined, respectively, by

$$A : T_p\Sigma \to T_p\Sigma, \quad A(X) := \overline{\nabla}_X \nu, \quad X \in T_p\Sigma,$$
$$H := \operatorname{tr} A = -\sum_{i=1}^{n} \langle \overline{\nabla}_{e_i} e_i, \nu \rangle, \quad \boldsymbol{H} := -H\nu.$$

Define the weighted mean curvature vector $\boldsymbol{H_f}$ and the weighted mean curvature H_f of $(\Sigma, \bar{g})$ by

$$\boldsymbol{H_f} := \boldsymbol{H} - (\overline{\nabla} f)^{\perp} \quad \text{and} \quad \boldsymbol{H_f} = -H_f \nu,$$

where $\perp$ denotes the projection to the normal bundle of Σ. It follows that

$$H_f = H - \langle \overline{\nabla} f, \nu \rangle.$$

Definition. A hypersurface Σ immersed in $(M^{n+1}, \bar{g}, e^{-f}\,d\mu)$ with the induced metric $\bar{g}$ is called f-minimal if its weighted mean curvature H_f vanishes identically or, equivalently, if it satisfies

$$H = \langle \overline{\nabla} f, \nu \rangle. \tag{3}$$

Definition. The weighted volume of $(\Sigma, \bar{g})$ is defined by

$$V_f(\Sigma) := \int_\Sigma e^{-f}\,d\sigma. \tag{4}$$

It is well known that Σ is f-minimal if and only if Σ is a critical point of the weighted volume functional. Namely it holds that

Proposition 1. *If T is a compactly supported normal variational vector field on Σ (i.e., $T = T^{\perp}$), then the first variation formula of the weighted volume of $(\Sigma, \bar{g})$ is given by*

$$\frac{d}{dt} V_f(\Sigma_t)\Big|_{t=0} = -\int_\Sigma \langle T, \boldsymbol{H_f} \rangle_{\bar{g}} e^{-f}\,d\sigma. \tag{5}$$

On the other hand, an f-minimal hypersurface can be viewed as a minimal hypersurface under a conformal metric. More precisely, define the new metric $\tilde{g} = e^{-2f/n}\bar{g}$ on M, which is conformal to $\bar{g}$. Then the immersion $i : \Sigma \to M$ induces a metric $i^*\tilde{g}$ on Σ from $(M, \tilde{g})$. In the following, $i^*\tilde{g}$ is still denoted by $\tilde{g}$ for simplicity of notation. The volume of $(\Sigma, \tilde{g})$ is

$$\widetilde{V}(\Sigma) := \int_\Sigma d\tilde{\sigma} = \int_\Sigma e^{-f}\, d\sigma = V_f(\Sigma). \tag{6}$$

Hence Proposition 1 and (6) imply that

$$\int_\Sigma \langle T, \widetilde{\boldsymbol{H}}\rangle_{\tilde{g}}\, d\tilde{\sigma} = \int_\Sigma \langle T, \boldsymbol{H}_f\rangle_{\bar{g}} e^{-f}\, d\sigma, \tag{7}$$

where $d\tilde{\sigma} = e^{-f}\, d\sigma$ and $\widetilde{\boldsymbol{H}}$ denote the volume element and the mean curvature vector of Σ with respect to the conformal metric $\tilde{g}$, respectively.

Equation (7) implies that $\widetilde{\boldsymbol{H}} = e^{2f/n}\boldsymbol{H}_f$. Therefore $(\Sigma, \bar{g})$ is f-minimal in $(M, \bar{g})$ if and only if $(\Sigma, \tilde{g})$ is minimal in $(M, \tilde{g})$.

In this paper, for a closed hypersurface, we choose ν to be the outer unit normal.

3. Some facts on the weighted Laplacian and f-minimal hypersurfaces

In this section, we give some known results which will be used later in this paper. Recall that Reilly [1977] proved an integral version of the Bochner formula for compact domains of a Riemannian manifold, which is called the Reilly formula. Ma and Du [2010] obtained a Reilly formula for metric measure spaces, which is the following proposition. We include its proof in an appendix for the sake of completeness.

Proposition 2. *Let Ω be a compact Riemannian manifold with boundary $\partial\Omega$ and $(\Omega, \bar{g}, e^{-f}\, d\mu)$ a smooth metric measure space. Then*

$$\begin{aligned}\int_\Omega (\overline{\Delta}_f u)^2 e^{-f} &= \int_\Omega |\overline{\nabla}^2 u|^2 e^{-f} + \int_\Omega \overline{\mathrm{Ric}}_f(\overline{\nabla} u, \overline{\nabla} u) e^{-f} \\ &\quad + 2\int_{\partial\Omega} u_\nu (\Delta_f u) e^{-f} + \int_{\partial\Omega} A(\nabla u, \nabla u) e^{-f} + \int_{\partial\Omega} u_\nu^2 H_f e^{-f},\end{aligned} \tag{8}$$

where ν is the outward pointing unit normal to $\partial\Omega$ and A is the second fundamental form of $\partial\Omega$ with respect to the normal ν, the quantities with bars denote the ones on $(\Omega, \bar{g})$ (for instance, $\overline{\mathrm{Ric}}_f$ denotes the Bakry–Émery Ricci curvature on $(\Omega, \bar{g})$), and Δ_f and H_f denote the weighted Laplacian on $\partial\Omega$ and the weighted mean curvature of $\partial\Omega$, respectively.

A Riemannian manifold with Bakry–Émery Ricci curvature bounded below by a positive constant has some properties similar to a Riemannian manifold with Ricci curvature bounded below by a positive constant. We refer to [Wei and Wylie 2009; Munteanu and Wang 2014; 2012] and the references therein.

Proposition 3 [Morgan 2005] (see also [Wei and Wylie 2009, Corollary 5.1]). *If a complete smooth metric measure space $(N, g, e^{-\omega} d\mu)$ has $\mathrm{Ric}_w \geq k$, where k is a positive constant, then N has finite weighted volume and finite fundamental group.*

For f-minimal hypersurfaces, the following intersection theorem holds.

Proposition 4 [Wei and Wylie 2009, Theorem 7.4]. *Any two closed f-minimal hypersurfaces immersed in a complete smooth metric measure space $(M, \bar{g}, e^{-f} d\mu)$ with $\overline{\mathrm{Ric}}_f > 0$ must intersect. Thus a closed f-minimal hypersurface in M must be connected.*

In [Cheng and Zhou 2013] it was proved that the weighted volume of a self-shrinker Σ^n immersed in $\mathbb{R}^m$ being finite implies it is properly immersed. This result extends to f-minimal submanifolds:

Proposition 5 [Cheng et al. 2012]. *Let Σ^n be an n-dimensional complete f-minimal submanifold immersed in an m-dimensional Riemannian manifold M^m, $n < m$. If Σ has finite weighted volume, then Σ is properly immersed in M.*

An f-minimal hypersurface is an f-minimal submanifold with codimension 1. See more properties of f-minimal submanifolds in [Cheng et al. 2012].

4. Lower bound for $\lambda_1(\Delta_f)$

In this section, we apply the Reilly formula for metric measure spaces to prove Theorems 2 and 5.

Proof of Theorem 2. Since $\overline{\mathrm{Ric}}_f \geq k$, where $k > 0$ is constant, Proposition 3 implies that M has finite fundamental group. We first assume that M is simply connected. Since Σ is connected (Proposition 4) and embedded in M, Σ is orientable and divides M into two components (see its proof in [Choi and Schoen 1985]). Thus Σ divides D into two bounded components Ω_1 and Ω_2. That is $D\backslash\Sigma = \Omega_1 \cup \Omega_2$ with $\partial\Omega_1 = \Sigma$ and $\partial\Omega_2 = \partial D \cup \Sigma$.

For simplicity, we denote by λ_1 the first eigenvalue $\lambda_1(\Delta_f)$ of the weighted Laplacian Δ_f on Σ. Let h be a corresponding eigenfunction so that on Σ

$$\Delta_f h + \lambda_1 h = 0 \quad \text{with} \int_\Sigma h^2 e^{-f} = 1.$$

Consider the solution of the Dirichlet problem on Ω_1 so that

$$\begin{cases} \bar{\Delta}_f u = 0 & \text{in } \Omega_1, \\ u = h & \text{on } \partial\Omega_1 = \Sigma. \end{cases} \tag{9}$$

Substitute Ω_1 for Ω and put the solution u of (9) in Proposition 2. Then the assumption on $\overline{\mathrm{Ric}}_f$ implies that

$$0 \geq k \int_{\Omega_1} |\overline{\nabla} u|^2 e^{-f} - 2\lambda_1 \int_\Sigma u_\nu h e^{-f} + \int_\Sigma A(\nabla h, \nabla h) e^{-f},$$

where ν is the outer unit normal of Σ with respect to Ω_1. By Stokes' theorem and (9),

$$\int_\Sigma u_\nu h e^{-f} = \int_{\Omega_1} (|\overline{\nabla} u|^2 + u\overline{\Delta}_f u) e^{-f} = \int_{\Omega_1} |\overline{\nabla} u|^2 e^{-f}.$$

Thus

$$0 \geq (k - 2\lambda_1) \int_{\Omega_1} |\overline{\nabla} u|^2 e^{-f} + \int_\Sigma A(\nabla h, \nabla h) e^{-f}.$$

If $\int_\Sigma A(\nabla h, \nabla h) e^{-f} \geq 0$, by $u \not\equiv C$, we have

$$\lambda_1 \geq \frac{k}{2}.$$

If $\int_\Sigma A(\nabla h, \nabla h) e^{-f} < 0$, we consider the compact domain Ω_2 with the boundary $\partial\Omega_2 = \Sigma \cup \partial D$. Let u be the solution of the mixed problem

$$\begin{cases} \overline{\Delta}_f u = 0 & \text{in } \Omega_2, \\ u = h & \text{on } \Sigma, \\ u_{\tilde{\nu}} = 0 & \text{on } \partial D, \end{cases} \tag{10}$$

where $\tilde{\nu}$ denotes the outer unit normal of ∂D with respect to Ω_2.

Substituting Ω_2 for Ω and putting the solution u of (10) in Proposition 2, we have

$$0 \geq \int_{\Omega_2} |\overline{\nabla}^2 u|^2 e^{-f} + k \int_{\Omega_2} |\overline{\nabla} u|^2 e^{-f} - 2\lambda_1 \int_\Sigma h u_{\tilde{\nu}} e^{-f} \\ + \int_\Sigma \widetilde{A}(\nabla h, \nabla h) e^{-f} + \int_{\partial D} \widetilde{A}(\nabla u, \nabla u) e^{-f},$$

where $\tilde{\nu}$ denotes the outer unit normal of Σ with respect to Ω_2 and $\widetilde{A}$ denotes the second fundamental form of Σ with respect to the normal $\tilde{\nu}$.

On the other hand, Stokes' theorem and (10) imply

$$\int_{\Omega_2} |\overline{\nabla} u|^2 e^{-f} = \int_{\partial\Omega_2} u u_{\tilde{\nu}} e^{-f} = \int_\Sigma h u_{\tilde{\nu}} e^{-f}.$$

Thus we have

$$0 \geq (k - 2\lambda_1) \int_{\Omega_2} |\overline{\nabla} u|^2 e^{-f} + \int_\Sigma \widetilde{A}(\nabla h, \nabla h) e^{-f} + \int_{\partial D} \widetilde{A}(\nabla u, \nabla u) e^{-f}. \tag{11}$$

Since ∂D is assumed convex, the last term on the right side of (11) is nonnegative. Observe that the orientations of Σ are opposite for Ω_1 and Ω_2. Namely $\tilde{\nu} = -\nu$.

Then $\widetilde{A}(\nabla u, \nabla u) = -A(\nabla u, \nabla u)$ on Σ. This implies that the second term on the right side of (11) is nonnegative. Thus

$$0 \geq (k - 2\lambda_1) \int_{\Omega_2} |\overline{\nabla} u|^2 e^{-f}.$$

Since u is not a constant function, we conclude that $k - 2\lambda_1 \leq 0$. Again we have

$$\lambda_1 \geq \frac{k}{2}.$$

Therefore we obtain that $\lambda_1(\Delta_f) \geq k/2$ if M is simply connected.

Second, if M is not simply connected, we consider its universal covering $\widehat{M}$, which is a finite $|\pi_1|$-fold covering. $\widehat{M}$ is simply connected and the covering map $\pi : \widehat{M} \to M$ is a locally isometry.

Take $\hat{f} = f \circ \pi$. Obviously $\widehat{M}$ has $\widehat{\mathrm{Ric}}_{\hat{f}} \geq k$, and the lift $\widehat{\Sigma}$ of Σ is also $\hat{f}$-minimal, embedded and closed. By Proposition 4, $\widehat{\Sigma}$ must be connected. Since $\widehat{M}$ is simply connected, the closed embedded connected $\widehat{\Sigma}$ must be orientable and thus divides $\widehat{M}$ into two components. Moreover the connectedness of $\widehat{\Sigma}$ implies that the lift $\widehat{D}$ of D is also a connected domain. Also $\partial \widehat{D} = \widehat{\partial D}$ is smooth and convex. Hence the assertion obtained for the simply connected ambient space can be applied here. Thus the first eigenvalue of the weighted Laplacian $\widehat{\Delta}_{\hat{f}}$ on $\widehat{\Sigma}$ satisfies $\lambda_1(\widehat{\Delta}_{\hat{f}}) \geq k/2$.

Observing the lift of the first eigenfunction of Σ is also an eigenfunction of $\widehat{M}$, we have

$$\lambda_1(\Delta_f) \geq \lambda_1(\widehat{\Delta}_{\hat{f}}) \geq \frac{k}{2}. \qquad \square$$

Remark. In Theorem 2, the boundary ∂D is not necessarily smooth. ∂D can be assumed to be C^1, which is sufficient for the existence of the solution of the mixed problem (10).

Theorem 5 holds by the same argument as that of Theorem 2.

5. Upper bound on area and total curvature of f-minimal surfaces

In this section, we study surfaces in a 3-manifold. First we estimate the corresponding upper bounds on the area and weighted area of an embedded closed f-minimal surface by applying the first eigenvalue estimate in Section 4. Next we discuss the upper bound on the total curvature. We begin with a result of Yang and Yau [1980]:

Proposition 6. *Let Σ^2 be a closed orientable Riemannian surface with genus g. Then the first eigenvalue $\lambda_1(\Delta)$ of the Laplacian Δ on Σ satisfies*

$$\lambda_1(\Delta)\,\mathrm{Area}(\Sigma) \leq 8\pi(1+g).$$

Using Theorem 2 and Proposition 6, we obtain the following area estimates for closed embedded f-minimal surfaces if the ambient space is simply connected.

Proposition 7. *Let $(M^3, \bar{g}, e^{-f}\,d\mu)$ be a simply connected complete smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Let $\Sigma^2 \subset M$ be a closed embedded f-minimal surface with genus g. If Σ is contained in a bounded domain D with convex boundary ∂D, then its area and weighted area satisfy*

$$\mathrm{Area}(\Sigma) \leq \frac{16\pi(1+g)}{k} e^{\mathrm{osc}_\Sigma f} \tag{12}$$

and

$$\mathrm{Area}_f(\Sigma) \leq \frac{16\pi(1+g)}{k} e^{-\inf_\Sigma f}, \tag{13}$$

where $\mathrm{osc}_\Sigma f = \sup_\Sigma f - \inf_\Sigma f$.

Proof. Consider the conformal metric $\tilde{g} = e^{-f}\bar{g}$ on M. Let $\lambda_1(\widetilde{\Delta})$ be the first eigenvalue of the Laplacian $\widetilde{\Delta}$ on $(\Sigma, \tilde{g})$, which satisfies

$$\lambda_1(\widetilde{\Delta}) = \inf_{\substack{\int_\Sigma u\,d\tilde{\sigma}=0\\ u\neq 0}} \frac{\int_\Sigma |\widetilde{\nabla} u|^2_{\tilde{g}}\,d\tilde{\sigma}}{\int_\Sigma u^2\,d\tilde{\sigma}},$$

where $\widetilde{\Delta}$, $\widetilde{\nabla}$ and $d\tilde{\sigma}$ are the Laplacian, gradient and area element of Σ with respect to the metric $\tilde{g}$, respectively.

On the other hand, the first eigenvalue of the weighted Laplacian $\lambda_1(\Delta_f)$ on $(\Sigma, \bar{g})$ satisfies

$$\lambda_1(\Delta_f) = \inf_{\substack{\int_\Sigma u e^{-f}\,d\sigma=0\\ u\neq 0}} \frac{\int_\Sigma |\nabla u|^2_{\bar{g}} e^{-f}\,d\sigma}{\int_\Sigma u^2 e^{-f}\,d\sigma}.$$

Since $\widetilde{\nabla} u = e^f \nabla u$, $d\tilde{\sigma} = e^{-f}\,d\sigma$ and $\tilde{g} = e^{-f}\bar{g}$,

$$\begin{aligned}
\lambda_1(\widetilde{\Delta}) &= \inf_{\substack{\int_\Sigma u e^{-f}\,d\sigma=0\\ u\neq 0}} \frac{\int_\Sigma |\nabla u|^2_{\bar{g}}\,d\sigma}{\int_\Sigma u^2 e^{-f}\,d\sigma} \\
&\geq \inf_{\substack{\int_\Sigma u e^{-f}\,d\sigma=0\\ u\neq 0}} \frac{\int_\Sigma |\nabla u|^2_{\bar{g}} e^{-f+\inf_\Sigma(f)}\,d\sigma}{\int_\Sigma u^2 e^{-f}\,d\sigma} \\
&= e^{\inf_\Sigma f}\lambda_1(\Delta_f).
\end{aligned}$$

By this inequality, Theorem 2 and Proposition 6, we have the estimate

$$\mathrm{Area}(\Sigma, \tilde{g}) \leq \frac{16\pi(1+g)}{k} e^{-\inf_\Sigma(f)}. \tag{14}$$

Since $\mathrm{Area}_f(\Sigma) = \int_\Sigma e^{-f}\, d\sigma = \mathrm{Area}(\Sigma, \tilde{g})$,

$$\mathrm{Area}_f(\Sigma) \le \frac{16\pi(1+g)}{k} e^{-\inf_\Sigma(f)},$$

which is (13). Thus

$$\mathrm{Area}(\Sigma) \le \frac{16\pi(1+g)}{k} e^{\sup_\Sigma(f)-\inf_\Sigma(f)} = \frac{16\pi(1+g)}{k} e^{\mathrm{osc}_\Sigma(f)}.$$

That is, (12) holds. □

Now, suppose that M is not simply connected. We use a covering argument as in [Choi and Schoen 1985].

Proposition 8. *Let $(M^3, \bar{g}, e^{-f} d\mu)$ be a complete smooth metric measure space with $\overline{\mathrm{Ric}}_f \ge k$, where k is a positive constant. Let Σ^2 be a closed embedded f-minimal surface. If Σ is contained in a bounded domain D of M with convex boundary ∂D, then*

$$\mathrm{Area}_f(\Sigma) \le \frac{16\pi}{k}\left(\frac{2}{|\pi_1|} - \frac{1}{2}\chi(\Sigma)\right) e^{-\inf_\Sigma f} \tag{15}$$

and

$$\mathrm{Area}(\Sigma) \le \frac{16\pi}{k}\left(\frac{2}{|\pi_1|} - \frac{1}{2}\chi(\Sigma)\right) e^{\mathrm{osc}_\Sigma f}, \tag{16}$$

where $|\pi_1|$ is the order of the fundamental group of M, and $\chi(\Sigma)$ is the Euler characteristic of Σ.

Proof. Let $\widehat{M}$ be the universal covering manifold of M. By Proposition 3, the covering is a finite $|\pi_1|$-fold covering. Let $\widehat{\Sigma}$ be the lifting of Σ. In the proof of Theorem 2, we have shown that $\widehat{\Sigma}$ is orientable and satisfies the assumption of Theorem 2. Hence Theorem 2 implies that the first eigenvalue of the weighted Laplacian of $\widehat{\Sigma}$ satisfies $\lambda_1(\widehat{\Delta}_{\hat{f}}) \ge k/2$, where $\hat{f}$ is the lift of f. By Proposition 7, we conclude that

$$\mathrm{Area}(\widehat{\Sigma}) \le \frac{16\pi}{k}\big(2 - \tfrac{1}{2}\chi(\widehat{\Sigma})\big) e^{\mathrm{osc}_{\widehat{\Sigma}}(\tilde{f})}$$

and

$$\mathrm{Area}_{\hat{f}}(\widehat{\Sigma}) = \int_{\widehat{\Sigma}} e^{-\hat{f}}\, d\sigma \le \frac{16\pi}{k}\big(2 - \tfrac{1}{2}\chi(\widehat{\Sigma})\big) e^{-\inf_{\widehat{\Sigma}}(\hat{f})}.$$

Thus (15) and (16) follow from the equalities

$$\chi(\widehat{\Sigma}) = |\pi_1| \cdot \chi(\Sigma), \quad \inf_{\widehat{\Sigma}}(\hat{f}) = \inf_\Sigma(f), \quad \mathrm{osc}_{\widehat{\Sigma}}(\hat{f}) = \mathrm{osc}_\Sigma(f),$$
$$\mathrm{Area}(\widehat{\Sigma}) = |\pi_1| \cdot \mathrm{Area}(\Sigma) \quad \text{and} \quad \mathrm{Area}_{\hat{f}}(\widehat{\Sigma}) = |\pi_1| \cdot \mathrm{Area}_f(\Sigma).$$ □

In the following, we will give the upper bound for the total curvature of f-minimal surfaces. Here the term the *total curvature* of Σ means $\int_\Sigma |A|^2\, d\sigma$ not $\int_\Sigma K\, d\sigma$.

Proposition 9. *Let $(M^3, \bar{g}, e^{-f}\,d\mu)$ be a complete smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Let $\Sigma^2 \subset M$ be a closed embedded f-minimal surface with genus g. If Σ is contained in a bounded domain D of M with convex boundary ∂D, then Σ satisfies*

$$\int_\Sigma |A|^2\,d\sigma \leq C, \tag{17}$$

where A is the second fundamental form of $(\Sigma, \bar{g})$ and C is a constant depending on the genus g of Σ, the order $|\pi_1|$ of the fundamental group of M, the maximum $\sup_\Sigma \bar{K}$ of the sectional curvature of M on Σ, the lower bound k of the Bakry–Émery Ricci curvature of M, the oscillation $\operatorname{osc}_\Sigma(f)$ and the maximum $\sup_\Sigma |\bar{\nabla} f|$ on Σ.

Proof. By the Gauss equation and Gauss–Bonnet formula,

$$\begin{aligned}\int_\Sigma |A|^2\,d\sigma = \int_\Sigma H^2 - 2\int_\Sigma (K - \bar{K}) &= \int_\Sigma \langle \bar{\nabla} f, \boldsymbol{n}\rangle^2 - 4\pi\chi(\Sigma) + 2\int_\Sigma \bar{K} \\ &\leq (\sup_\Sigma |\bar{\nabla} f|)^2 \operatorname{Area}(\Sigma) + 8\pi(g-1) + 2(\sup_\Sigma \bar{K}) \operatorname{Area}(\Sigma).\end{aligned}$$

Using (16), we have the conclusion of the theorem. □

To prove the compactness theorem in Section 6, we need the following total curvature estimate for $(\Sigma, \tilde{g})$, which is a minimal surface in $(M, \tilde{g})$.

Proposition 10. *Let $(M^3, \bar{g}, e^{-f}\,d\mu)$ be a complete smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Let $\Sigma^2 \subset M$ be a closed embedded f-minimal surface with genus g. If Σ is contained in a bounded domain D of M with convex boundary $\partial\Omega$, then Σ satisfies*

$$\int_\Sigma |\widetilde{A}|^2_{\tilde{g}}\,d\tilde{\sigma} \leq C, \tag{18}$$

where $\widetilde{A}$ is the second fundamental form of $(\Sigma, \tilde{g})$ with respect to the conformal metric $\tilde{g} = e^{-f}\bar{g}$ of M and C is a constant depending on the genus g of Σ, the order $|\pi_1|$ of the fundamental group of M, the maximum $\sup_\Sigma \widetilde{K}$ of the sectional curvature of $(M, \tilde{g})$ on Σ, the lower bound k of the Bakry–Émery Ricci curvature of M and the oscillation $\operatorname{osc}_\Sigma(f)$ on Σ.

Proof. By the Gauss equation and the Gauss–Bonnet formula, we have

$$\begin{aligned}\int_\Sigma |\widetilde{A}|^2_{\tilde{g}}\,d\tilde{\sigma} = \int_\Sigma \widetilde{H}^2 - 2\int_\Sigma (\widetilde{K}^\Sigma - \widetilde{K}^M)\,d\tilde{\sigma} &= -4\pi\chi(\Sigma) + 2\int_\Sigma \widetilde{K}\,d\tilde{\sigma} \\ &\leq 8\pi(g-1) + 2(\sup_\Sigma \widetilde{K}) \operatorname{Area}((\Sigma, \tilde{g})) \\ &= 8\pi(g-1) + 2(\sup_\Sigma \widetilde{K}) \operatorname{Area}_f(\Sigma).\end{aligned}$$

We have used $\widetilde{H} = e^f H_f = 0$ and $\mathrm{Area}((\Sigma, \tilde{g})) = \mathrm{Area}_f(\Sigma)$. Now (18) follows from (15). □

6. Compactness of compact f-minimal surfaces

We will prove some compactness theorems for closed embedded f-minimal surfaces in a 3-manifold. We have two ways to prove Theorem 4.

The first proof roughly follows the one in [Colding and Minicozzi 2011] (cf. [Choi and Schoen 1985]) with some modifications. The modifications can be made because we have the assumptions that f-minimal surfaces are contained in the closure of a bounded domain Ω of M and $\overline{\Omega}$ is contained in a bounded domain U with convex boundary. The second proof will need a compactness theorem of complete embedded f-minimal surfaces that was proved in [Cheng et al. 2012].

We prefer to give two proofs here since the first one is independent of the compactness theorem of complete embedded f-minimal surfaces. But the compactness theorem of complete embedded f-minimal surfaces needs a theorem about nonexistence of L_f-stable minimal surfaces (see [Cheng et al. 2012, Theorem 3]).

First proof. We first prove a singular compactness theorem, which is a variation of a result from [Choi and Schoen 1985] (compare [Colding and Minicozzi 2011, Proposition 7.14; Anderson 1985; White 1987]):

Proposition 11. *Let $(M^3, \bar{g})$ be a 3-manifold. Assume that Ω is a bounded domain in M. Let Σ_i be a sequence of closed embedded minimal surfaces contained in $\overline{\Omega}$, with genus g, and satisfying*

$$\mathrm{Area}(\Sigma_i) \le C_1 \tag{19}$$

and

$$\int_{\Sigma_i} |A_{\Sigma_i}|^2 \le C_2. \tag{20}$$

Then there exists a finite set of points $\mathcal{S} \subset \overline{\Omega}$ and a subsequence, still denoted by Σ_i, that converges uniformly in the C^m topology ($m \ge 2$) on compact subsets of $M \setminus \mathcal{S}$ to a complete minimal surface $\Sigma \subset \overline{\Omega}$ (possibly with multiplicity).

The subsequence also converges to Σ in extrinsic Hausdorff distance. Σ is smooth, embedded in M, has genus at most g and satisfies (19) and (20).

Proof. We may use the same argument as that of [Colding and Minicozzi 2011, Proposition 7.14]. Moreover $\Sigma_i \subset \overline{\Omega}$ implies that the singular set $S \subset \overline{\Omega}$ and the smooth surface $\Sigma \subset \overline{\Omega}$. Here we omit the details of proof. □

We can apply Proposition 11 to the f-minimal surfaces which are minimal in the conformal metric.

Lemma. *Let $(M^3, \bar{g}, e^{-f} d\mu)$ be a smooth metric measure space. Assume that Ω is a bounded domain in M. Let $\Sigma_i \subset \overline{\Omega}$ be a sequence of closed embedded f-minimal surfaces of genus g. Suppose that $\tilde{g} = e^{-f}\bar{g}$ on M and $(\Sigma_i, \tilde{g})$ satisfy*

$$\mathrm{Area}((\Sigma_i, \tilde{g})) = \mathrm{Area}_f(\Sigma_i) \le C_1 \tag{21}$$

and

$$\int_{\Sigma_i} |\widetilde{A}_{\Sigma_i}|^2_{\tilde{g}} \, d\tilde{\sigma} \le C_2, \tag{22}$$

where $\widetilde{A}_{\Sigma_i}$ and $d\tilde{\sigma}$ denote the second fundamental form and the volume element of $(\Sigma_i, \tilde{g})$, respectively. Then there exists a finite set of points $\mathcal{S} \subset \overline{\Omega}$ and a subsequence, still denoted by Σ_i, that converges uniformly in the C^m topology ($m \ge 2$) on compact subsets of $M \backslash \mathcal{S}$ to a complete f-minimal surface $\Sigma \subset \overline{\Omega}$ (possibly with multiplicity).

The subsequence also converges to Σ in extrinsic Hausdorff distance. Σ is smooth, embedded in M, has genus at most g, and satisfies (21) and (22).

Proof. Since an f-minimal surface in the original metric $\bar{g}$ is equivalent to it being minimal in the conformal metric $\tilde{g}$, we can apply Proposition 11 to get the conclusion of the lemma. □

Proof of Theorem 4. First assume M is simply connected. Since $\Sigma_i \subset \overline{\Omega} \subset U$, we see from Proposition 7 and Proposition 10 that

$$\mathrm{Area}((\Sigma_i, \tilde{g})) = \mathrm{Area}_f(\Sigma_i) \le C_1$$

and

$$\int_{\Sigma_i} |\widetilde{A}_{\Sigma_i}|^2_{\tilde{g}} \, d\sigma_{\tilde{g}} \le C_2,$$

where C_1 and C_2 depend on g, $\sup_{\Omega_j} f$, $\sup_{\Omega_j} \widetilde{K}$ and k.

By the lemma, there exists a finite set of points $\mathcal{S} \subset \tilde{\Omega}$ and a subsequence $\Sigma_{i'}$ that converges uniformly in the C^m topology (any $m \ge 2$) on compact subsets of $M \backslash \mathcal{S}$ to a complete f-minimal surface $\Sigma \subset \overline{\Omega}$ without boundary (possibly with multiplicity). Σ is smooth, embedded in M and has genus at most g. Equivalently, with respect to the conformal metric $\tilde{g}$, a subsequence $\Sigma_{i'}$ of minimal surfaces converges uniformly in the C^m topology on compact subsets of $M \backslash \mathcal{S}$ to a complete minimal surface Σ, where $\Sigma \subset \overline{\Omega}$.

Since complete embedded $\Sigma \subset \overline{\Omega}$ satisfies (21), it must be properly embedded (Proposition 5), thus closed and orientable.

We need to prove that the convergence is smooth across the points $\mathcal{S}$. By Allard's regularity theorem, it suffices to prove that the convergence has multiplicity one. If the multiplicity is not one, by a proof similar to that of [Choi and Schoen 1985] (see also [Colding and Minicozzi 2011, p. 249]), we can show that there is an i big enough and a Σ_i in the convergent subsequence, so that the first eigenvalue of the Laplacian $\widetilde{\Delta}^{\Sigma_i}$ on Σ_i with the conformal metric $\tilde{g}$ satisfies $\lambda_1(\widetilde{\Delta}^{\Sigma_i}) < ke^{\inf_\Omega f}/2$. We have

$$
\begin{aligned}
\lambda_1(\widetilde{\Delta}^{\Sigma_i}) &= \inf\left\{\frac{\int_{\Sigma_i} |\widetilde{\nabla}\phi|^2_{\tilde{g}}\, d\tilde{\sigma}}{\int_{\Sigma_i} \phi^2\, d\tilde{\sigma}}, \int_{\Sigma_i} \phi\, d\tilde{\sigma} = 0\right\} \\
&= \inf\left\{\frac{\int_{\Sigma_i} |\nabla\phi|^2 d\sigma}{\int_{\Sigma_i} \phi^2 e^{-f}\, d\sigma}, \int_{\Sigma_i} \phi e^{-f}\, d\sigma = 0\right\} \\
&\geq \lambda_1(\Delta_f^{\Sigma_i}) e^{\inf_\Omega f}.
\end{aligned}
$$

By Theorem 2, $\Sigma_i \subset \overline{\Omega} \subset U$ implies $\lambda_1(\Delta_f^{\Sigma_i}) \geq k/2$. Thus we have a contradiction.

When M is not simply connected, we use a covering argument. The assumption of $\overline{\mathrm{Ric}}_f \geq k$, where $k > 0$ is constant, implies that M has finite fundamental group π_1 (Proposition 3). We consider the finite-fold universal covering $\widehat{M}$. By the proof of Theorem 2, we know that the corresponding lifts of Σ_i, $\overline{\Omega}$ and U satisfy $\widehat{\Sigma}_i \subset \widehat{\overline{\Omega}} \subset \widehat{U}$. Then Propositions 8 and 10 give uniform bounds on the area and total curvature in the conformal metric $\hat{\tilde{g}}$ on $\widehat{M}$. By the assertion on the simply connected ambient manifold before, we have the smooth convergence of a subsequence of $\widehat{\Sigma}_i$. This implies the smooth convergence of a subsequence of Σ_i. □

Second Proof. In [Cheng et al. 2012], we proved the following:

Theorem 6. *Let $(M^3, \bar{g}, e^{-f} d\mu)$ be a complete smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Given an integer $g \geq 0$ and a constant $V > 0$, the space $S_{g,V}$ of smooth complete embedded f-minimal surfaces $\Sigma \subset M$ with*

- *genus at most g,*
- *$\partial\Sigma = \varnothing$, and*
- *$\int_\Sigma e^{-f}\, d\sigma \leq V$*

is compact in the C^m topology, for any $m \geq 2$. Namely any sequence of $S_{g,V}$ has a subsequence that converges in the C^m topology on compact subsets to a surface in $S_{D,g}$, for any $m \geq 2$.

Proof of Theorem 4. Since a surface in S is contained in $\bar{\Omega} \subset U$, by Proposition 8, we have the uniform bound V of the weighted volume of closed embedded f-minimal surfaces in S. Hence Theorem 6 can be applied. Moreover $\Sigma_i \subset \overline{\Omega}$ implies that the smooth limit surface $\Sigma \subset \overline{\Omega}$. Otherwise, since the subsequence $\{\Sigma_i\}$ converges uniformly in the C^m topology ($m \geq 2$) on any compact subset of M to Σ, there is a surface Σ_i (with index i big enough) in the subsequence that would not satisfy $\Sigma_i \subset \overline{\Omega}$.

By Proposition 5, Σ must be properly embedded. Thus Σ must be closed. □

To prove Theorem 3 we require a lemma.

Lemma. *Let $(M^3, \bar{g}, e^{-f}\,d\mu)$ be a complete noncompact smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k > 0$. If Σ is any closed f-minimal surface in M with genus at most g and diameter at most D, then $\Sigma \subset \overline{B}_r(p)$ for some $r > 0$ (independent of Σ), where $B_r(p)$ is a ball in M with radius r centered at $p \in M$.*

Proof. Fix a closed f-minimal surface Σ_0. Obviously $\Sigma_0 \subset B_{r_0}(p)$ for some $r_0 > 0$. Proposition 4 says that Σ and Σ_0 must intersect. Then, for $x \in \Sigma$,

$$d(p, x) \leq d(p, x_0) + d(x_0, x) \leq r_0 + D, x_0 \in \Sigma_0.$$

Taking $r = r_0 + D$, we have $\Sigma \subset \overline{B}_{r_0+D}$. □

Remark. In the lemma and hence in Theorem 3, D is a bound on the intrinsic diameter of closed f-minimal surfaces or a bound on the extrinsic diameter of closed f-minimal surfaces. Also, by Proposition 4, the assumption that f-minimal surfaces are contained in the closure of a bounded domain Ω in Theorem 4 is equivalent to that of a uniform upper bound on the extrinsic diameter of f-minimal surfaces.

Proof of Theorem 3. By the lemma immediately above, we may apply Theorem 4 to the space $S_{D,g}$. Next the closed embedded limit Σ must have diameter at most D. Otherwise, since the subsequence $\{\Sigma_i\}$ converges uniformly in the C^m topology ($m \geq 2$) on any compact subset of M to Σ, there is a surface Σ_i (with the index i big enough) in the subsequence that would have diameter greater than D. So Σ must be in $S_{D,g}$. □

As a corollary, Theorem 3 implies:

Theorem 7. *Let $(M^3, \bar{g}, e^{-f}\,d\mu)$ be a complete noncompact smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Assume that M admits an exhaustion by bounded domains with convex boundary. Then the space of closed embedded f-minimal surface in M of fixed topological type and diameter at most D is compact in the C^m topology, for any $m \geq 2$.*

Proof of Theorem 7. By Theorem 3, it suffices to prove that the limit f-minimal surface of a convergent subsequence in the given space has the same topological type, which holds by the Gauss–Bonnet formula and smooth convergence. □

Similar to the proof of Theorem 7, Theorem 4 implies:

Theorem 8. *Let $(M^3, \bar{g}, e^{-f}d\mu)$ be a complete noncompact smooth metric measure space with $\overline{\mathrm{Ric}}_f \geq k$, where k is a positive constant. Assume that Ω is a bounded domain and U is a bounded domain with convex boundary so that $\overline{\Omega} \subset U$. Then the space of closed embedded f-minimal surface in M of fixed topological type and contained in the closure $\overline{\Omega}$ is compact in the C^m topology, for any $m \geq 2$.*

Appendix: Proof of Proposition 2

The Bochner formula implies that

$$\tfrac{1}{2}\bar{\Delta}_f|\overline{\nabla}u|^2 - \langle\overline{\nabla}u, \overline{\nabla}(\bar{\Delta}_f u)\rangle = |\overline{\nabla}^2 u|^2 + \overline{\mathrm{Ric}}_f(\overline{\nabla}u, \overline{\nabla}u).$$

Integrating this equation on Ω with respect to the weighted measure $e^{-f}d\mu$, we obtain

$$\int_\Omega \bigl(\tfrac{1}{2}\bar{\Delta}_f|\overline{\nabla}u|^2 - \langle\overline{\nabla}u, \overline{\nabla}(\bar{\Delta}_f u)\rangle\bigr)e^{-f} = \int_\Omega |\overline{\nabla}^2 u|^2 e^{-f} + \int_\Omega \overline{\mathrm{Ric}}_f(\overline{\nabla}u, \overline{\nabla}u)e^{-f}.$$

On the other hand, by the divergence formula, we have

$$\begin{aligned}\tfrac{1}{2}\bar{\Delta}_f|\overline{\nabla}u|^2 &- \langle\overline{\nabla}u, \overline{\nabla}(\bar{\Delta}_f u)\rangle \\ &= \tfrac{1}{2}\overline{\mathrm{div}}\bigl(e^{-f}\overline{\nabla}|\overline{\nabla}u|^2\bigr)e^f - \overline{\mathrm{div}}\bigl(e^{-f}\bar{\Delta}_f(u)\overline{\nabla}u\bigr)e^f + (\bar{\Delta}_f u)^2.\end{aligned}$$

Integrating and applying Stokes' theorem, we have

$$\begin{aligned}\int_\Omega &\bigl(\tfrac{1}{2}\bar{\Delta}_f|\overline{\nabla}u|^2 - \langle\overline{\nabla}u, \overline{\nabla}(\bar{\Delta}_f u)\rangle\bigr)e^{-f} \\ &= \int_{\partial\Omega}\bigl(\tfrac{1}{2}|\overline{\nabla}u|^2_\nu - (\bar{\Delta}_f u)u_\nu\bigr)e^{-f} + \int_\Omega (\bar{\Delta}_f u)^2 e^{-f}.\end{aligned} \tag{23}$$

Then

$$\begin{aligned}\tfrac{1}{2}&|\overline{\nabla}u|^2_\nu - (\bar{\Delta}_f u)u_\nu \\ &= \langle\overline{\nabla}_\nu\overline{\nabla}u, \overline{\nabla}u\rangle - (\bar{\Delta}_f u)u_\nu = \langle\overline{\nabla}_{\overline{\nabla}u}\overline{\nabla}u, \nu\rangle - (\bar{\Delta}_f u)u_\nu \\ &= \langle\overline{\nabla}_\nu\overline{\nabla}u, \nu\rangle u_\nu + \langle\overline{\nabla}_{\nabla u}\overline{\nabla}u, \nu\rangle - (\bar{\Delta}_f u)u_\nu \\ &= \bigl(\langle\overline{\nabla}_\nu\overline{\nabla}u, \nu\rangle - \bar{\Delta}u + \langle\overline{\nabla}f, \overline{\nabla}u\rangle\bigr)u_\nu + \langle\nabla u, \nabla u_\nu\rangle - \langle\overline{\nabla}u, \overline{\nabla}_{\nabla u}\nu\rangle \\ &= \bigl(-\Delta u - Hu_\nu + \langle\nabla f, \nabla u\rangle + \langle\overline{\nabla}f, \nu\rangle u_\nu\bigr)u_\nu + \langle\nabla u, \nabla u_\nu\rangle - \langle\nabla u, \overline{\nabla}_{\nabla u}\nu\rangle \\ &= -(\Delta_f u + H_f u_\nu)u_\nu + \langle\nabla u, \nabla u_\nu\rangle - A(\nabla u, \nabla u),\end{aligned} \tag{24}$$

where $H_f = H - \langle \overline{\nabla} f, \nu \rangle$. By substituting (24) into (23), we obtain

$$\begin{aligned}\int_\Omega \big(\tfrac{1}{2}\overline{\Delta}_f |\overline{\nabla} u|^2 - \langle \overline{\nabla} u, \overline{\nabla}(\overline{\Delta}_f u)\rangle\big) e^{-f} \\ = -\int_{\partial\Omega} (\Delta_f u) u_\nu e^{-f} - \int_{\partial\Omega} H_f u_\nu^2 e^{-f} + \int_{\partial\Omega} \big(\langle \nabla u, \nabla u_\nu\rangle - A(\nabla u, \nabla u)\big) e^{-f} \\ + \int_\Omega (\overline{\Delta}_f u)^2 e^{-f} \\ = -2\int_{\partial\Omega} (\Delta_f u) u_\nu e^{-f} - \int_{\partial\Omega} H_f u_\nu^2 e^{-f} - \int_{\partial\Omega} A(\nabla u, \nabla u) e^{-f} + \int_\Omega (\overline{\Delta}_f u)^2 e^{-f}.\end{aligned}$$

This immediately implies (8).

References

[Anderson 1985] M. T. Anderson, "Curvature estimates for minimal surfaces in 3-manifolds", *Ann. Sci. École Norm. Sup.* (4) **18**:1 (1985), 89–105. MR 87e:53098 Zbl 0578.49027

[Cao 2010] H.-D. Cao, "Recent progress on Ricci solitons", pp. 1–38 in *Recent advances in geometric analysis* (Taipei, 2007), edited by Y.-I. Lee et al., Adv. Lect. Math. **11**, International Press, Somerville, MA, 2010. MR 2011d:53061 Zbl 1201.53046 arXiv 0908.2006

[Cao and Zhou 2010] H.-D. Cao and D. Zhou, "On complete gradient shrinking Ricci solitons", *J. Differential Geom.* **85**:2 (2010), 175–185. MR 2011k:53040 Zbl 1246.53051

[Cheeger and Gromoll 1972] J. Cheeger and D. Gromoll, "On the structure of complete manifolds of nonnegative curvature", *Ann. of Math.* (2) **96** (1972), 413–443. MR 46 #8121 Zbl 0246.53049

[Cheng and Zhou 2013] X. Cheng and D. Zhou, "Volume estimate about shrinkers", *Proc. Amer. Math. Soc.* **141**:2 (2013), 687–696. MR 2996973 Zbl 1262.53030

[Cheng et al. 2012] X. Cheng, T. Mejia, and D. Zhou, "Stability and compactness for complete f-minimal surfaces", preprint, 2012. To appear in Trans. Amer. Math. Soc. arXiv 1210.8076

[Choi and Schoen 1985] H. I. Choi and R. Schoen, "The space of minimal embeddings of a surface into a three-dimensional manifold of positive Ricci curvature", *Invent. Math.* **81**:3 (1985), 387–394. MR 87a:58040 Zbl 0577.53044

[Choi and Wang 1983] H. I. Choi and A. N. Wang, "A first eigenvalue estimate for minimal hypersurfaces", *J. Differential Geom.* **18**:3 (1983), 559–562. MR 85d:53028 Zbl 0523.53055

[Colding and Minicozzi 2002] T. H. Colding and W. P. Minicozzi, II, "Estimates for parametric elliptic integrands", *Int. Math. Res. Not.* **2002**:6 (2002), 291–297. MR 2002k:53060 Zbl 1002.53035

[Colding and Minicozzi 2011] T. H. Colding and W. P. Minicozzi, II, *A course in minimal surfaces*, Graduate Studies in Mathematics **121**, American Mathematical Society, Providence, RI, 2011. MR 2780140 Zbl 1242.53007

[Colding and Minicozzi 2012a] T. H. Colding and W. P. Minicozzi, II, "Generic mean curvature flow, I: Generic singularities", *Ann. of Math.* (2) **175**:2 (2012), 755–833. MR 2993752 Zbl 1239.53084

[Colding and Minicozzi 2012b] T. H. Colding and W. P. Minicozzi, II, "Smooth compactness of self-shrinkers", *Comment. Math. Helv.* **87**:2 (2012), 463–475. MR 2914856 Zbl 1258.53069

[Ding and Xin 2013] Q. Ding and Y. L. Xin, "Volume growth, eigenvalue and compactness for self-shrinkers", *Asian J. Math.* **17**:3 (2013), 443–456. MR 3119795 Zbl 1283.53062

[Li and Wei 2012] H. Li and Y. Wei, "f-minimal surface and manifold with positive m-Bakry–Émery Ricci curvature", preprint, 2012. To appear in J. Geom. Anal. arXiv 1209.0895v1

[Ma and Du 2010] L. Ma and S.-H. Du, "Extension of Reilly formula with applications to eigenvalue estimates for drifting Laplacians", *C. R. Math. Acad. Sci. Paris* **348**:21-22 (2010), 1203–1206. MR 2011m:58051 Zbl 1208.58028

[Morgan 2005] F. Morgan, "Manifolds with density", *Notices Amer. Math. Soc.* **52**:8 (2005), 853–858. MR 2006g:53044 Zbl 1118.53022

[Munteanu and Wang 2012] O. Munteanu and J. Wang, "Analysis of weighted Laplacian and applications to Ricci solitons", *Comm. Anal. Geom.* **20**:1 (2012), 55–94. MR 2903101 Zbl 1245.53039

[Munteanu and Wang 2014] O. Munteanu and J. Wang, "Geometry of manifolds with densities", *Adv. Math.* **259** (2014), 269–305. MR 3197658 Zbl 1290.53048

[Reilly 1977] R. C. Reilly, "Applications of the Hessian operator in a Riemannian manifold", *Indiana Univ. Math. J.* **26**:3 (1977), 459–472. MR 57 #13799 Zbl 0391.53019

[Wei and Wylie 2009] G. Wei and W. Wylie, "Comparison geometry for the Bakry–Émery Ricci tensor", *J. Differential Geom.* **83**:2 (2009), 377–405. MR 2011a:53064 Zbl 1189.53036

[White 1987] B. White, "Curvature estimates and compactness theorems in 3-manifolds for surfaces that are stationary for parametric elliptic functionals", *Invent. Math.* **88**:2 (1987), 243–256. MR 88g:58037 Zbl 0615.53044

[Yang and Yau 1980] P. C. Yang and S. T. Yau, "Eigenvalues of the Laplacian of compact Riemann surfaces and minimal submanifolds", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **7**:1 (1980), 55–63. MR 81m:58084 Zbl 0446.58017

Received July 4, 2013. Revised May 14, 2014.

XU CHENG
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE FEDERAL FLUMINENSE - UFF
24020-140 CENTRO, NITERÓI-RJ
BRAZIL
xcheng@impa.br

TITO MEJIA
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE FEDERAL FLUMINENSE - UFF
24020-140 CENTRO, NITERÓI-RJ
BRAZIL
tmejia.uff@gmail.com

DETANG ZHOU
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE FEDERAL FLUMINENSE - UFF
24020-140 CENTRO, NITERÓI-RJ
BRAZIL
zhou@impa.br

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.369

LEFSCHETZ NUMBERS OF SYMPLECTIC INVOLUTIONS ON ARITHMETIC GROUPS

STEFFEN KIONKE

The reduced norm-one group G of a central simple algebra is an inner form of the special linear group, and an involution on the algebra induces an automorphism of G. We study the action of such automorphisms in the cohomology of arithmetic subgroups of G. The main result is a precise formula for Lefschetz numbers of automorphisms induced by involutions of symplectic type. Our approach is based on a careful study of the smoothness properties of group schemes associated with orders in central simple algebras. Along the way we also derive an adelic reformulation of Harder's Gauss–Bonnet theorem.

1. Introduction

Let G be a semisimple linear algebraic group defined over the field $\mathbb{Q}$ of rational numbers. Given a torsion-free arithmetic subgroup $\Gamma \subset G(\mathbb{Q})$, it is in general a very difficult task to compute the (cohomological) Betti numbers of Γ. However Harder's Gauss–Bonnet theorem [Harder 1971] makes it possible to determine the Euler characteristic of arithmetic groups. If the Euler characteristic is nonzero, one can extract information on the Betti numbers. Moreover, whether or not the Euler characteristic vanishes only depends on the structure of the associated real Lie group $G(\mathbb{R})$ (see the remark on page 384). If the Euler characteristic vanishes, Lefschetz numbers of automorphisms of finite order of G are a suitable substitute to gain insight into the cohomology of Γ. The idea to study Lefschetz numbers in the cohomology of arithmetic groups goes back to Harder [1975]. A general method was developed by J. Rohlfs, first for Galois automorphisms [1978] and later in a general adelic setting [1990]. Lefschetz numbers were also studied in [Lee and Schwermer 1983; Lai 1991]. However, only very few groups have been considered in detail; most frequently Lefschetz numbers on Bianchi groups have been studied (see [Krämer 1985; Rohlfs 1985; Sengün and Türkelli 2012; Kionke and Schwermer 2012]). In this article we describe a method (based on Rohlfs'

The author was supported by FWF Austrian Science Fund, grant P 21090-N13.
MSC2010: primary 11F75; secondary 20H10, 20G35.
Keywords: arithmetic group, cohomology, Lefschetz number, involution.

approach) to compute Lefschetz numbers of specific automorphisms on arithmetic subgroups of inner forms of the special linear group.

More precisely, let F be an algebraic number field and let A be a central simple F-algebra. The reduced norm $\mathrm{nrd}_{A/F}$ is a polynomial function on A and the associated reduced norm-one group $G = \mathrm{SL}_A$ is a linear algebraic group defined over F. Indeed, the algebraic group G is an inner form of the special linear group. If A has an involution σ of symplectic type (see the definition on page 377), then the composition of σ with the group inversion yields an automorphism σ^* of G. We study the Lefschetz numbers of such automorphisms induced by involutions of symplectic type.

1A. ***The main result.*** Let F be an algebraic number field and let $\mathbb{O}$ denote its ring of integers. Let A be a central simple F-algebra. For our purposes we may assume that $A = M_n(D)$ for some quaternion F-algebra D (see Section 1C).

Let $\Lambda_D \subseteq D$ be a maximal $\mathbb{O}$-order in D; then $\Lambda := M_n(\Lambda_D)$ is a maximal $\mathbb{O}$-order in A. For a nontrivial ideal $\mathfrak{a} \subseteq \mathbb{O}$ we study the cohomology of the principal congruence subgroups

$$\Gamma(\mathfrak{a}) := \{g \in M_n(\Lambda_D) \mid \mathrm{nrd}_A(g) = 1 \text{ and } g \equiv 1 \bmod \mathfrak{a}\}$$

of G. In fact, for $n \geq 2$ the groups $\Gamma(\mathfrak{a})$ have vanishing Euler characteristic.

The quaternion algebra D is equipped with a unique involution of symplectic type $\tau_c : D \to D$, called *conjugation*, which induces an involution of symplectic type $\tau : A \to A$ by $\tau(x) := \tau_c(x)^T$; that is, apply τ_c to every entry of the matrix and then transpose the matrix. We will call τ the *standard involution of symplectic type* on $M_n(D)$. Composition of τ with the group inversion yields an automorphism τ^* of order two on G. Note that the congruence groups $\Gamma(\mathfrak{a})$ are stable under τ^*. Fix a rational representation $\rho : G \times_F \overline{F} \to \mathrm{GL}(W)$ of G (defined over the algebraic closure of F) on a finite dimensional vector space. If W is equipped with a compatible τ^*-action (see the definition on page 389), then we can define the Lefschetz number $\mathcal{L}(\tau^*, \Gamma(\mathfrak{a}), W)$ of τ^* in the cohomology $H^\bullet(\Gamma(\mathfrak{a}), W)$.

Main Theorem. *Assume that $\Gamma(\mathfrak{a})$ is torsion-free. If D is totally definite*, *we assume further that $n \geq 2$. The Lefschetz number $\mathcal{L}(\tau^*, \Gamma(\mathfrak{a}), W)$ is zero if F is not totally real.*

If F is totally real, *the Lefschetz number is given by the formula*

$$\mathcal{L}(\tau^*, \Gamma(\mathfrak{a}), W) = 2^{-r}\, \mathrm{N}(\mathfrak{a})^{n(2n+1)} \Delta_{\mathrm{rd}}(D)^{n(n+1)/2}\, \mathrm{Tr}(\tau^*|W) \prod_{j=1}^{n} M(j, \mathfrak{a}, D).$$

Here $\Delta_{\mathrm{rd}}(D)$ denotes the signed reduced discriminant of D (see the definition on page 390), *r denotes the number of real places of F ramified in D*, *and*

$$M(j,\mathfrak{a},D) := \zeta_F(1-2j)\prod_{\mathfrak{p}|\mathfrak{a}}\bigl(1-\mathrm{N}(\mathfrak{p})^{-2j}\bigr)\prod_{\substack{\mathfrak{p}\in\mathrm{Ram}_f(D)\\ \mathfrak{p}\nmid\mathfrak{a}}}\bigl(1+(-\mathrm{N}(\mathfrak{p}))^{-j}\bigr),$$

where $\mathrm{Ram}_f(D)$ *denotes the set of finite places of* F *where* D *ramifies and* ζ_F *denotes the Dedekind zeta-function of* F. *If* F *is totally real*, *then the Lefschetz number is zero if and only if* $\mathrm{Tr}(\tau^*|W) = 0$.

1B. ***Applications.*** We briefly give three applications of the above formula where we always assume F to be a totally real number field.

1B1. *Growth of the total Betti number.* The analysis of the asymptotic behaviour of Betti numbers of arithmetic groups is an important topic. Calegari and Emerton [2009] have provided strong asymptotic upper bounds. We can use the main theorem to obtain an asymptotic lower bound result.

Let G be the reduced norm-one group associated with the central simple F-algebra $M_n(D)$. For a torsion-free arithmetic subgroup $\Gamma \subseteq G(F)$ we define the *total Betti number* $B(\Gamma)$ as $\sum_{i=0}^{\infty} \dim H^i(\Gamma,\mathbb{C})$. Note that this is a finite sum since torsion-free arithmetic groups are of type (FL) (see [Borel and Serre 1973, Theorem 11.4.4]).

Corollary 1.1. *Let* $\Gamma_0 \subset G(F)$ *be an arithmetic subgroup. For any ideal* $\mathfrak{a} \subset \mathbb{O}$ *we define* $\Gamma_0(\mathfrak{a}) := \Gamma_0 \cap \Gamma(\mathfrak{a})$. *There is a positive real number* $\kappa > 0$, *depending on* F, D, Γ_0, *and* n, *such that*

$$B(\Gamma_0(\mathfrak{a})) \geq \kappa[\Gamma_0 : \Gamma_0(\mathfrak{a})]^{\frac{n(2n+1)}{4n^2-1}}$$

for every ideal $\mathfrak{a}$ *such that* $\Gamma(\mathfrak{a})$ *is torsion-free.*

A proof of this corollary will be given in Section 5E.

1B2. *Rationality of zeta values.* Note that the Lefschetz number is an integer since τ^* is of order two. We obtain a new proof of a classical theorem of Siegel [1969] and Klingen [1962].

Corollary 1.2. *If* F *is a totally real number field*, *then* $\zeta_F(1-2m)$ *is a nonzero rational number for all integers* $m \geq 1$.

Proof. Apply the main theorem with $D = M_2(F)$, $\Lambda_D = M_2(\mathbb{O})$ and choose W to be the trivial one-dimensional representation. We see that for every $n \geq 1$ and all sufficiently small ideals $\mathfrak{a} \subseteq \mathbb{O}$, the number

$$\mathrm{N}(\mathfrak{a})^{n(2n+1)}\prod_{j=1}^{n}\Bigl(\zeta_F(1-2j)\prod_{\mathfrak{p}|\mathfrak{a}}\bigl(1-\mathrm{N}(\mathfrak{p})^{-2j}\bigr)\Bigr)$$

is a nonzero integer. The claim follows by induction on m. □

1B3. *Cohomology of cocompact Fuchsian groups.* Let D be a division quaternion algebra over F such that D is split at precisely one real place v_0 of F. Therefore $r = [F:\mathbb{Q}] - 1$ is the number of real places ramified in D.

Let $\Lambda = \Lambda_D$ be a maximal $\mathbb{O}$-order in D. We consider the reduced norm-one group $G = \mathrm{SL}_D$ defined over F. The associated real Lie group is

$$G_\infty \cong \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_1(\mathbb{H})^r.$$

Note that the group $\mathrm{SL}_1(\mathbb{H})$ is compact and so the projection $p_1 : G_\infty \to \mathrm{SL}_2(\mathbb{R})$ onto the first factor is a proper and open homomorphism of Lie groups. In particular, every discrete torsion-free subgroup $\Gamma \subseteq G_\infty$ maps via p_1 isomorphically to a discrete subgroup in $\mathrm{SL}_2(\mathbb{R})$.

Let $\mathfrak{a} \subseteq \mathbb{O}$ be a proper ideal such that $\Gamma(\mathfrak{a})$ is torsion-free. We will interpret $\Gamma(\mathfrak{a})$ as a subgroup of $\mathrm{SL}_2(\mathbb{R})$. Note that since we assumed D to be a division algebra, the group $\Gamma(\mathfrak{a})$ is a cocompact Fuchsian group [Katok 1992, Theorem 5.4.1].

Let $\mathfrak{h} = \mathrm{SL}_2(\mathbb{R})/\,\mathrm{SO}(2)$ be the Poincaré upper half-plane.

Corollary 1.3. *The compact Riemann surface* $\mathfrak{h}/\,\Gamma(\mathfrak{a})$ *has genus*

$$g = 1 + 2^{-[F:\mathbb{Q}]}\,\mathrm{N}(\mathfrak{a})^3 |\Delta_{\mathrm{rd}}(D)\zeta_F(-1)| \prod_{\mathfrak{p}|\mathfrak{a}} \bigl(1-\mathrm{N}(\mathfrak{p})^{-2}\bigr) \prod_{\substack{\mathfrak{p}\in \mathrm{Ram}_f(D)\\ \mathfrak{p}\nmid\mathfrak{a}}} \bigl(1-\mathrm{N}(\mathfrak{p})^{-1}\bigr).$$

This implies an explicit formula for the first Betti number $b_1(\Gamma(\mathfrak{a}))$ *since*

$$b_1(\Gamma(\mathfrak{a})) = \dim H^1(\Gamma(\mathfrak{a}), \mathbb{C}) = 2g.$$

Proof. Consider the main theorem for $n = 1$. Note that for $n = 1$ the automorphism τ^* is actually the identity. This means that, using the main theorem with the trivial representation,

$$\mathscr{L}(\tau^*, \Gamma(\mathfrak{a}), \mathbb{C}) = \chi(\Gamma(\mathfrak{a})) = \chi(\mathfrak{h}/\,\Gamma(\mathfrak{a})).$$

Note that the sign of the Lefschetz number is -1. Since $\chi(\mathfrak{h}/\,\Gamma(\mathfrak{a})) = 2-2g$, the claim follows immediately. □

In fact Corollary 1.3 yields a precise formula for the dimension of the space of holomorphic weight-k modular forms for the group $\Gamma(\mathfrak{a})$ [Shimura 1971, Theorems 2.24 and 2.25].

1C. ***Reduction to quaternion algebras.*** Let A be a central simple F-algebra. If A has an involution σ of symplectic type (see the definition on page 377), then A is isomorphic to the opposed F-algebra A^{op}. This means that the class of A has order two in the Brauer group of F. Since the dimension of A is even, it follows from [Reiner 2003, Theorem (32.19)] that A is isomorphic to a matrix algebra $M_n(D)$ over a quaternion algebra D. Therefore we always assume $A = M_n(D)$.

Let τ be the standard involution of symplectic type on $M_n(D)$. Note that in this case $\sigma = \mathrm{int}(g) \circ \tau$ for an element $g \in \mathrm{GL}_n(D)$ with $\tau(g) = g$. Due to this observation it is only a minor restriction if we focus on the standard symplectic involution τ.

1D. ***Structure of this article.*** In Section 2 we give a short general treatment of smooth group schemes over Dedekind rings which are associated with orders in central simple algebras. In particular we treat integral models of inner forms of the special linear group. Further, we consider the fixed points groups attached to involutions. An important tool in the proof of the main theorem will be the pfaffian as a map in nonabelian Galois cohomology (Section 2D). In Section 3 we give an adelic reformulation of Harder's Gauss–Bonnet theorem which hinges on the notion of smooth group scheme. The calculation of the Lefschetz number is based on Rohlfs' method which we summarise in Section 4. Finally the proof of the main theorem is contained in Section 5. It consists of two major steps. The first is the analysis of various nonabelian Galois cohomology sets which occur in Rohlfs' decomposition. The second step is the calculation of the Euler characteristics of the fixed point groups using Harder's Gauss–Bonnet theorem.

Notation. Apart from Section 2, where we work in a more general setting, we use the following notation: F is an algebraic number field and $\mathbb{O}$ denotes its ring of integers. Let V denote the set of places of F. We have $V = V_\infty \cup V_f$, where V_∞ and V_f denote the set of Archimedean and finite places of F, respectively. Let $v \in V$ be a place of F; we denote the completion of F at v by F_v. The valuation ring of F_v is denoted by $\mathbb{O}_v$ and the prime ideal in $\mathbb{O}_v$ is denoted by $\mathfrak{p}_v$. For a nonzero ideal $\mathfrak{a} \subseteq \mathbb{O}$ the ideal norm is defined by $\mathrm{N}(\mathfrak{a}) := |\mathbb{O}/\mathfrak{a}|$. As usual $\mathbb{A}$ denotes the ring of adeles of F and $\mathbb{A}_f$ is the ring of finite adeles.

2. Group schemes associated with orders in central simple algebras

In this section we will investigate the smoothness properties of group schemes attached to orders in central simple algebras. Throughout, R denotes a Dedekind ring and k denotes its field of fractions. For simplicity we assume $\mathrm{char}(k) = 0$. In our applications R is usually the ring of integers of an algebraic number field or a complete discrete valuation ring.

The term *scheme* always refers to an affine scheme of finite type; the same holds for group schemes. Recall that a scheme $\mathfrak{X}$ defined over R is *smooth* if for every commutative R-algebra C and every nilpotent ideal $I \subseteq C$ the induced map $\mathfrak{X}(C) \to \mathfrak{X}(C/I)$ is surjective. Suppose R is a complete discrete valuation ring and let $\mathfrak{p}$ denote its prime ideal. We will frequently use the following property: if $\mathfrak{X}$ is a smooth R-scheme, then the induced map $\mathfrak{X}(R) \to \mathfrak{X}(R/\mathfrak{p}^e)$ is surjective for every integer $e \geq 1$ [Grothendieck 1964, Corollary 19.3.11]. If G is a group scheme, then we denote the Lie algebra of G by $\mathrm{Lie}(G)$.

2A. ***The general linear group over an order.*** Let A be a central simple k-algebra and let Λ be an R-order in A. Since Λ is a finitely generated torsion-free R-module, it is a finitely generated *projective* R-module [Reiner 2003, Theorem (4.13)]. The functor Λ_a from the category of commutative R-algebras to the category of rings defined by $C \mapsto \Lambda \otimes_R C$ is represented by the symmetric algebra $S_R(\Lambda^*)$, where $\Lambda^* = \mathrm{Hom}_R(\Lambda, R)$. In fact it defines a smooth R-scheme [Grothendieck 1964, Proposition 19.3.2].

Recall that, since Λ is finitely generated and projective, one can attach to every R-linear endomorphism φ of Λ its determinant $\det(\varphi) \in R$. More precisely, here the determinant of φ is just the determinant of the k-linear extension $\varphi \otimes \mathrm{Id}_k : A \to A$. As usual one defines the *norm* of an element $x \in \Lambda$ to be the determinant of the left multiplication with x. One can check that the norm defines a morphism of schemes over R

$$\mathrm{N}_{\Lambda/R} : \Lambda_a \to \mathbb{A}^1/R$$

to the affine line $\mathbb{A}^1$ defined over R. This can be seen, for instance, by observing that the norm is a natural transformation of functors. Let C be a commutative R-algebra. An element $x \in \Lambda \otimes_R C$ is a unit if and only if $\mathrm{N}_{\Lambda/R}(x) \in C^\times$. It follows from the next lemma that the associated unit group functor $\mathrm{GL}_\Lambda : C \mapsto (\Lambda \otimes_R C)^\times$ is a smooth group scheme over R.

Lemma 2.1. *Let $\mathbb{A}^1$ denote the affine line over R. Let $\mathfrak{X}$ be an affine scheme over R with a morphism $f : \mathfrak{X} \to \mathbb{A}^1$. The subfunctor $\mathfrak{Y}$ (from the category of commutative R-algebras to the category of sets) defined by*

$$C \mapsto \{y \in \mathfrak{X}(C) \mid f(y) \in C^\times\}$$

is an affine scheme and the natural transformation $\mathfrak{Y} \to \mathfrak{X}$ is a morphism of schemes. If $\mathfrak{X}$ is smooth, then $\mathfrak{Y}$ has the same property.

Proof. Let $R[\mathfrak{X}]$ be the coordinate ring of $\mathfrak{X}$ and let $P \in R[\mathfrak{X}]$ be the polynomial defining f. Note that $\mathfrak{Y}$ is canonically isomorphic to the functor given by

$$C \mapsto \{(y, z) \in \mathfrak{X}(C) \times C \mid f(y)z = 1\}.$$

Using this it is easily checked that the R-algebra $S := R[\mathfrak{X}] \otimes_R R[T]/(P \otimes T - 1)$ represents $\mathfrak{Y}$. Clearly, S is of finite type since $R[\mathfrak{X}]$ is of finite type.

It remains to show that $\mathfrak{Y}$ is smooth whenever $\mathfrak{X}$ is smooth. Assume that $\mathfrak{X}$ is smooth and take a commutative R-algebra C with an ideal J such that $J^2 = 0$. By assumption $\mathfrak{X}(C) \to \mathfrak{X}(C/J)$ is surjective, so given $y \in \mathfrak{Y}(C/J)$ we find $x \in \mathfrak{X}(C)$ projecting to y. By assumption $f(x) + J$ is a unit in C/J. In particular, we can find $z \in C$ with $f(x)z \in 1+J$. However, $1+J$ consists entirely of units and thus $f(x) \in C^\times$. We deduce that $\mathfrak{Y}$ is smooth. □

We stress this once more: in this article GL_Λ is always a functor and not a group. If we take $\Lambda = R$ then we call GL_Λ the *multiplicative group* (or multiplicative group scheme) defined over R, and we denote it by $\mathbb{G}_m$. Note that the norm defines a homomorphism of R-group schemes

$$\mathrm{N}_{\Lambda/R} : \mathrm{GL}_\Lambda \to \mathbb{G}_m.$$

We also point out that the Lie algebra of GL_Λ can be (and will be) identified with Λ_a in a natural way.

2B. *The special linear group over an order.*

2B1. *Reduced norm and trace.* Let A be a central simple k-algebra. We consider the reduced norm and trace (for definitions see [Reiner 2003, Section 9] or [Weil 1995, Chapter IX, §2]). It was observed by Weil that the reduced norm and trace are polynomial functions. We reformulate this in schematic language: there is a unique element $\mathrm{nrd}_{A/k}$ in the symmetric algebra $S_k(A^*)$ (here $A^* = \mathrm{Hom}_k(A,k)$) such that for every splitting field ℓ of A and every splitting $\varphi : A \otimes_k \ell \xrightarrow{\simeq} M_n(\ell)$ the induced map

$$S(\varphi^*) : S_\ell(M_n(\ell)^*) \to S_k(A^*) \otimes_k \ell$$

maps the determinant to $\mathrm{nrd}_{A/k} \otimes 1$. Similarly there is the reduced trace $\mathrm{trd}_{A/k} \in A^*$ with an analogous property.

Let $\Lambda \subseteq A$ be an R-order. We show that the reduced norm and trace are defined over R in the appropriate sense. For the reduced trace this is easy: elements in Λ are integral over R, hence the reduced trace takes values in R on the order Λ and defines an R-linear map $\Lambda \to R$. In particular we obtain a morphism of schemes over R:

$$\mathrm{trd}_{\Lambda/R} : \Lambda_a \to \mathbb{A}^1/R.$$

Consider the reduced norm. From [Reiner 2003, (9.7)] one can deduce that $\mathrm{nrd}^n_{A/k}$ and $\mathrm{N}_{\Lambda/R}$ agree as elements in the coordinate ring $S_k(A^*)$. However, the coordinate ring $S_R(\Lambda^*)$ of Λ_a is integrally closed in $S_k(A^*)$ and we conclude that the reduced norm is defined over R. This means that there is a polynomial $\mathrm{nrd}_{\Lambda/R} \in S_R(\Lambda^*)$ that defines the reduced norm as a morphism of R-schemes:

$$\mathrm{nrd}_{\Lambda/R} : \Lambda_a \to \mathbb{A}^1.$$

We can also restrict the reduced norm to the unit group and obtain a homomorphism of group schemes:

$$\mathrm{nrd}_{\Lambda/R} : \mathrm{GL}_\Lambda \to \mathbb{G}_m/R.$$

Definition. The *special linear group* SL_Λ over the order Λ is the group scheme over R defined by the kernel of the reduced norm:

$$\mathrm{SL}_\Lambda = \ker(\mathrm{nrd}_{\Lambda/R} : \mathrm{GL}_\Lambda \to \mathbb{G}_m).$$

2B2. *Smoothness of the special linear group.* Whereas the general linear group is always smooth, independent of the chosen order, the smoothness of the special linear group depends on the underlying order. Recall the following useful result.

Proposition 2.2 (smoothness of kernels). *Let $f : G \to H$ be a morphism between two smooth group schemes over R. If the derivative* $\mathrm{d}(f) : \mathrm{Lie}(G)(R) \to \mathrm{Lie}(H)(R)$ *is surjective, then the group scheme $K := \ker(f)$ is smooth over R.*

Proof. This follows from the theorem of infinitesimal points (see [Demazure and Gabriel 1970, p. 208]) and some easy diagram chasing. □

As a matter of fact the derivative of the reduced norm $\mathrm{d}(\mathrm{nrd}_{\Lambda/R}) : \Lambda_a \to \mathbb{A}^1$ is the reduced trace. Having this in mind we make the following definition.

Definition. An R-order Λ in a central simple k-algebra is called *smooth* if the reduced trace $\mathrm{trd}_{\Lambda/R} : \Lambda \to R$ is surjective.

Note that smoothness of orders is a local property.

Corollary 2.3. *If the order Λ is smooth then the scheme SL_Λ is smooth.*

Proof. This follows immediately from Proposition 2.2 using the fact that the derivative of the reduced norm is the reduced trace. □

In fact, the converse statement also holds under the assumption $\mathrm{char}(R) = 0$. However, we shall not need this result. The next proposition shows that smooth orders exist.

Proposition 2.4. *Assume that $R/\mathfrak{p}$ is finite for every prime ideal $\mathfrak{p}$. Then every maximal R-order in a central simple k-algebra is smooth.*

Proof. Let A be a central simple k-algebra and let $\Lambda \subset A$ be a maximal R-order. Since Λ is maximal in A if and only if all $\mathfrak{p}$-adic completions are maximal orders [Reiner 2003, Corollary (11.6)], and since smoothness of Λ is a local property, we may assume that R is a complete discrete valuation ring. Recall that A is isomorphic to a matrix algebra $M_r(D)$ over a central division algebra D. Moreover, D has a unique maximal R-order $\Delta \subseteq D$ and Λ is (up to conjugation) the maximal order $M_r(\Delta)$ in A [Reiner 2003, Theorem (17.3)]. It is known that the reduced trace of a matrix $x = (x_{ij})_{i,j=1}^r \in M_r(D)$ is given by

$$\mathrm{trd}_{A/k}(x) = \sum_{i=1}^{r} \mathrm{trd}_{D/k}(x_{ii})$$

[Weil 1995, Corollary 2, Chapter IX, §2]. Hence we may assume that $A = D$ is a division algebra and $\Lambda = \Delta$ is the unique maximal order. Let $\dim_k D = n^2$ and let ℓ/k be the unique unramified extension of k of degree $[\ell : k] = n$. The field ℓ embeds into D as a maximal subfield and the reduced trace $\mathrm{trd}_{D/k}$ on the elements of ℓ agrees with the field trace $\mathrm{Tr}_{\ell/k}$ [Reiner 2003, proof of Theorem (14.9)]. Let o_ℓ denote the valuation ring of ℓ. The image of o_ℓ under the embedding $\ell \to D$ lies in the maximal order Δ. Finally the surjectivity of $\mathrm{trd}_{D/k} : \Delta \to R$ follows from the well-known surjectivity of the field trace $\mathrm{Tr}_{\ell/k} : o_\ell \to R$. □

2C. ***Involutions and fixed point groups.*** Let A be a central simple k-algebra. An *involution* τ on A is an additive mapping $\tau : A \to A$ of order two such that $\tau(xy) = \tau(y)\tau(x)$ for all $x, y \in A$. We say that τ is of the *first kind* if τ is k-linear. Otherwise, we say that τ is of the *second kind*. In this article all involutions are of the first kind unless the contrary is explicitly stated. We will mostly focus on involutions of symplectic type.

Definition. We say that an involution τ on A is of *symplectic type* if there is a splitting field ℓ of the algebra A, a splitting

$$\varphi : A \otimes_k \ell \xrightarrow{\simeq} M_{2n}(\ell),$$

and a skew-symmetric matrix $a \in M_{2n}(\ell)$ satisfying $\varphi(\tau(x)) = a\varphi(x)^T a^{-1}$ for all elements $x \in A \otimes_k \ell$. If this is the case, then every splitting (over any splitting field) admits such a matrix.

Let $\tau : A \to A$ be an involution of the first kind. Let Λ be an R-order in A and assume that Λ is τ-stable. Since $\tau : \Lambda \to \Lambda$ is R-linear, we obtain a morphism of R-schemes

$$\tau : \Lambda_a \to \Lambda_a.$$

We restrict τ to the unit group GL_Λ and compose it with the group inversion to obtain a homomorphism of group schemes

$$\tau^* : \mathrm{GL}_\Lambda \to \mathrm{GL}_\Lambda.$$

We define $G(\Lambda, \tau)$ to be the group of fixed points of τ^*, that is, for every commutative R-algebra C we obtain

$$G(\Lambda, \tau)(C) = \{x \in (\Lambda \otimes_R C)^\times \mid \tau(x)x = 1\}.$$

We analyse the smoothness properties of group schemes constructed in this way. Define the R-submodule $\mathrm{Sym}(\Lambda, \tau) = \{x \in \Lambda \mid \tau(x) = x\}$ of Λ and note that it is a direct summand.

Lemma 2.5. *For every commutative R-algebra C, every $y \in \Lambda \otimes_R C$ and every $x \in \mathrm{Sym}(\Lambda, \tau) \otimes_R C$ we have $\tau(y)xy \in \mathrm{Sym}(\Lambda, \tau) \otimes_R C$.*

Proof. We can write $y = \sum_i u_i \otimes c_i$ for certain $u_i \in \Lambda$ and $c_i \in C$. The claim is linear in x, hence we may assume $x = e \otimes c$ with $e \in \mathrm{Sym}(\Lambda, \tau)$ and $c \in C$. We calculate

$$\begin{aligned}\tau(y)xy &= \sum_{i,j} \tau(u_i)eu_j \otimes cc_ic_j \\ &= \sum_i \tau(u_i)eu_i \otimes cc_i^2 + \sum_{i<j}(\tau(u_i)eu_j + \tau(u_j)eu_i) \otimes cc_ic_j,\end{aligned}$$

and see that $\tau(y)xy \in \mathrm{Sym}(\Lambda, \tau) \otimes_R C$ since $\tau(u_i)eu_i$ and $\tau(u_i)eu_j + \tau(u_j)eu_i$ are elements of $\mathrm{Sym}(\Lambda, \tau)$. □

Definition. The order Λ is called *τ-smooth* if the map $s : \Lambda \to \mathrm{Sym}(\Lambda, \tau)$ defined by $x \mapsto x + \tau(x)$ is surjective. Clearly τ-smoothness is a local property.

Proposition 2.6. *If an R-order Λ is τ-smooth, then the scheme $G(\Lambda, \tau)$ is smooth.*

Proof. We set $G := G(\Lambda, \tau)$. Let C be a commutative R-algebra with an ideal $I \subseteq C$ such that $I^2 = 0$. We have to show that the canonical map $G(C) \to G(C/I)$ is surjective. Take $\bar{y} \in G(C/I)$. Since the unit group scheme GL_Λ is smooth (see Section 2A), we can find $y \in \mathrm{GL}_\Lambda(C) = (\Lambda \otimes C)^\times$ mapping to $\bar{y}$ modulo I. Since $\bar{y}$ is in the fixed point group of τ^*, this implies that

$$\tau(y)y = 1 + \rho$$

with some $\rho \in \Lambda \otimes I$.

We consider $E := \mathrm{Sym}(\Lambda, \tau)$ and we obtain $\tau(y)y \in E \otimes_R C$ by Lemma 2.5. Consequently, there is $u \in \Lambda \otimes_R C$ such that $\tau(u) + u = y$. Moreover, $1 \in E$; thus there is some $v \in \Lambda \otimes_R C$ with $\tau(v) + v = 1$. We deduce that $\rho = \tau(u - v) + (u - v)$ is an element in $E \otimes_R C$, and thus

$$\rho \in (E \otimes_R C) \cap (\Lambda \otimes_R I) = E \otimes_R I.$$

As a last step we use once again that Λ is τ-smooth and deduce that there is some $w \in \Lambda \otimes I$ with $\rho = \tau(w) + w$. We put $y' := y(1 - w)$, which is congruent to $\bar{y}$ modulo I and satisfies

$$\begin{aligned}\tau(y')y' &= (1 - \tau(w))\tau(y)y(1 - w) = (1 - \tau(w))(1 + \rho)(1 - w) \\ &= 1 + \rho - \tau(w) - w = 1.\end{aligned}$$

Therefore $y' \in G(C)$ and y' maps to $\bar{y} \in G(C/I)$ under the canonical map. □

2D. ***Involutions of symplectic type and the pfaffian.*** Let A be a central simple k-algebra with an involution of symplectic type τ. Let Λ be a τ-stable R-order in A.

2D1. *The pfaffian.* Set $E := \mathrm{Sym}(\Lambda, \tau)$ in the notation of Section 2C. The inclusion $\iota : E \to \Lambda$ induces a morphism of R-algebras

$$S(\iota^*) : S_R(\Lambda^*) \to S_R(E^*).$$

Recall that the reduced norm is given by a polynomial function $\mathrm{nrd}_{\Lambda/R} \in S_R(\Lambda^*)$ (see Section 2B1). We define $\mathrm{nrd}_{|E} := S(\iota^*)(\mathrm{nrd}_{\Lambda/R}) \in S_R(E^*)$. We will construct a *pfaffian*, that is, a polynomial $\mathrm{pf}_\tau \in S_R(E^*)$ such that $\mathrm{nrd}_{|E} = \mathrm{pf}_\tau^2$.

Let L/k be any field extension. It follows from [Knus et al. 1998, Proposition 2.9] that for every $x \in E \otimes_R L$ the reduced norm $\mathrm{nrd}_{|E}(x)$ is a square in L. Therefore, we may deduce that there is a polynomial $f \in S_R(E^*)$ such that

$$f^2 = \mathrm{nrd}_{|E}.$$

We normalise this polynomial $\mathrm{pf}_\tau := \pm f$ such that $\mathrm{pf}_\tau(1) = 1$ and we call pf_τ the *pfaffian* with respect to τ.

Lemma 2.7. *Let $S(\tau^*)$ denote the automorphism of the symmetric R-algebra $S_R(\Lambda^*)$ which is induced by τ. The following assertions hold*:

(i) $S(\tau^*)(\mathrm{nrd}_{\Lambda/R}) = \mathrm{nrd}_{\Lambda/R}$.

(ii) *For all $y \in \Lambda \otimes_R C$ and all $x \in \mathrm{Sym}(\Lambda, \tau) \otimes_R C$, we have*

$$\mathrm{pf}_\tau(\tau(y)xy) = \mathrm{nrd}_{\Lambda/R}(y)\,\mathrm{pf}_\tau(x),$$

where C is any commutative R-algebra.

Proof. To prove the first claim we may work over fields. However, over fields this is the well-known statement [Knus et al. 1998, Corollary 2.2].

The same proof works for the second statement. Note that $\tau(y)xy$ lies in $\mathrm{Sym}(\Lambda, \tau) \otimes_R C$ by Lemma 2.5. Both are polynomial functions on $\Lambda \times \mathrm{Sym}(\Lambda, \tau)$. If they agree over all fields then they agree as polynomials. However, over fields this is the result [Knus et al. 1998, Proposition 2.13]. □

Remark. Consider the fixed point group scheme $G = G(\Lambda, \tau)$ associated with τ. Let $x \in G(C)$ for some commutative R-algebra C. We see from $\tau(x)x = 1$ and Lemma 2.7 that

$$\mathrm{nrd}_{\Lambda/R}(x) = \mathrm{pf}_\tau(\tau(x)x) = \mathrm{pf}_\tau(1) = 1.$$

Hence the reduced norm restricts to the trivial character on $G(\Lambda, \tau)$.

2D2. *The cohomological pfaffian.* We study nonabelian Galois cohomology of τ^* with values in the groups $\mathrm{GL}_\Lambda(C)$ and $\mathrm{SL}_\Lambda(C)$. For the definition of nonabelian cohomology we refer the reader to [Serre 1994; 1979, pages 123–126] or [Knus et al. 1998, Chapter VII]. We shall in this context often denote τ and τ^* by left exponents, that is, we write ${}^{\tau^*}\!x$ for $\tau^*(x)$.

Let C be a commutative R-algebra and assume that C is flat as an R-module. A *cocycle* b in $Z^1(\tau^*, \mathrm{GL}_\Lambda(C))$ is an element of $(\Lambda \otimes C)^\times$ which satisfies $b\,{}^{\tau^*}b = 1$, or equivalently $b = {}^{\tau}b$. In other words

$$Z^1(\tau^*, \mathrm{GL}_\Lambda(C)) = \mathrm{Sym}(\Lambda \otimes_R C, \tau) \cap \mathrm{GL}_\Lambda(C).$$

The assumption that C is flat yields that $\mathrm{Sym}(\Lambda \otimes_R C, \tau) = \mathrm{Sym}(\Lambda, \tau) \otimes_R C$. Therefore we can apply the pfaffian associated with τ to cocycles in $Z^1(\tau^*, \mathrm{GL}_\Lambda(C))$. Two cocycles b and c are *cohomologous* if there is $y \in \mathrm{GL}_\Lambda(C)$ such that $b = {}^{\tau}ycy$. In this case it follows from Lemma 2.7 that $\mathrm{pf}_\tau(b) = \mathrm{nrd}_{\Lambda/R}(y)\,\mathrm{pf}_\tau(c)$. Therefore the pfaffian defines a morphism of pointed sets

$$\mathrm{pf}_\tau : H^1(\tau^*, \mathrm{GL}_\Lambda(C)) \to C^\times / \mathrm{nrd}_{\Lambda/R}(\mathrm{GL}_\Lambda(C)).$$

By the same reasoning we obtain a morphism of pointed sets

$$\mathrm{pf}_\tau : H^1(\tau^*, \mathrm{SL}_\Lambda(C)) \to \{x \in C^\times \mid x^2 = 1\}.$$

For simplicity we define $C^{(2)} := \{x \in C^\times \mid x^2 = 1\}$ and $C^\times_\Lambda := \mathrm{nrd}_{\Lambda/R}(\mathrm{GL}_\Lambda(C))$.

Proposition 2.8 (cohomological diagram for symplectic involutions). *Let τ be an involution of symplectic type on A and let Λ be a τ-stable R-order. For every commutative R-algebra C which is flat as an R-module, there is a commutative diagram of pointed sets with exact rows*:

$$\begin{tikzcd}
C^{(2)} \cap C^\times_\Lambda \arrow[r, "\delta"] \arrow[d, equal] & H^1(\tau^*, \mathrm{SL}_\Lambda(C)) \arrow[r, "j_*"] \arrow[d, "\mathrm{pf}_\tau"'] & H^1(\tau^*, \mathrm{GL}_\Lambda(C)) \arrow[r, "\mathrm{nrd}"] \arrow[d, "\mathrm{pf}_\tau"'] & C^\times/(C^\times_\Lambda)^2 \arrow[d, equal] \\
C^{(2)} \cap C^\times_\Lambda \arrow[r] & C^{(2)} \arrow[r] & C^\times/C^\times_\Lambda \arrow[r, "\cdot^2"] & C^\times/(C^\times_\Lambda)^2.
\end{tikzcd}$$

The map δ is injective and the lower row is an exact sequence of groups. Here j_* denotes the map induced by the inclusion $j : \mathrm{SL}_\Lambda(C) \to \mathrm{GL}_\Lambda(C)$.

Proof. The short exact sequence of groups

$$1 \longrightarrow \mathrm{SL}_\Lambda(C) \xrightarrow{j} \mathrm{GL}_\Lambda(C) \xrightarrow{\mathrm{nrd}} C^\times_\Lambda \longrightarrow 1$$

is an exact sequence of groups with τ^*-action, where τ^* acts on $C^\times_\Lambda$ by inversion. Consider the initial segment of the associated long exact sequence in the cohomology (see [Serre 1994, Proposition I.38]):

$$1 \longrightarrow \mathrm{SL}_\Lambda(C)^{\tau^*} \xrightarrow{j} \mathrm{GL}_\Lambda(C)^{\tau^*} \xrightarrow{\mathrm{nrd}} C^\times_\Lambda \cap C^{(2)} \longrightarrow \cdots .$$

It follows from the remark on page 379 that $\mathrm{SL}_\Lambda(C)^{\tau^*} \xrightarrow{j} \mathrm{GL}_\Lambda(C)^{\tau^*}$ is bijective. Thus the long exact sequence takes the form

$$1 \longrightarrow C_\Lambda^\times \cap C^{(2)} \xrightarrow{\delta} H^1(\tau^*, \mathrm{SL}_\Lambda(C)) \longrightarrow H^1(\tau^*, \mathrm{GL}_\Lambda(C)) \longrightarrow H^1(\tau^*, C_\Lambda^\times).$$

It is easy to see that $H^1(\tau^*, C_\Lambda^\times) = C_\Lambda^\times/(C_\Lambda^\times)^2$, which is a subgroup of $C^\times/(C_\Lambda^\times)^2$. Hence we simply replace the last term by $C^\times/(C_\Lambda^\times)^2$. This yields the upper row of the diagram. It is an easy exercise to verify that the lower row is an exact sequence of groups.

It remains to verify the commutativity of the rectangles. The middle one is obviously commutative by definition of the pfaffian in the cohomology. For the last rectangle we simply use that $\mathrm{pf}_\tau(g)^2 = \mathrm{nrd}(g)$ for all $g \in Z^1(\tau^*, \mathrm{GL}_\Lambda(C))$ by the construction of the pfaffian.

Consider the first rectangle. We recall the definition of the connecting morphism δ: given $c \in C_\Lambda^\times \cap C^{(2)}$, we can find an element $g \in \mathrm{GL}_\Lambda(C)$ such that $\mathrm{nrd}_{\Lambda/R}(g) = c$; then $\delta(c)$ is defined to be the class of $g^{-1}\,{}^{\tau^*}\!g$. The pfaffian of $g^{-1}\,{}^{\tau^*}\!g$ is

$$\mathrm{pf}_\tau(g^{-1}\,{}^{\tau^*}\!g) = \mathrm{nrd}(g)^{-1} = c^{-1} = c$$

(see Lemma 2.7). This proves the commutativity of the first rectangle.

Finally, note that δ is injective since $\mathrm{pf}_\tau \circ \delta$ is injective. $\square$

Corollary 2.9. *An element $x \in H^1(\tau^*, \mathrm{GL}_\Lambda(C))$ lies in the image of j_* if and only if $\mathrm{pf}_\tau(x)$ lies in the image of the canonical map $C^{(2)} \to C^\times/C_\Lambda^\times$.*

Proof. Let $\alpha : C^{(2)} \to C^\times/C_\Lambda^\times$ denote the canonical map. Suppose the class $x \in H^1(\tau^*, \mathrm{GL}_\Lambda(C))$ is in the image of j_*, then we obtain immediately that $\mathrm{pf}_\tau(x)$ lies in the image of α.

Conversely, suppose $\mathrm{pf}_\tau(x) = \alpha(u)$ for some $u \in C^{(2)}$. Then the diagram shows that $\mathrm{nrd}_{\Lambda/R}(x)$ is 1 in $C^\times/(C_\Lambda^\times)^2$ and therefore x lies in the image of j_*. $\square$

Remark (twisting involutions). Let A be a central simple k-algebra with an involution τ of symplectic type and let Λ be a τ-stable R-order. Given an element $b \in \mathrm{Sym}(\Lambda, \tau) \cap \Lambda^\times$, we can twist the involution τ with b. More precisely, we define $\tau|b : A \to A$ by $x \mapsto b\,{}^\tau\!x b^{-1}$. It is easily verified that this is again an involution on A, and since $b \in \Lambda^\times$, the order Λ is $\tau|b$-stable. Note that $\tau|b$ is again an involution of symplectic type.

Suppose Λ is τ-smooth, we claim that Λ is $\tau|b$-smooth as well. Take some element y in $\mathrm{Sym}(\Lambda, \tau|b)$; this is $y = b\,{}^\tau\!y b^{-1}$. Consequently, $yb \in \mathrm{Sym}(\Lambda, \tau)$ and by τ-smoothness there is an element $z \in \Lambda$ which satisfies ${}^\tau z + z = yb$. The element b is a unit in Λ, hence we may write $z = wb$ for $w = zb^{-1} \in \Lambda$ and it follows that ${}^{\tau|b}w + w = y$. We have shown that Λ is $\tau|b$-smooth.

Finally, for all $b \in \mathrm{Sym}(\Lambda, \tau) \cap \Lambda^\times$ we have $(\tau|b)^* = \mathrm{int}(b) \circ \tau^*$ on the group scheme GL_Λ. Since $b = {}^\tau b$ is equivalent to $b\,{}^{\tau^*}\!b = 1$, such an element b is a cocycle for $H^1(\tau^*, \Lambda^\times)$. If we now twist τ^* with the cocycle b (see Section 4), we obtain

$$\tau^*|b := \mathrm{int}(b) \circ \tau^* = (\tau|b)^*.$$

2E. ***Hermitian forms and nonabelian Galois cohomology.*** We shall also need a result due to Fainsilber and Morales from the theory of hermitian forms. Let A be a central simple k-algebra and let τ be an involution on A. In this short section it is not important whether or not τ is of the first or of the second kind.

The notion of τ-smoothness is related to the theory of even hermitian forms. Let Λ be a τ-stable R-order in A and let M be a finitely generated and projective right Λ-module. A hermitian form h (or more precisely a 1-hermitian form) with respect to τ on M is said to be *even* if there is a τ-sesquilinear form $s : M \times M \to \Lambda$ such that $h = s + s^*$. Here s^* is the sesquilinear form defined by

$$s^*(x, y) := {}^{\tau}s(y, x).$$

It follows immediately that Λ is τ-smooth if and only if every hermitian form on Λ (considered as a right Λ-module) is even. This is useful since even hermitian forms can be handled more easily than arbitrary hermitian forms.

We consider the automorphism τ^* of $\Lambda^\times$ defined as the composition of τ and the group inversion. Similarly we obtain τ^* on $A^\times$. Here it is not necessary to consider τ^* as a morphism of group schemes, which is a little bit more tedious if τ is of the second kind. We will need a theorem from [Fainsilber and Morales 1999] in the following paraphrase:

Theorem 2.10. *Let k be a field which is complete for a discrete valuation and let R be its valuation ring. Let A be a central simple k-algebra with involution τ. Suppose Λ is a τ-stable maximal R-order in A. If Λ is τ-smooth, the canonical map*

$$j_* : H^1(\tau^*, \Lambda^\times) \to H^1(\tau^*, A^\times)$$

is injective.

Compared with [Fainsilber and Morales 1999] we have added the assumption of τ-smoothness to eliminate the restriction on the residual characteristic. The proof is almost identical.

3. An adelic reformulation of Harder's Gauss–Bonnet theorem

We briefly describe an adelic reformulation of Harder's Gauss–Bonnet theorem [Harder 1971] that hinges on the notion of a smooth group scheme. In fact, the Euler characteristic of an arithmetic group can also be computed using G. Prasad's [1989] general volume formula. Since we have explicit underlying smooth integral models of the algebraic groups, we think that the adelic volume formula derived in this section is adapted much better to the applications given in this article.

Let F be an algebraic number field and let $\mathbb{O}$ denote its ring of integers. Let G be a connected semisimple algebraic group defined over F. We denote by G_∞ the associated real semisimple Lie group

$$G_\infty = G(F \otimes_{\mathbb{Q}} \mathbb{R}) = \prod_{v \in V_\infty} G(F_v).$$

3A. ***The Euler–Poincaré measure.*** We define what we mean by the *compact dual group* of G_∞, since the definition differs from author to author. Let $\mathfrak{g}_\infty$ be the real Lie algebra of G_∞ and let $\mathfrak{g}_{\infty,\mathbb{C}}$ denote its complexification. Moreover, let K_∞ be a maximal compact subgroup of G_∞ and consider the associated Cartan decomposition

$$\mathfrak{g}_\infty = \mathfrak{k}_\infty \oplus \mathfrak{p}.$$

The real vector space $\mathfrak{u} := \mathfrak{k}_\infty \oplus i\mathfrak{p} \subseteq \mathfrak{g}_{\infty,\mathbb{C}}$ is a real Lie subalgebra of $\mathfrak{g}_{\infty,\mathbb{C}}$ and is even a compact real form of $\mathfrak{g}_{\infty,\mathbb{C}}$ [Knapp 2002, page 360]. Let G_u be the unique connected (a priori virtual) Lie subgroup of $G(F \otimes_{\mathbb{Q}} \mathbb{C})$ with Lie algebra $\mathfrak{u}$. Since the real semisimple Lie algebra $\mathfrak{u}$ is a compact form, the Lie group G_u is compact and thus closed in $G(F \otimes_{\mathbb{Q}} \mathbb{C})$ [Knapp 2002, Chapter IV, Theorem 4.69]. Further we see that the connected component K_∞^0 is a subgroup of G_u. We say that G_u is the *compact dual group* of G_∞ containing K_∞^0. Note that the dual group depends on the algebraic group G.

Let $B : \mathfrak{g}_\infty \times \mathfrak{g}_\infty \to \mathbb{R}$ be a nondegenerate $\mathbb{R}$-bilinear form such that $\mathfrak{k}_\infty$ and $\mathfrak{p}$ are orthogonal. We extend B to a $\mathbb{C}$-bilinear form (again denoted by B) on $\mathfrak{g}_{\infty,\mathbb{C}}$. Note that B restricted to $\mathfrak{u}$ is a nondegenerate $\mathbb{R}$-bilinear form $\mathfrak{u}\times\mathfrak{u} \to \mathbb{R}$. We obtain corresponding right-invariant volume densities on G_∞ and on G_u which will be denoted by $|\mathrm{vol}_B|$.

We define $X := K_\infty \backslash G_\infty$. Let $\Gamma \subseteq G(F)$ be a torsion-free arithmetic group. Harder's Gauss–Bonnet theorem shows that integration over G_∞/Γ with the Euler–Poincaré measure μ_χ [Serre 1971, §3] yields the Euler characteristic of Γ — even if Γ is not cocompact. Via Hirzebruch's proportionality principle one has the following formula for the Euler–Poincaré measure on G_∞ [Harder 1971; Serre 1971].

Theorem 3.1. *If* $\dim(X)$ *is odd or if* $\mathrm{rk}(\mathfrak{k}_{\infty,\mathbb{C}}) < \mathrm{rk}(\mathfrak{g}_{\infty,\mathbb{C}})$, *then* $\mu_\chi = 0$ *is the Euler–Poincaré measure. Otherwise, if* $\mathrm{rk}(\mathfrak{g}_{\infty,\mathbb{C}}) = \mathrm{rk}(\mathfrak{k}_{\infty,\mathbb{C}})$ *and* $\dim(X) = 2p$ *is even, then*

$$\mu_\chi := \frac{(-1)^p |W(\mathfrak{g}_{\infty,\mathbb{C}})|}{|\pi_0(G_\infty)|\,|W(\mathfrak{k}_{\infty,\mathbb{C}})|}\ \mathrm{vol}_B(G_u)^{-1} |\mathrm{vol}_B|.$$

Here $\pi_0(G_\infty) = G_\infty/G_\infty^0$ *and* $W(\mathfrak{g}_{\infty,\mathbb{C}})$, $W(\mathfrak{k}_{\infty,\mathbb{C}})$ *denote the Weyl groups of the complexified Lie algebras* $\mathfrak{g}_{\infty,\mathbb{C}}$, $\mathfrak{k}_{\infty,\mathbb{C}}$.

3B. *The adelic reformulation.* Let $\mathbb{A}$ denote the ring of adeles of F and $\mathbb{A}_f$ the ring of finite adeles. Let G be a connected semisimple algebraic group defined over F. Let $K_f \subseteq G(\mathbb{A}_f)$ be an open compact subgroup of the locally compact group $G(\mathbb{A}_f)$. Borel showed that $G(\mathbb{A})$ is the disjoint union of a finite number m of double cosets, that is,

$$G(\mathbb{A}) = \bigsqcup_{i=1}^{m} G_\infty K_f x_i G(F)$$

for some representatives $x_1, \dots, x_m \in G(\mathbb{A}_f)$ [Borel 1963, Theorem 5.1]. For every $i = 1, \dots, m$ we obtain an arithmetic subgroup $\Gamma_i \subseteq G(F)$ defined by

$$\Gamma_i := G(F) \cap x_i^{-1} K_f x_i.$$

There is a G_∞-equivariant homeomorphism

$$K_f \backslash G(\mathbb{A})/G(F) \xrightarrow{\simeq} \bigsqcup_{i=1}^{m} G_\infty/\Gamma_i. \tag{1}$$

Here the right-hand side denotes the topologically disjoint union.

Remark. Define $X = K_\infty \backslash G_\infty$. Suppose $G(F)$ acts freely on $K_\infty K_f \backslash G(\mathbb{A})$. This is the case if and only if the groups Γ_i are torsion-free for all $i = 1, \dots, m$. If $\dim(X)$ is odd or if $\mathrm{rk}(\mathfrak{k}_{\infty,\mathbb{C}}) < \mathrm{rk}(\mathfrak{g}_{\infty,\mathbb{C}})$, then

$$\chi(K_\infty K_f \backslash G(\mathbb{A})/G(F)) = 0.$$

This follows immediately from Harder's Gauss–Bonnet theorem and the homeomorphism in (1).

Note further that if F has a complex place, then $\mathrm{rk}(\mathfrak{k}_{\infty,\mathbb{C}}) < \mathrm{rk}(\mathfrak{g}_{\infty,\mathbb{C}})$ is always satisfied. Therefore we may restrict to the case where F is totally real.

3B1. *The Tamagawa measure.* We derive a description of the Tamagawa measure in terms of the local volume densities. For a thorough definition of the Tamagawa measure we refer the reader to [Oesterlé 1984]. Let G be a connected semisimple linear algebraic F-group of dimension d. Let $\mathfrak{g} = \mathrm{Lie}(G)(F)$ be the Lie algebra of G over F.

Fix a nondegenerate F-bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to F$ on the Lie algebra. For every place $v \in V$ we have the left invariant volume density $|\mathrm{vol}_B|_v$ attached to B on the F_v-analytic manifold $G(F_v)$. The volume density is uniquely determined by $|\mathrm{vol}_B|(e_1 \wedge \dots \wedge e_d) = |\det(B(e_i, e_j))|^{1/2}$ for all $e_1, \dots, e_d \in \mathfrak{g}$. We fix Haar measures μ_v on F_v for every place v such that

(i) $\mu_v(\mathbb{O}_v) = 1$ if $v \in V_f$ is a finite place,

(ii) $\mu_v([0, 1]) = 1$ if v is a real place, and

(iii) $\mu_v([0, 1] + [0, 1]i) = 2$ if v is a complex place.

Using these choices of Haar measures, a density on $G(F_v)$ defines a measure on the analytic manifold $G(F_v)$.

Lemma 3.2. *Let G be a d-dimensional semisimple connected linear algebraic group defined over F. Fix a nondegenerate F-bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to F$ on the Lie algebra. Then the Tamagawa measure on $G(\mathbb{A})$ is given by*

$$\tau = |d_F|^{-d/2} \prod_{v \in V} |\mathrm{vol}_B|_v.$$

Proof. Let $e_1, \dots, e_d$ be a basis of $\mathfrak{g}$ over F and take the dual basis $\varepsilon_1, \dots, \varepsilon_d$ of $\mathrm{Hom}_F(\mathfrak{g}, F)$. We define a nontrivial form of highest degree $\omega = \varepsilon_1 \wedge \cdots \wedge \varepsilon_d$ on $\mathfrak{g}$. By definition of the volume density we have

$$|\mathrm{vol}_B|_v = |\det(B(e_i, e_j))|_v^{1/2} |\omega|_v.$$

By the product formula we know that $|\det(B(e_i, e_j))|_v = 1$ for almost all places v and further that $\prod_{v \in V} |\det(B(e_i, e_j))|_v = 1$. $\square$

3B2. *The modulus factor.* We focus on the case where the algebraic group has a smooth $\mathbb{O}$-model. Let G be a smooth group scheme defined over $\mathbb{O}$. For any commutative $\mathbb{O}$-algebra R we write $\mathfrak{g}_R := \mathrm{Lie}(G)(R)$ to denote the R-points of the Lie algebra of G. Let $B : \mathfrak{g}_F \times \mathfrak{g}_F \to F$ be a nondegenerate F-bilinear form. For every finite place $v \in V_f$ we define the *modulus factor* $m(B)_v$ as follows: take an $\mathbb{O}_v$-basis $e_1, \dots, e_n$ of the free $\mathbb{O}_v$-module $\mathfrak{g}_v := \mathfrak{g}_{\mathbb{O}_v}$, and define

$$m(B)_v := \left|\det(B(e_i, e_j))\right|_v^{1/2}.$$

For almost all finite places $v \in V_f$ we have $m(B)_v = 1$. To see this, take an F-basis of $\mathfrak{g}_F$ and note that it is an $\mathbb{O}_v$-basis of $\mathfrak{g}_v$ for almost all finite places v. This allows us to define the *global modulus factor* $m(B) := \prod_{v \in V_f} m(B)_v$.

3B3. *Congruence groups.* In the adelic formulation of Harder's Gauss–Bonnet theorem we focus on congruence groups which are given by local data. Let G be a smooth $\mathbb{O}$-group scheme. For every finite place $v \in V_f$, let $\alpha_v \geq 1$ be a natural number and we assume that $\alpha_v = 1$ for almost all $v \in V_f$. Let v be a finite place and let $\mathfrak{p}_v \subseteq \mathbb{O}_v$ be the unique prime ideal in $\mathbb{O}_v$. We define π_v to be the reduction morphism

$$\pi_v : G(\mathbb{O}_v) \to G(\mathbb{O}_v/\mathfrak{p}_v^{\alpha_v}).$$

Further, we assume that we are given a subgroup U_v of the finite group $G(\mathbb{O}_v/\mathfrak{p}_v^{\alpha_v})$ for every place $v \in V_f$. For a place $v \in V_f$ the group $K_v(U) := \pi_v^{-1}(U_v)$ is an open compact subgroup of $G(\mathbb{O}_v)$. If we additionally impose the assumption that

$U_v = G(\mathcal{O}_v/\mathfrak{p}_v^{\alpha_v})$ for almost all v, then the group

$$K(U) := \prod_{v \in V_f} K_v(U)$$

is an open compact subgroup of the locally compact group $G(\mathbb{A}_f)$. We say that $K(U)$ is the congruence group associated with the *local datum*

$$U = (U_v)_v = (U_v, \alpha_v)_v$$

(usually the numbers α_v are considered to be implicitly a part of the datum U).

3B4. *The adelic Euler characteristic formula.* Let F be a totally real number field. Let G be a smooth group scheme over $\mathcal{O}$ such that $G \times_{\mathcal{O}} F$ is a connected semisimple group. For every real place v we choose a maximal compact subgroup $K_v \subseteq G(F_v)$. The real Lie algebra of K_v will be denoted $\mathfrak{k}_v$. The product $K_\infty = \prod_{v \in V_\infty} K_v$ is a maximal compact subgroup of the associated real Lie group G_∞. We denote the Lie algebra of K_∞ by $\mathfrak{k}$.

Let $B : \mathfrak{g}_F \times \mathfrak{g}_F \to F$ be an F-bilinear form. We say that B is *nice* with respect to K_∞ if B is nondegenerate and for every real place $v \in V_\infty$ the Cartan decomposition with respect to $\mathfrak{k}_v$ is orthogonal with respect to B. A nice form induces a nondegenerate bilinear form $B : \mathfrak{g}_\infty \times \mathfrak{g}_\infty \to \mathbb{R}$ by defining the Lie subalgebras $\mathfrak{g}_v = \mathrm{Lie}(G(F_v))$ to be orthogonal. Note that the form B satisfies the requirements of Theorem 3.1.

Theorem 3.3. *Let G be a smooth group scheme over $\mathcal{O}$ such that $G \times_{\mathcal{O}} F$ is a connected semisimple group of dimension d. We fix any nice form $B : \mathfrak{g}_F \times \mathfrak{g}_F \to F$. Furthermore, let $K_f = K(U)$ be a congruence subgroup of $G(\mathbb{A}_f)$ given by a local datum (U, α) such that $G(F)$ acts freely on $K_\infty K_f \backslash G(\mathbb{A})$.*

If $\dim(X) = 2p$ *is even and* $\mathrm{rk}(\mathfrak{k}_\mathbb{C}) = \mathrm{rk}(\mathfrak{g}_{\infty,\mathbb{C}})$, *then the Euler characteristic of the double coset space $K_\infty K_f \backslash G(\mathbb{A})/G(F)$ is given by*

$$\chi(K_\infty K_f \backslash G(\mathbb{A})/G(F)) = (-1)^p |d_F|^{d/2} \frac{|W(\mathfrak{g}_{\infty,\mathbb{C}})|\tau(G)}{|\pi_0(G_\infty)||W(\mathfrak{k}_\mathbb{C})|} \mathrm{vol}_B(G_u)^{-1} m(B)^{-1} \prod_{v \in V_f} \frac{\mathrm{N}(\mathfrak{p}_v)^{d\alpha_v}}{|U_v|}.$$

Here $\tau(G)$ is the Tamagawa number of G, $\mathrm{N}(\mathfrak{p}_v)$ denotes the cardinality of the residue class field $\mathcal{O}_v/\mathfrak{p}_v$, and G_u denotes the compact dual group of G_∞^0 (remaining notation is as in Theorem 3.1).

Proof. Let $x_1, \dots, x_m \in G(\mathbb{A}_f)$ be a collection of representatives of the finitely many elements of $G_\infty K_f \backslash G(\mathbb{A})/G(F)$. We consider the torsion-free arithmetic groups Γ_i defined as $\Gamma_i := G(F) \cap x_i^{-1} K_f x_i$. Let $\mathcal{F}_i$ be a Borel measurable fundamental domain for the right action of Γ_i on G_∞. Here we mean a fundamental domain in the strict sense, that is, $\mathcal{F}_i$ is a set of representatives for G_∞/Γ_i (for the existence of measurable fundamental domains see [Bourbaki 1963, Chapter VII §2, Example 12]). The set $\mathcal{F}$ defined as the union $\bigsqcup_{i=1}^{m} \mathcal{F}_i K_f x_i \subseteq G(\mathbb{A})$ is a Borel measurable fundamental domain for the right action of $G(F)$ on $G(\mathbb{A})$. We write $|\mathrm{vol}_B|_\infty = \prod_{v \in V_\infty} |\mathrm{vol}_B|_v$ and further $|\mathrm{vol}_B|_f := \prod_{v \in V_f} |\mathrm{vol}_B|_v$. Due to Theorem 3.1 we have

$$\chi(K_\infty K_f \backslash G(\mathbb{A})/G(F)) = \sum_{i=1}^{m} \chi(X/\Gamma_i) = \lambda \sum_{i=1}^{m} \int_{\mathcal{F}_i} |\mathrm{vol}_B|_\infty,$$

where

$$\lambda = (-1)^p \frac{|W(\mathfrak{g}_{\infty,\mathbb{C}})|}{|\pi_0(G_\infty)|\,|W(\mathfrak{k}_\mathbb{C})|} \mathrm{vol}_B(G_u)^{-1}.$$

By multiplication with the volume of K_f, which is simply $\mathrm{vol}_B(K_f) = \int_{K_f} |\mathrm{vol}_B|_f$, and by Lemma 3.2, we obtain

$$\sum_{i=1}^{m} \int_{\mathcal{F}_i} |\mathrm{vol}_B|_\infty \mathrm{vol}_B(K_f) = \int_{\mathcal{F}} \prod_{v \in V} |\mathrm{vol}_B|_v = |d_F|^{d/2} \int_{\mathcal{F}} \tau = |d_F|^{d/2} \tau(G).$$

This means we have

$$\chi(K_\infty K_f \backslash G(\mathbb{A})/G(F)) = \lambda |d_F|^{d/2} \tau(G) \mathrm{vol}_B(K_f)^{-1}.$$

Finally we are left with the task of determining $\mathrm{vol}_B(K_f)$. We shall exploit that K_f is given by the local datum (U, α). Since $\mathrm{vol}_B(K_f) = \prod_{v \in V_f} \mathrm{vol}_B(K_v(U))$ and the scheme G is smooth, we can apply a theorem of Weil (for a modern formulation see [Oesterlé 1984, Section I.2.5] or [Batyrev 1999, Theorem 2.5]) in every finite place to deduce

$$\mathrm{vol}_B(K_f) = \prod_{v \in V_f} m(B)_v \frac{|U_v|}{\mathrm{N}(\mathfrak{p}_v)^{d\alpha_v}}.$$

Now the claim follows readily. □

4. Rohlfs' method

In this section we give a short summary of Rohlfs' method for the computation of Lefschetz numbers.

Let F be an algebraic number field and let G be a linear algebraic group defined over F. We assume that G has strong approximation. For example, unipotent groups and F-simple, simply connected groups with a noncompact associated Lie group have strong approximation [Platonov and Rapinchuk 1994, page 427]. Choose a maximal compact subgroup $K_\infty \subseteq G_\infty$ and set $X := K_\infty \backslash G_\infty$. Furthermore, let $K_f \subseteq G(\mathbb{A}_f)$ be an open compact subgroup and let $\Gamma := G(F) \cap K_f$ be the arithmetic group defined by this open compact subgroup. There is a homeomorphism

$$X/\Gamma \xrightarrow{\simeq} K_\infty K_f \backslash G(\mathbb{A})/G(F).$$

To see this, consider the inclusion $G_\infty \to G(\mathbb{A})$ and use strong approximation to observe that it factors to such a homeomorphism. Recall that Γ is torsion-free if and only if $G(F)$ acts freely on $K_\infty K_f \backslash G(\mathbb{A})$.

Let τ be an automorphism of finite order of G. We can choose K_∞ such that it is τ-stable. We further assume that $K_f \subset G(\mathbb{A}_f)$ is a τ-stable open compact subgroup. We obtain an action of τ on the double coset space

$$S(K_f) := K_\infty K_f \backslash G(\mathbb{A})/G(F).$$

We describe the set $S(K_f)^\tau$ of τ-fixed points following [Rohlfs 1990] under the assumption that $G(F)$ acts freely on $K_\infty K_f \backslash G(\mathbb{A})$.

Consider the finite set $\mathcal{H}^1(\tau)$ defined as the fibred product

$$\mathcal{H}^1(\tau) := H^1(\tau, K_\infty K_f) \underset{H^1(\tau, G(\mathbb{A}))}{\times} H^1(\tau, G(F))$$

of nonabelian cohomology sets. Here we usually write τ instead of the finite group $\langle\tau\rangle$ generated by τ. We consider $\mathcal{H}^1(\tau)$ as a topological space with the discrete topology. Rohlfs [1990, Section 3.5] constructed a surjective and continuous map

$$\vartheta : S(K_f)^\tau \to \mathcal{H}^1(\tau).$$

In particular the fibres are open and closed in $S(K_f)^\tau$ and we get a decomposition

$$S(K_f)^\tau = \bigsqcup_{\eta \in \mathcal{H}^1(\tau)} \vartheta^{-1}(\eta). \tag{2}$$

Let $\gamma \in Z^1(\tau, G(F))$ be a cocycle. The γ-twisted τ-action on G, defined by ${}^{\tau|\gamma}(x) = \gamma_\tau\, {}^\tau x \gamma_\tau^{-1}$, is an automorphism defined over F and the group of fixed points is a linear algebraic group which will be denoted $G(\gamma)$. Similarly, given a cocycle $k \in Z^1(\tau, K_\infty K_f)$ we define the k-twisted action of τ on $K_\infty K_f$ by ${}^{\tau|k}g := k_\tau\, {}^\tau g k_\tau^{-1}$. The corresponding group of fixed points under this action will be written $(K_\infty K_f)^{\tau|k}$. Rohlfs obtained the following description of the fibres of ϑ.

Lemma 4.1 [Rohlfs 1990, Section 3.5]. *Let $K_f \subset G(\mathbb{A}_f)$ be a τ-stable open compact subgroup such that $G(F)$ acts freely on $K_\infty K_f \backslash G(\mathbb{A})$.*

Let $\eta \in \mathcal{H}^1(\tau)$ be a class represented by a pair of cocycles (k, γ) with $(k_s)_s$ in $Z^1(\tau, K_\infty K_f)$ and $(\gamma_s)_s \in Z^1(\tau, G(F))$. Take $a \in G(\mathbb{A})$ such that ${}^s a = k_s^{-1} a \gamma_s$ for all $s \in \langle\tau\rangle$. There is a homeomorphism

$$a^{-1}(K_\infty K_f)^{\tau|k} a \backslash G(\gamma)(\mathbb{A})/G(\gamma)(F) \xrightarrow{\simeq} \vartheta^{-1}(\eta).$$

Combined with Theorem 4.2 below this yields a method for the computation of Lefschetz numbers which we simply call *Rohlfs' method*.

Definition. Let $\rho : G \to \mathrm{GL}(W)$ be a rational representation defined over the algebraic closure $\overline{F}$ of F. Here W is a finite dimensional $\overline{F}$-vector space. Given an action of the finite group $\langle\tau\rangle$ on W, we say that this action is *compatible* with ρ if

$$^{s}(\rho(g)v) = \rho(^{s}g)\,^{s}v$$

for all $v \in V$, $s \in \langle\tau\rangle$ and $g \in G(\overline{F})$. In other words W is a $(G(\overline{F}) \rtimes \langle\tau\rangle)$-module.

Let $\rho : G \to \mathrm{GL}(W)$ be a rational representation and let $\Gamma \subseteq G(F)$ be a torsion-free arithmetic subgroup. If W is equipped with a compatible τ-action then we define the *Lefschetz number* of τ with values in W as

$$\mathscr{L}(\tau,\Gamma,W) := \sum_{i=0}^{\infty}(-1)^i \operatorname{Tr}\bigl(\tau^i : H^i(\Gamma,W) \to H^i(\Gamma,W)\bigr).$$

Since torsion-free arithmetic groups are of type (FL), this is a finite sum.

Given a cocycle $b = (b_s)_s \in H^1(\tau, G(\overline{F}))$ one can define the b-twisted τ-action on W by

$$^{\tau|b}w = b_\tau\,^{\tau}w$$

for all $w \in W$. We write $W(b)$ to denote the space W with the b-twisted τ-action. We need the following slight paraphrase of a theorem of Rohlfs.

Theorem 4.2 [Rohlfs 1990]. *Let G be an algebraic F-group with strong approximation and let τ be an automorphism of finite order defined over F. Let $K_f \subset G(\mathbb{A}_f)$ be a τ-stable open compact subgroup such that $\Gamma := G(F) \cap K_f$ is torsion-free. Let $\rho : G \to \mathrm{GL}(W)$ be a rational representation defined over $\overline{F}$ with a compatible τ-action. Then we have*

$$\mathscr{L}(\tau,\Gamma,W) = \sum_{\eta \in \mathscr{H}^1(\tau)} \chi(\vartheta^{-1}(\eta)) \operatorname{Tr}(\tau|W(b_\eta)),$$

where $b_\eta \in G(F)$ is any representative of the $H^1(\tau, G(F))$ component of η.

Proof. This follows from Rohlfs' decomposition — see (2) — and a suitable Lefschetz fixed point principle — for instance [Rohlfs and Schwermer 1998, §2.3] or [Kionke 2012]. □

5. Proof of the main theorem

5A. ***Introduction.*** In this section we compute the Lefschetz number of an involution of symplectic type on principal congruence subgroups of inner forms of the special linear group. For this purpose we combine the tools developed in the previous sections.

One should keep in mind that the central result is the adelic Lefschetz number formula in Theorem 4.2. Whenever we want to apply this theorem, there are two important steps to do. First step: understand the involved first nonabelian Galois cohomology sets. Second step: compute the Euler characteristics of the fixed point components. In the second step we use the adelic formula in Theorem 3.3 obtained from Harder's Gauss–Bonnet theorem.

First we introduce some notation, then we begin to determine various nonabelian cohomology sets. In the third subsection we describe the fixed point groups and we compute their Euler characteristics. Finally we prove the main theorem.

As before F denotes an algebraic number field and $\mathcal{O}$ denotes its ring of integers. Let D be a quaternion algebra over F, that is, a central simple F-algebra of dimension four. Note that even though we use the symbol D, the quaternion algebra D is in general not assumed to be a division algebra. Given a place v, we define $D_v := D \otimes_F F_v$. If D_v is isomorphic to $M_2(F_v)$, we say that D *splits* at the place v. Otherwise D_v is a division algebra and we say that D is *ramified* at v. Let $\mathrm{Ram}(D) \subset V$ be the finite set of places where D ramifies, and let $\mathrm{Ram}_f(D)$ (resp. $\mathrm{Ram}_\infty(D)$) denote the subset of finite (resp. Archimedean) places.

Definition. The *signed reduced discriminant* $\Delta_{\mathrm{rd}}(D)$ of D is the integer

$$\Delta_{\mathrm{rd}}(D) := (-1)^r \prod_{\mathfrak{p} \in \mathrm{Ram}_f(D)} \mathrm{N}(\mathfrak{p}),$$

where $r = |\mathrm{Ram}_\infty(D)|$.

5A1. *The canonical involution.* On the quaternion algebra D we have the *canonical involution*

$$\tau_c : D \to D, \quad \tau_c(x) =: \bar{x},$$

sometimes called *conjugation*. Given a description $D = Q(a, b|F)$ of D with $a, b \in F^\times$ — meaning there is a basis $1, i, j, ij$ of D with $i^2 = a$, $j^2 = b$ and $ij = -ji$ — conjugation is defined by

$$\tau_c : x_0 + x_1 i + x_2 j + x_3 ij \mapsto x_0 - x_1 i - x_2 j - x_3 ij.$$

Note that the conjugation is F-linear; that is, it is an involution of the first kind on D. Moreover, τ_c is an involution of symplectic type.

The elements fixed by conjugation are precisely the elements of F. The reduced norm and trace of D are related to conjugation by

$$\mathrm{trd}_D(x) = x + \bar{x}, \quad \mathrm{nrd}_D(x) = x\bar{x} = \bar{x}x$$

for all $x \in D$.

5A2. *Orders.* Let Λ_D be an $\mathbb{O}$-order in D. We show that Λ_D is τ_c-stable: let $x \in \Lambda_D$, then

$$\bar{x} = x + \bar{x} - x = \mathrm{trd}_D(x) - x.$$

Recall that $\mathrm{trd}_D(x) \in \mathbb{O}$ because x is integral. Since $\mathbb{O} \subseteq \Lambda_D$, we obtain $\bar{x} \in \Lambda_D$. Moreover, it follows directly from the definitions that Λ_D is smooth if and only if Λ_D is τ_c-smooth (see the second definition on page 376, and the one on page 378).

We will assume from now on that Λ_D is a maximal $\mathbb{O}$-order in D. In particular, it is a smooth and τ_c-smooth order (see Proposition 2.4).

Let n be a positive integer. Consider the central simple F-algebra

$$A := M_n(D)$$

of $n \times n$-matrices with entries in the quaternion algebra D. The canonical involution on D induces an involution τ on A defined by

$$\tau(x) := {}^{\tau}x := \bar{x}^T;$$

that is, conjugate every entry in the matrix x and then transpose the matrix. It is easily checked that this defines an involution of symplectic type on A [Knus et al. 1998, Proposition 2.23].

Lemma 5.1. *Let $\Lambda_D \subseteq D$ be a maximal $\mathbb{O}$-order. The $\mathbb{O}$-order $\Lambda = M_n(\Lambda_D)$ in A is maximal, τ-stable, smooth and τ-smooth.*

Proof. Since Λ_D is stable under conjugation, it is obvious that Λ is τ-stable. Moreover, it follows from [Reiner 2003, Theorem (21.6)] that Λ is a maximal $\mathbb{O}$-order. In turn Proposition 2.4 shows that Λ is also a smooth order.

Finally we need to check that Λ is τ-smooth. Let $x \in \mathrm{Sym}(\Lambda, \tau)$ be an element which is fixed by τ. This means that $x = (x_{ij})$ satisfies

$$x_{ij} = \bar{x}_{ji} \quad \text{for all } i \neq j$$

and

$$x_{ii} \in \mathbb{O}.$$

The order Λ_D is smooth, therefore there is, for every $i = 1, \ldots, n$, an element $z_i \in \Lambda_D$ with $\mathrm{trd}_D(z_i) = z_i + \bar{z}_i = x_{ii}$. Now we define the upper triangular element $y \in \Lambda$ by

$$y_{ij} := \begin{cases} 0 & \text{if } i > j, \\ x_{ij} & \text{if } i < j, \\ z_i & \text{if } i = j, \end{cases}$$

and it is easy to see that $y + {}^{\tau}y = x$. We deduce that Λ is τ-smooth. $\square$

5A3. *Setting and assumptions.* We define $G := \mathrm{SL}_\Lambda$ to be the special linear group over the order Λ (see the first definition on page 376). From the previous lemma and Corollary 2.3 we deduce that G is a smooth group scheme over $\mathbb{O}$. Moreover, the involution τ induces an automorphism τ^* of GL_Λ where $\tau^* = \mathrm{inv} \circ \tau$ (see Section 2C). Clearly τ^* has order (at most) two and restricts to an automorphism of $G = \mathrm{SL}_\Lambda$.

The real Lie group G_∞ associated with G is

$$G_\infty := \prod_{v \in V_\infty} G(F_v) \cong \mathrm{SL}_{2n}(\mathbb{R})^s \times \mathrm{SL}_n(\mathbb{H})^r \times \mathrm{SL}_{2n}(\mathbb{C})^t.$$

Here s denotes the number real places of F where D splits, r is the number of real places where D ramifies, and t is the number of complex places of F. The symbol $\mathbb{H}$ is used for Hamilton's quaternion division algebra and $\mathrm{SL}_n(\mathbb{H})$ is the group of elements with reduced norm one in the central simple $\mathbb{R}$-algebra $M_n(\mathbb{H})$. Note that $[F : \mathbb{Q}] = r + s + 2t$. For every Archimedean place v we fix a τ^*-stable maximal compact subgroup $K_v \subseteq G(F_v)$; then the group $K_\infty := \prod_{v \in V_\infty} K_v$ is a τ^*-stable maximal compact subgroup of G_∞.

We study the cohomology of congruence subgroups arising from the group SL_Λ. Let $\mathfrak{a} \subseteq \mathbb{O}$ be a proper ideal; we define the principal congruence subgroup

$$\Gamma(\mathfrak{a}) := \ker(G(\mathbb{O}) \to G(\mathbb{O}/\mathfrak{a}))$$

of level $\mathfrak{a}$. We shall always assume that $\Gamma(\mathfrak{a})$ is torsion-free (which holds for almost all ideals). Note that the groups $\Gamma(\mathfrak{a})$ are always τ^*-stable.

These groups can be described by local data. Let $\mathfrak{p} \subseteq \mathbb{O}$ be a prime ideal of $\mathbb{O}$ and let v be the associated finite place. Let $\nu_\mathfrak{p}(\mathfrak{a})$ be the maximal exponent e such that $\mathfrak{p}^e$ divides $\mathfrak{a}$; then $\mathfrak{a}\mathbb{O}_v = \mathfrak{p}^{\nu_\mathfrak{p}(\mathfrak{a})}\mathbb{O}_v$. We obtain an open and compact subgroup $K_v \subseteq G(\mathbb{O}_v)$ defined as

$$K_v := \ker(G(\mathbb{O}_v) \longrightarrow G(\mathbb{O}_v/\mathfrak{a}\mathbb{O}_v)).$$

We form the direct product $K_f := \prod_{v \in V_f} K_v$, which is an open and compact subgroup of the locally compact group $G(\mathbb{A}_f)$. Clearly, $\Gamma(\mathfrak{a}) = G(F) \cap K_f$.

We keep the notation introduced in this section. We always assume that

(i) the order Λ_D is a maximal order in D, and

(ii) the ideal $\mathfrak{a} \subseteq \mathbb{O}$ is nontrivial and chosen such that $\Gamma(\mathfrak{a})$ is torsion-free.

5B. ***Hermitian forms and Galois cohomology.*** In this section we determine the nonabelian Galois cohomology set $\mathcal{H}^1(\tau^*)$. Recall that $\mathcal{H}^1(\tau^*)$ is the fibred product

$$\mathcal{H}^1(\tau^*) := H^1(\tau^*, K_\infty K_f) \underset{H^1(\tau^*, G(\mathbb{A}))}{\times} H^1(\tau^*, G(F)).$$

In order to determine this set we need to calculate local and global cohomology sets. The global problem is to determine $H^1(\tau^*, G(F))$, whereas locally we have to calculate $H^1(\tau^*, G(F_v))$ and $H^1(\tau^*, K_v)$ for every place v. We start by determining the corresponding cohomology sets for GL_Λ. This task amounts to the classification of certain hermitian forms over quaternion algebras, which is well known (see for instance [Shimura 1963, §2] or [Scharlau 1985, Chapter 10]). Afterwards we use the pfaffian to obtain results for the special linear group.

5B1. *Local results for* GL_Λ. We introduce the following notation: given two integers $p, q \geq 0$, we define the diagonal matrix

$$I_{p,q} = \mathrm{diag}(\underbrace{1, \ldots, 1}_{p}, \underbrace{-1, \ldots, -1}_{q}).$$

Proposition 5.2. *Let $v \in V$ be a place of F. If v is a real place where D is ramified, then*

$$H^1(\tau^*, \mathrm{GL}_\Lambda(F_v)) \cong \{I_{p,q} \mid p, q \geq 0 \textit{ with } p + q = n\}.$$

This means that the matrices $I_{p,q}$ are a system of representatives for the cohomology classes. The cohomology is trivial for all places $v \in V \setminus \mathrm{Ram}_\infty(D)$, that is,

$$H^1(\tau^*, \mathrm{GL}_\Lambda(F_v)) = \{1\}.$$

Proof. Let $b \in Z^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$ be a cocycle; b is an element of $\mathrm{GL}_n(D_v)$ satisfying $b = {}^\tau b$. Such a matrix b defines a regular hermitian form on the free right D_v-module D_v^n.

If $v \in V \setminus \mathrm{Ram}_\infty(D)$ (i.e., v is not a real ramified place), regular hermitian forms over D_v are classified by their dimension over F_v; this follows from [Scharlau 1985, Chapter 10, Theorem 1.7 and Example 1.8]. Note that these results cover the case where D_v is a division algebra. However it is easy to obtain an analogous result if $D_v \cong M_2(F_v)$ (at least for free regular hermitian spaces). Thus we find $g \in \mathrm{GL}_n(D_v)$ with $gb\,{}^\tau g = 1$, and so the second assertion follows immediately.

Let $v \in \mathrm{Ram}_\infty(D)$; then $D_v \cong \mathbb{H}$. In this case τ_c-hermitian forms are classified by dimension and signature. Translated to the setting of nonabelian Galois cohomology, this means that the set $\{I_{p,q} \mid p, q \geq 0 \text{ with } p + q = n\}$ is a system of representatives for $H^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$. □

Definition. Let $v \in \mathrm{Ram}_\infty(D)$. For a cocycle $b \in Z^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$ which is cohomologous to $I_{p,q}$ we say that the *signature* of b is the pair (p, q).

Corollary 5.3. *Let $v \in V_f$ be a finite place; then $H^1(\tau^*, \mathrm{GL}_\Lambda(\mathbb{O}_v)) = \{1\}$.*

Proof. The $\mathbb{O}$-order Λ is maximal and τ-smooth (see Lemma 5.1) and the same holds for the $\mathbb{O}_v$-order $\Lambda \otimes \mathbb{O}_v$ (see [Reiner 2003, Corollary (11.6)] and note that

τ-smoothness is a local property). By Theorem 2.10 (Fainsilber and Morales), the canonical map

$$H^1(\tau^*, \mathrm{GL}_\Lambda(\mathbb{O}_v)) \to H^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$$

is injective, and hence the assertion follows immediately from Proposition 5.2. □

5B2. *Global results for* GL_Λ. We analyse $H^1(\tau^*, \mathrm{GL}_\Lambda(F))$, again using the classification of τ_c-hermitian forms.

Proposition 5.4 (Hasse principle). *The canonical map*

$$H^1(\tau^*, \mathrm{GL}_\Lambda(F)) \longrightarrow \prod_{v \in \mathrm{Ram}_\infty(D)} H^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$$

induced by the inclusions is bijective. This means a class in $H^1(\tau^*, \mathrm{GL}_\Lambda(F))$ *is uniquely determined by its signatures at the real ramified places.*

Proof. If D is not a division algebra it is easily checked that $H^1(\tau^*, \mathrm{GL}_\Lambda(F)) = \{1\}$. Thus there is nothing to show.

Assume that D is a division algebra. The regular hermitian forms over D (with respect to τ_c) are classified by dimension and their signatures at the real places of F where D ramifies [Scharlau 1985, Chapter 10, Example 1.8]. The claim follows as in the local case. □

5B3. *The pfaffian associated with* τ. We explain how to compute the pfaffian associated with τ (see Section 2D1) for diagonal matrices. Let k be any extension field of F, for example a local completion. Given a diagonal matrix $x = \mathrm{diag}(x_1, \dots, x_n)$ with entries in k, we can consider x as a τ-fixed matrix in $A \otimes_F k = M_n(D \otimes_F k)$.

Lemma 5.5. *For* $x = \mathrm{diag}(x_1, \dots, x_n)$ *with entries in some extension field* k *of* F, *the pfaffian of* x *is the product of all entries*:

$$\mathrm{pf}_\tau(x) = x_1 x_2 \cdots x_n.$$

Proof. We can assume without loss of generality that k is algebraically closed. In this case $D \otimes_F k \cong M_2(k)$ and the reduced norm $\mathrm{nrd}_D : D \otimes_F k \to k$ agrees with the determinant, in particular it is surjective. This means, for given $i \in \{1, \dots, n\}$, we can write $x_i = \mathrm{nrd}_D(y_i) = \bar{y}_i y_i$ for some $y_i \in D \otimes_F k$. Consider the matrix $y = \mathrm{diag}(y_1, \dots, y_n) \in M_n(D \otimes_F k)$: it satisfies $\tau(y) y = x$. By Lemma 2.7 we obtain

$$\mathrm{pf}_\tau(x) = \mathrm{nrd}_A(y) = \prod_{i=1}^{n} \mathrm{nrd}_D(y_i) = \prod_{i=1}^{n} x_i.$$

Here we used that the reduced norm of a diagonal matrix in $M_n(D \otimes_F k)$ is the product of the reduced norms of the entries; see [Weil 1995, Chapter IX, §2, Corollary 2]. □

Note in particular that the pfaffian $\mathrm{pf}_\tau : \mathrm{Sym}(\Lambda, \tau) \to \mathbb{O}$ is surjective.

5B4. *Transfer of results to* SL_Λ. The final step in this section is to transfer the results on nonabelian Galois cohomology with values in GL_Λ to the group $G = \mathrm{SL}_\Lambda$. Our main tool is the cohomological diagram for symplectic involutions.

Lemma 5.6. *Let $v \in V \setminus \mathrm{Ram}_\infty(D)$ be a place of F. Then the pfaffian induces a bijection*

$$\mathrm{pf}_\tau : H^1(\tau^*, \mathrm{SL}_\Lambda(F_v)) \xrightarrow{\simeq} \{\pm 1\}.$$

Proof. If follows from Proposition 5.2 that $H^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$ is trivial. The cohomological diagram for symplectic involutions (see Proposition 2.8) collapses to

$$\begin{tikzcd}
\{\pm 1\} \arrow[r, "\delta"] \arrow[d, equal] & H^1(\tau^*, \mathrm{SL}_\Lambda(F_v)) \arrow[r] \arrow[d, "\mathrm{pf}_\tau"] & 1 \\
\{\pm 1\} \arrow[r, equal] & \{\pm 1\}. &
\end{tikzcd}$$

Here we used that $\mathrm{nrd}_\Lambda : \mathrm{GL}_\Lambda(F_v) \to F_v^\times$ is surjective, see [Reiner 2003, Theorem (33.4)]. By Proposition 2.8 the morphism δ is injective, and thus bijective. □

Lemma 5.7. *Let $v \in \mathrm{Ram}_\infty(D)$. The canonical map*

$$j_* : H^1(\tau^*, \mathrm{SL}_\Lambda(F_v)) \to H^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$$

is bijective.

Proof. In this case the reduced norm takes only positive values in $F_v \cong \mathbb{R}$. Therefore the cohomological diagram for symplectic involutions (Proposition 2.8) yields

$$\begin{tikzcd}
1 \arrow[r] & H^1(\tau^*, \mathrm{SL}_\Lambda(F_v)) \arrow[r, "j_*"] \arrow[d, "\mathrm{pf}_\tau"] & H^1(\tau^*, \mathrm{GL}_\Lambda(F_v)) \arrow[d, "\mathrm{pf}_\tau"] \\
1 \arrow[r] & \{\pm 1\} \arrow[r, "\simeq"] & \mathbb{R}^\times / \mathbb{R}^\times_{>0}.
\end{tikzcd}$$

It follows directly from Corollary 2.9 that j_* is surjective. Moreover, twisting the upper row with cocycles for $H^1(\tau^*, \mathrm{SL}_\Lambda(F_v))$ shows that j_* is indeed injective. For more details on twisting in nonabelian cohomology the reader may consult [Serre 1994, Chapter I, §5.4]. Note that twisting an involution of symplectic type gives an involution of symplectic type (see the remark on page 381). □

Lemma 5.8. *Let v be a finite place and let $\mathfrak{p}_v \subseteq \mathcal{O}_v$ be the prime ideal. For an integer $m \geq 0$ we define $K_v(m) := \ker(G(\mathcal{O}_v) \to G(\mathcal{O}_v/\mathfrak{p}_v^m))$. Then the pfaffian induces a bijection*

$$\mathrm{pf}_\tau : H^1(\tau^*, K_v(m)) \xrightarrow{\simeq} \begin{cases} \{\pm 1\} & \text{if } -1 \equiv 1 \mod \mathfrak{p}_v^m, \\ \{1\} & \text{otherwise.} \end{cases}$$

Proof. We start with the special case $m = 0$; here $K_v(m) = \mathrm{SL}_\Lambda(\mathbb{O}_v)$. Here the claim follows just as in the proof of Lemma 5.6 from Proposition 2.8, Corollary 5.3, and the fact that the reduced norm $\mathrm{nrd}_\Lambda : \mathrm{GL}_\Lambda(\mathbb{O}_v) \to \mathbb{O}_v^\times$ is onto [Reiner 2003, Theorem (14.1) and Exercise 5 on page 152].

For $m \geq 1$ consider the short exact sequence of groups

$$1 \longrightarrow K_v(m) \longrightarrow \mathrm{SL}_\Lambda(\mathbb{O}_v) \longrightarrow \mathrm{SL}_\Lambda(\mathbb{O}_v/\mathfrak{p}_v^m) \longrightarrow 1.$$

Note that this sequence uses that the order Λ, and hence the group scheme SL_Λ, is smooth by Lemma 5.1. We obtain a long exact sequence of pointed sets

$$G^{\tau^*}(\mathbb{O}_v) \xrightarrow{\pi} G^{\tau^*}(\mathbb{O}_v/\mathfrak{p}_v^m) \xrightarrow{\delta} H^1(\tau^*, K_v(m)) \xrightarrow{j_m} H^1(\tau^*, G(\mathbb{O}_v)).$$

It follows from the remark on page 379 that the fixed point group G^{τ^*} is just the group scheme $G(\Lambda, \tau)$ defined in Section 2C. Since the group scheme $G(\Lambda, \tau)$ is smooth (Proposition 2.6), the canonical map π is surjective, and so δ is trivial. Via twisting (see the remark on page 381) we obtain that j_m is injective.

We use that the pfaffian is a morphism of schemes defined over $\mathbb{O}$ (as explained in Section 2D1). Given a cocycle $b \in Z^1(\tau^*, K_v(m))$, we have $\mathrm{pf}_\tau(b) \equiv 1 \mod \mathfrak{p}_v^m$. Consequently, if 1 and -1 are not congruent modulo $\mathfrak{p}_v^m$, then $H^1(\tau^*, K_v(m)) = \{1\}$ and the claim follows.

Assume now that $-1 \equiv 1 \mod \mathfrak{p}_v^m$. Then the matrix $\mathrm{diag}(-1, 1, \ldots, 1)$ lies in $K_v(m)$ and has pfaffian -1 (see Section 5B3). □

For a real place $v \in V_\infty$ we denote the associated embedding $F \to \mathbb{R}$ by ι_v. Define

$$F_D^\times = \{x \in F^\times \mid \iota_v(x) > 0 \text{ for all } v \in \mathrm{Ram}_\infty(D)\}.$$

By the Hasse–Schilling–Maass theorem [Reiner 2003, Theorem (33.15)] the image of the reduced norm $\mathrm{nrd}_A : A^\times \to F^\times$ is $F_D^\times$.

Lemma 5.9. *Assume that* $\mathrm{Ram}_\infty(D)$ *is not empty. Then the canonical morphism of pointed sets*

$$j_* : H^1(\tau^*, \mathrm{SL}_\Lambda(F)) \longrightarrow H^1(\tau^*, \mathrm{GL}_\Lambda(F))$$

is injective. The image consists of precisely those classes $x \in H^1(\tau^*, \mathrm{GL}_\Lambda(F))$ *which satisfy* $\mathrm{pf}_\tau(x) = \pm 1 \cdot F_D^\times$.

If otherwise D splits at every real place, then the pfaffian induces a bijection

$$\mathrm{pf}_\tau : H^1(\tau^*, \mathrm{SL}_\Lambda(F)) \xrightarrow{\simeq} \{\pm 1\}.$$

Proof. Assume that $\mathrm{Ram}_\infty(D)$ is empty. By the Hasse–Schilling–Maass theorem the reduced norm $\mathrm{GL}_\Lambda(F) \to F^\times$ is surjective and the second assertion follows as in Lemma 5.6.

Now we assume that $\mathrm{Ram}_\infty(D)$ is not empty. The image of the reduced norm $\mathrm{nrd}_A : A^\times \to F^\times$ is $F_D^\times$. Note that $F_D^\times$ cannot contain the element -1 since $\mathrm{Ram}_\infty(D)$ is not empty. Consider the cohomological diagram for symplectic involutions (Proposition 2.8)

$$\begin{array}{ccccc} 1 & \longrightarrow & H^1(\tau^*, \mathrm{SL}_\Lambda(F)) & \xrightarrow{j_*} & H^1(\tau^*, \mathrm{GL}_\Lambda(F)) \\ & & \Big\downarrow \mathrm{pf}_\tau & & \Big\downarrow \mathrm{pf}_\tau \\ 1 & \longrightarrow & \{\pm 1\} & \longrightarrow & F^\times / F_D^\times. \end{array}$$

Twisting shows that the map j_* is injective. The assertion about the image of j_* follows immediately from Corollary 2.9. □

Remark. Assume that $\mathrm{Ram}_\infty(D)$ is not empty. Let $x \in H^1(\tau^*, \mathrm{GL}_\Lambda(F))$ be a cohomology class. For every place $v \in \mathrm{Ram}_\infty(D)$ the class x considered as a class in $H^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$ has a local signature (p_v, q_v). Then according to Lemma 5.9 the class x lies in the image of j_* if and only if

$$q_v \equiv q_w \mod 2$$

for every pair of places $v, w \in \mathrm{Ram}_\infty(D)$. This means that either all q_v are even or all q_v are odd.

Theorem 5.10. *Let $K_f = \prod_{v \in V_f} K_v \subseteq G(\mathbb{A}_f)$ be the open compact subgroup associated with the congruence subgroup* $\Gamma(\mathfrak{a})$ (see Section 5A3). *Consider the set* $\mathcal{H}^1(\tau^*)$ (see the beginning of Section 5B). *The projection $\pi : \mathcal{H}^1(\tau^*) \to H^1(\tau^*, G(F))$ is injective and there is a short exact sequence of pointed sets*

$$1 \longrightarrow \mathcal{H}^1(\tau^*) \xrightarrow{\pi} H^1(\tau^*, G(F)) \xrightarrow{\mathrm{pf}_\tau} \{\pm 1\} \longrightarrow 1.$$

Proof. Consider the nonabelian cohomology set $H^1(\tau^*, K_\infty K_f)$, which agrees with the direct product $H^1(\tau^*, K_\infty) \times H^1(\tau^*, K_f)$. The canonical map

$$H^1(\tau^*, K_\infty) \to H^1(\tau^*, G_\infty)$$

is bijective (see [An and Wang 2008] or [Rohlfs 1981, Lemma 1.4]). Moreover, for every finite place $v \in V_f$ the group K_v is of the form

$$K_v(m) = \ker(G(\mathcal{O}_v) \to G(\mathcal{O}_v/\mathfrak{p}_v^m))$$

for some integer m. It follows from Lemma 5.8 that the inclusion $K_v \to G(F_v)$ induces an injection $H^1(\tau^*, K_v) \to H^1(\tau^*, G(F_v))$. Therefore the canonical map $H^1(\tau^*, K_\infty K_f) \to H^1(\tau^*, G(\mathbb{A}))$ is injective and we conclude that the projection $\pi : \mathcal{H}^1(\tau^*) \to H^1(\tau^*, G(F))$ is injective.

Moreover, it follows from the considerations on diagonal matrices in Section 5B3 that the pfaffian $\mathrm{pf}_\tau : H^1(\tau^*, G(F)) \to \{\pm 1\}$ is surjective.

It remains to understand the image of π. Since $\Gamma(\mathfrak{a})$ is (by assumption) torsion-free, we know that -1 is not congruent to 1 modulo $\mathfrak{a}$. In particular there is a prime ideal $\mathfrak{p}$ which divides $\mathfrak{a}$, say $e = \nu_{\mathfrak{p}}(\mathfrak{a})$, such that 1 and -1 are not congruent modulo $\mathfrak{p}^e$. Let $v \in V_f$ be the finite place associated with $\mathfrak{p}$, then $K_v = K_v(e)$ and $H^1(\tau^*, K_v) = \{1\}$ by Lemma 5.8. Let $\gamma \in H^1(\tau^*, G(F))$ be in the image of π, say (x, γ) is the inverse image in $\mathcal{H}^1(\tau^*)$. Let x_v be the projection of the class x to $H^1(\tau^*, K_v)$. Since x and γ have the same image in $H^1(\tau^*, G(\mathbb{A}))$, we can deduce that $\mathrm{pf}_\tau(\gamma) = \mathrm{pf}_\tau(x_v) = 1$.

Conversely, given $\gamma \in H^1(\tau^*, G(F))$ in the kernel of the pfaffian, then γ lies in the image of π. Let $c_\infty \in H^1(\tau^*, K_\infty)$ be a cohomology class such that c_∞ and γ define the same class in $H^1(\tau^*, G_\infty)$. Let 1_f denote the trivial class in $H^1(\tau^*, K_f)$, then the triple $(c_\infty, 1_f, \gamma)$ is a class in $\mathcal{H}^1(\tau^*)$ which is mapped to γ by π. □

5C. ***The fixed point groups.*** Up to Section 5C6 the number field F is assumed to be *totally real*.

Definition. Let R be a commutative $\mathcal{O}$-algebra (for example $\mathcal{O}_v$ or F_v). For every cocycle γ in $Z^1(\tau^*, G(R))$ the R-group scheme $G(\gamma)$ of $\tau^*|\gamma$-fixed points is defined by

$$G(\gamma)(C) := \{g \in G(C) \mid g = {}^{\tau^*|\gamma}g\}$$

for any commutative R-algebra C. Recall that the γ-twisted τ^*-action is given by ${}^{\tau^*|\gamma}g = \gamma\, {}^{\tau^*}\!g \gamma^{-1}$.

We define the symplectic group Sp_n over $\mathbb{Z}$ by

$$\mathrm{Sp}_n(R) := \{g \in \mathrm{GL}_{2n}(R) \mid g^T J g = J\},$$

for every commutative ring R, where J is the standard symplectic matrix

$$J = \begin{pmatrix} 0_n & 1_n \\ -1_n & 0_n \end{pmatrix}.$$

Note that in this notation Sp_n is of rank n, but consists of matrices of size $2n \times 2n$.

Given a cocycle $\gamma \in Z^1(\tau^*, G(\mathcal{O}))$, we want to understand the associated group scheme $G(\gamma)$. In particular we want to calculate the Euler characteristic of congruence subgroups of this group. We start with some basic observations and afterwards we collect all the ingredients necessary for an application of the adelic Euler characteristic formula (Theorem 3.3).

Remark. If $\gamma \in Z^1(\tau^*, G(\mathcal{O}))$ then $G(\gamma) = G(\Lambda, \tau|\gamma)$ in the notation of Section 2C. The reason for this identity is that $G(\Lambda, \tau|\gamma)$ is always a closed subscheme of SL_Λ, that is, all elements have reduced norm one (see the remark on page 379). Here $\tau|\gamma$ is the γ-twisted involution on A (see the remark on page 381). Recall that

$\tau|\gamma$ is of symplectic type, and that twisting and the operation $*$ commute, that is $(\tau|\gamma)^* = \tau^*|\gamma$.

Lemma 5.11. *For every $\gamma \in Z^1(\tau^*, G(\mathbb{O}))$ the group scheme $G(\gamma)$ is smooth.*

Proof. By Lemma 5.1 the order Λ is τ-smooth. The remark on page 381 implies further that Λ is $(\tau|\gamma)$-smooth as well, and thus Proposition 2.6 yields that $G(\gamma) = G(\Lambda, \tau|\gamma)$ is smooth. $\square$

Lemma 5.12. *Let R be a commutative $\mathbb{O}$-algebra. Suppose the two cocycles $\gamma, \gamma' \in Z^1(\tau^*, G(R))$ define the same class in $H^1(\tau^*, \mathrm{GL}_\Lambda(R))$. Then $G(\gamma)$ and $G(\gamma')$ are isomorphic as group schemes over R.*

Proof. There is $c \in \mathrm{GL}_\Lambda(R)$ satisfying $\gamma' = c\gamma\, {}^\tau c$. We define a morphism of group schemes $f : G(\gamma) \to G(\gamma')$ by

$$f_C : g \mapsto cgc^{-1}$$

for every commutative R-algebra C and all $g \in G(\gamma)(C)$. This map is well-defined:

$${}^{\tau^*|\gamma'}(cgc^{-1}) = \gamma'\, {}^{\tau^*}c\, {}^{\tau^*}g\, {}^{\tau^*}c^{-1} \gamma'^{-1} = c\gamma\, {}^{\tau^*}g \gamma^{-1} c^{-1} = cgc^{-1}.$$

Obviously the inverse map of f is given by $g \mapsto c^{-1}gc$, so f is an isomorphism. $\square$

Corollary 5.13. *Let $\gamma \in Z^1(\tau^*, G(\mathbb{O}))$ be a cocycle, and let R be a commutative $\mathbb{O}$-algebra with $H^1(\tau^*, \mathrm{GL}_\Lambda(R)) = \{1\}$. There is an isomorphism of R-group schemes*

$$G(\gamma) \times_{\mathbb{O}} R \xrightarrow{\simeq} G(1) \times_{\mathbb{O}} R.$$

In particular, this holds if $R = \mathbb{O}_v$ for $v \in V_f$ (see Corollary 5.3).

Moreover, if k is a splitting field of D, then $G(\gamma) \times_{\mathbb{O}} k$ is isomorphic to the symplectic group $\mathrm{Sp}_n \times_{\mathbb{Z}} k$ defined over k.

Proof. The first part follows immediately from Lemma 5.12. For the second assertion note that we can choose a splitting $\varphi : A \otimes k \to M_{2n}(k)$ such that $\varphi(\tau(x))$ equals $J\varphi(x)^T J^{-1}$, where J denotes the standard symplectic matrix. $\square$

5C1. *The associated real Lie groups.* Let $\gamma \in Z^1(\tau^*, G(\mathbb{O}))$ be a cocycle. Consider the real Lie group

$$G(\gamma)_\infty = \prod_{v \in V_\infty} G(\gamma)(F_v)$$

associated with the group $G(\gamma)$.

Lemma 5.14. *Let $\gamma \in Z^1(\tau^*, G(\mathbb{O}))$ be a cocycle and let $v \in \mathrm{Ram}_\infty(D)$ be a real ramified place. If the class of γ in $H^1(\tau^*, G(F_v))$ has signature (p, q), then there is an isomorphism of real Lie groups*

$$G(\gamma)(F_v) \xrightarrow{\simeq} \mathrm{Sp}(p, q).$$

Here $\mathrm{Sp}(p,q)$ *is the real Lie group defined by*

$$\mathrm{Sp}(p,q) := \{g \in \mathrm{GL}_n(\mathbb{H}) \mid \bar{g}^T I_{p,q} g = I_{p,q}\}.$$

Proof. This follows from Lemma 5.12 and the description of the cohomology set $H^1(\tau^*, \mathrm{GL}_\Lambda(F_v))$ in Proposition 5.2. □

For a real ramified place $v \in \mathrm{Ram}_\infty(D)$, let (p_v, q_v) denote the local signature of the cohomology class of γ in $H^1(\tau^*, G(F_v))$. It follows from Corollary 5.13 and Lemma 5.14 that there is an isomorphism of real Lie groups

$$G(\gamma)_\infty \xrightarrow{\simeq} \mathrm{Sp}_n(\mathbb{R})^s \times \prod_{v \in \mathrm{Ram}_\infty(D)} \mathrm{Sp}(p_v, q_v).$$

Here s denotes the number of real places of F which split D. Note that $G(\gamma)_\infty$ is connected and semisimple. The real Lie algebra $\mathfrak{g}(\gamma)_\infty$ of $G(\gamma)_\infty$ is isomorphic to

$$\mathfrak{g}(\gamma)_\infty \cong \mathfrak{sp}(n, \mathbb{R})^s \oplus \bigoplus_{v \in \mathrm{Ram}_\infty(D)} \mathfrak{sp}(p_v, q_v).$$

Recall that every maximal compact subgroup of the real Lie group $\mathrm{Sp}_n(\mathbb{R})$ is isomorphic to the unitary group $\mathrm{U}(n)$.

Consider the group $\mathrm{Sp}(n) := \mathrm{Sp}(n, 0)$. One can check that this is a compact connected semisimple real Lie group [Knapp 2002, page 111]. Moreover, it is a maximal compact subgroup of the special linear group $\mathrm{SL}_n(\mathbb{H})$.

Let $p, q \geq 0$ be integers with $p + q = n$. The Lie group $\mathrm{Sp}(p,q)$ is connected and semisimple [Knapp 2002, Proposition 1.145], and the compact subgroup $\mathrm{Sp}(p) \times \mathrm{Sp}(q)$ is a maximal compact subgroup. Given any maximal compact subgroup $K(\gamma)_\infty \subseteq G(\gamma)_\infty$, we obtain an isomorphism of Lie groups

$$K(\gamma)_\infty \xrightarrow{\simeq} \mathrm{U}(n)^s \times \prod_{v \in \mathrm{Ram}_\infty(D)} \mathrm{Sp}(p_v) \times \mathrm{Sp}(q_v).$$

5C2. *The symmetric space.* Consider the associated Riemannian symmetric space $X(\gamma)$ defined as $X(\gamma) := K(\gamma)_\infty \backslash G(\gamma)_\infty$. We have $\dim G(\gamma) = n(2n+1)$, thus

$$\dim G(\gamma)_\infty = n(2n+1)[F : \mathbb{Q}].$$

The dimension of the unitary group $\mathrm{U}(n)$ is n^2 and consequently

$$\dim K(\gamma)_\infty = sn^2 + \sum_{v \in \mathrm{Ram}_\infty(D)} p_v(2p_v+1) + q_v(2q_v+1).$$

Subtraction of both dimensions yields an obviously *even* number:

$$\dim X(\gamma) = sn(n+1) + \sum_{v \in \mathrm{Ram}_\infty(D)} 4p_v q_v.$$

5C3. *Lie algebras and complexifications.* We complexify the Lie algebra $\mathfrak{g}(\gamma)_\infty$ and we obtain an isomorphism

$$\mathfrak{g}(\gamma)_\infty \otimes_{\mathbb{R}} \mathbb{C} \cong \mathfrak{sp}(n, \mathbb{C})^{[F:\mathbb{Q}]}.$$

The rank of this complex semisimple Lie algebra is $\mathrm{rk}(\mathfrak{g}(\gamma)_{\infty,\mathbb{C}}) = n[F:\mathbb{Q}]$. Let $\mathfrak{k}(\gamma)_\infty$ denote the Lie algebra of the maximal compact subgroup $K(\gamma)_\infty$. The complexification of this Lie algebra is isomorphic to

$$\mathfrak{k}(\gamma)_\infty \otimes_{\mathbb{R}} \mathbb{C} \cong \mathfrak{gl}(n, \mathbb{C})^s \oplus \bigoplus_{v \in \mathrm{Ram}_\infty(D)} \mathfrak{sp}(p_v, \mathbb{C}) \oplus \mathfrak{sp}(q_v, \mathbb{C}).$$

The rank of $\mathfrak{k}(\gamma)_{\infty,\mathbb{C}}$ is $sn + \sum_{v \in \mathrm{Ram}_\infty(D)} p_v + q_v = n[F:\mathbb{Q}]$. Thus the complexified Lie algebras $\mathfrak{g}(\gamma)_{\infty,\mathbb{C}}$ and $\mathfrak{k}(\gamma)_{\infty,\mathbb{C}}$ have equal rank. The Weyl groups of these complex reductive Lie algebras are well known; in particular,

$$|W(\mathfrak{g}(\gamma)_{\infty,\mathbb{C}})| = (2^n n!)^{[F:\mathbb{Q}]}$$

and

$$|W(\mathfrak{k}(\gamma)_{\infty,\mathbb{C}})| = (n!)^s \prod_{v \in \mathrm{Ram}_\infty(D)} 2^{p_v} p_v! \cdot 2^{q_v} q_v!,$$

as can be found in [Humphreys 1972, page 66]. The quotient of the cardinalities of the two Weyl groups is given by

$$\frac{|W(\mathfrak{g}(\gamma)_{\infty,\mathbb{C}})|}{|W(\mathfrak{k}(\gamma)_{\infty,\mathbb{C}})|} = 2^{ns} \prod_{v \in \mathrm{Ram}_\infty(D)} \binom{n}{p_v}.$$

Remark. The linear algebraic F-group $G(\gamma) \times_{\mathbb{O}} F$ is an inner form of the symplectic group Sp_n; in particular it is a semisimple and simply connected group. Further this implies that the Tamagawa number $\tau(G(\gamma))$ is equal to one [Kottwitz 1988].

5C4. *The metric form B.* Recall that the Lie algebra of $G(\gamma)$ is a functor $\mathrm{Lie}(G(\gamma))$ which assigns to a commutative $\mathbb{O}$-algebra C the C-Lie algebra

$$\mathrm{Lie}(G(\gamma))(C) = \{x \in (\Lambda \otimes_{\mathbb{O}} C)^\times \mid (\tau|\gamma)(x) = -x\}.$$

For simplicity we write $\mathfrak{g}(\gamma)_C$ instead of $\mathrm{Lie}(G(\gamma))(C)$.

Consider the nondegenerate form $B : \mathfrak{g}(\gamma)_F \times \mathfrak{g}(\gamma)_F \to F$ defined by $B(x, y) := -\frac{1}{2}\,\mathrm{trd}_A(xy)$. Let $\iota : F \to \mathbb{C}$ be an embedding of F into the field of complex numbers. The central simple algebra $A = M_n(D)$ splits over $\mathbb{C}$ and we can choose a splitting $A \to M_{2n}(\mathbb{C})$ such that $\tau|\gamma$ is the standard symplectic involution. Via this splitting the Lie algebra $\mathfrak{g}(\gamma)_{\mathbb{C}}$ is isomorphic to the complex semisimple Lie algebra $\mathfrak{sp}(n, \mathbb{C})$.

Proposition 5.15. *Consider the compact Lie group* $\mathrm{Sp}(n)$ *and its Lie algebra*

$$\mathfrak{sp}(n) := \{x \in M_n(\mathbb{H}) \mid \bar{x}^T + x = 0\},$$

and the positive definite $\mathbb{R}$*-bilinear form* $B(x,y) := -\frac{1}{2}\operatorname{trd}(xy)$ *on* $\mathfrak{sp}(n)$*. With respect to the right-invariant Riemann metric induced by* B*, the group* $\mathrm{Sp}(n)$ *has the volume*

$$\mathrm{vol}_B(\mathrm{Sp}(n)) = \prod_{j=1}^{n} \frac{(2\pi)^{2j}}{2\cdot(2j-1)!}.$$

Proof. The $\mathbb{C}$-linear extension of B to $\mathfrak{sp}(n,\mathbb{C})$ is given by $B(x,y) = -\frac{1}{2}\operatorname{Tr}(xy)$. Recall that the Killing form β on $\mathfrak{sp}(n,\mathbb{C})$ is the form $\beta(x,y) = (2n+2)\operatorname{Tr}(xy)$ [Helgason 1978, Chapter III, §8] and hence $\beta = -4(n+1)B$. We conclude

$$\mathrm{vol}_\beta(\mathrm{Sp}(n)) = (4(n+1))^{\frac{n(2n+1)}{2}}\,\mathrm{vol}_B(\mathrm{Sp}(n)).$$

The assertion follows from Ono's formula for the volume of a compact Lie group with respect to the Killing form [Ono 1966, Equation 3.4.9], which yields

$$\mathrm{vol}_\beta(\mathrm{Sp}(n)) = (4(n+1))^{\frac{n(2n+1)}{2}} \prod_{j=1}^{n} \frac{(2\pi)^{2j}}{2\cdot(2j-1)!}. \qquad \square$$

5C5. *The modulus factor.* Consider the F-bilinear form $B : \mathfrak{g}(\gamma)_F \times \mathfrak{g}(\gamma)_F \to F$ defined by $B(x,y) := -\frac{1}{2}\operatorname{trd}_A(xy)$. In this paragraph we will calculate the global modulus factor $m(B) = \prod_{v\in V_f} m(B)_v$ (see Section 3B2). Note that Λ is in general not a free $\mathbb{O}$-module, therefore we have to work locally.

We start with the finite places $v \in V_f$ where D splits. The main observation is this: we can assume that $\Lambda \otimes_{\mathbb{O}} \mathbb{O}_v = M_{2n}(\mathbb{O}_v)$ and that $\tau|\gamma$ is the standard symplectic involution. This follows from the next lemma.

Lemma 5.16. *Let* R *be a complete discrete valuation ring with field of fractions* k *of characteristic* $\operatorname{char}(k) \neq 2$*. Let* σ *be an involution of symplectic type on* $M_{2n}(k)$ *and let* $\Lambda \subseteq M_{2n}(k)$ *be a maximal* R*-order which is* σ*-stable.*

There is an element $g \in \mathrm{GL}_{2n}(k)$ *such that*

(i) $g\Lambda g^{-1} = M_{2n}(R)$, *and*

(ii) $g\sigma(x)g^{-1} = J(gxg^{-1})^T J^{-1}$, *where* J *is the standard symplectic matrix.*

Proof. It follows from [Reiner 2003, Theorem (17.3)] that there is an invertible matrix $a \in \mathrm{GL}_{2n}(k)$ such that $a\Lambda a^{-1} = M_{2n}(R)$. Moreover, σ is an involution of symplectic type and we can consider $\operatorname{int}(a) : M_{2n}(k) \to M_{2n}(k)$ as a splitting of the central simple k-algebra $M_{2n}(k)$. There is a matrix $h \in \mathrm{GL}_{2n}(k)$ such that $h^T = -h$ and $\operatorname{int}(a)(\sigma(x)) = h(\operatorname{int}(a)(x))^T h^{-1}$ for every $x \in M_{2n}(k)$.

Because Λ is σ-stable, $hM_{2n}(R)h^{-1} = M_{2n}(R)$. After multiplication with some power of the prime element in R, we can assume $h \in \mathrm{GL}_{2n}(R)$. On a free module over a complete discrete valuation ring, there is only one regular symplectic form up to isogeny (since $\operatorname{char}(k) \neq 2$); this means that there is $b \in \mathrm{GL}_{2n}(R)$ such that

$bhb^T = J$. Finally, we define $g := ba$ and observe that

$$g\sigma(x)g^{-1} = bh(axa^{-1})^T h^{-1} b^{-1} = J(b^{-1})^T (axa^{-1})^T b^T J^{-1} = J(gxg^{-1})^T J^{-1}$$

for every $x \in M_{2n}(k)$. $\square$

Corollary 5.17. *Let $\gamma \in Z^1(\tau^*, G(\mathbb{O}))$ be a cocycle and let $v \in V_f$ be a finite place of F which splits D. There is an isomorphism of group schemes over $\mathbb{O}_v$:*

$$G(\gamma) \times_{\mathbb{O}} \mathbb{O}_v \xrightarrow{\simeq} \mathrm{Sp}_n \times_{\mathbb{Z}} \mathbb{O}_v.$$

Proof. This follows directly from the previous lemma since $\tau|\gamma$ is an involution of symplectic type. $\square$

Proposition 5.18. *Let $v \in V_f$ be a finite place which splits D. Consider the bilinear form $B : \mathfrak{g}(\gamma)_F \times \mathfrak{g}(\gamma)_F \to F$ defined by $B(x, y) = -\frac{1}{2}\operatorname{trd}(xy)$. The local modulus factor (see Section 3B2) is*

$$m(B)_v = |2|_v^{-n}.$$

Proof. By Corollary 5.17 we can assume $G(\gamma) = \mathrm{Sp}_n$ over $\mathbb{O}_v$. This means

$$\mathfrak{g}(\gamma)_{\mathbb{O}_v} = \mathfrak{sp}(n, \mathbb{O}_v) = \{x \in M_{2n}(\mathbb{O}_v) \mid x^T J + Jx = 0\}.$$

Note that the form B is given by the analogous formula $B(x, y) = -\frac{1}{2}\operatorname{Tr}(xy)$. Recall that the elements of $\mathfrak{sp}(n, \mathbb{O}_v)$ are matrices of the form

$$\begin{pmatrix} a & b \\ c & -a^T \end{pmatrix}$$

with $a, b, c \in M_n(\mathbb{O}_v)$, where b and c are symmetric matrices. Let $E_{s,t}$ denote the elementary $2n \times 2n$ matrix with exactly one entry 1 in position (s, t). We choose an $\mathbb{O}_v$-basis of $\mathfrak{sp}(n, \mathbb{O}_v)$ which is made up of the following elements:

(i) $a_{i,j} := E_{i,j} - E_{j+n,i+n}$ for all $i, j \in \{1, \dots, n\}$,

(ii) $b_{i,j} := E_{i,j+n} + E_{j,i+n}$ for all $1 \le i < j \le n$,

(iii) $c_{i,j} := E_{i+n,j} + E_{j+n,i}$ for all $1 \le i < j \le n$, and

(iv) $b_i := E_{i,i+n}$ and $c_i := E_{i+n,i}$ for all $i \in \{1, \dots, n\}$.

We evaluate the form B on all the basis vectors.

It is easy to observe that

$$\begin{aligned} 0 = B(a_{i,j}, c_{k,l}) &= B(a_{i,j}, b_{k,l}) = B(a_{i,j}, b_k) = B(a_{i,j}, c_k) \\ &= B(c_{i,j}, c_{k,l}) = B(b_{i,j}, b_{k,l}) = B(c_i, c_j) = B(b_i, b_j). \end{aligned}$$

for all i, j, k, l. Moreover, one readily verifies that $B(b_{i,j}, c_k) = B(c_{i,j}, b_k) = 0$ for all i, j, k. The remaining cases yield

- $B(a_{i,j}, a_{k,l}) = -\delta_{j,k}\delta_{i,l}$ for all $i, j, k, l \in \{1, \dots, n\}$,

- $B(b_{i,j}, c_{k,l}) = -\delta_{i,k}\delta_{j,l}$ for all $i < j \le n$ and $k < l \le n$, and
- $B(b_i, c_j) = -\frac{1}{2}\delta_{i,j}$ for all $i, j \in \{1, \dots, n\}$.

Using these results, we are able to calculate the modulus factor and obtain

$$m(B)_v = \left|\det\begin{pmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{pmatrix}\right|_v^{n/2} = |2|_v^{-n}. \qquad \square$$

Proposition 5.19. *Let $v \in \mathrm{Ram}_f(D)$ be a finite ramified place and $\mathfrak{p} \subset \mathcal{O}$ be the associated prime ideal. The local modulus factor for the group $G(\gamma)$ and the form B defined in Section 5C4 is*

$$m(B)_v = |2|_v^{-n}\, \mathrm{N}(\mathfrak{p})^{-n(n+1)/2}.$$

Proof. The F_v-algebra $D_v := D \otimes_F F_v$ is the unique quaternion division algebra over F_v and $\Delta := \Lambda_D \otimes_{\mathcal{O}} \mathcal{O}_v$ is the unique maximal order in D_v.

Set $H := G(1) \times \mathcal{O}_v$. Due to Corollary 5.13 we can assume that $\gamma = 1$, that is, $G(\gamma) \times \mathcal{O}_v$ is isomorphic to H. We define

$$\mathfrak{h} := \mathrm{Lie}(H)(\mathcal{O}_v) = \{x \in M_n(\Delta) \mid \tau(x) = -x\}.$$

Recall that $\tau(x) = \bar{x}^T$.

Take an $\mathcal{O}_v$-basis v_0, v_1, v_2, v_3 of Δ such that $\mathrm{trd}_D(v_0) = 1$ and $\mathrm{trd}_D(v_i) = 0$ for $i = 1, 2, 3$. Such a basis exists since $\mathrm{trd}_D : \Delta \to \mathcal{O}_v$ is surjective (maximal orders are smooth; see Proposition 2.4). We construct an $\mathcal{O}_v$-basis of the Lie algebra $\mathfrak{h}$, consisting of the following elements:

(i) $a_{s,i} := v_s E_{i,i}$ for all $s \in \{1, 2, 3\}$ and $i \in \{1, \dots, n\}$, and

(ii) $b_{s,i,j} := v_s E_{i,j} - \bar{v}_s E_{j,i}$ for all $s \in \{0, 1, 2, 3\}$ and $i, j \in \{1, \dots, n\}$ with $i < j$.

We calculate the form B on all basis vectors. Observe that $B(a_{s,i}, b_{t,k,l}) = 0$ for all s, t, i, k, l. Moreover, for $s, t \in \{1, 2, 3\}$ and $i, j \in \{1, \dots, n\}$ we find

$$B(a_{s,i}, a_{t,j}) = -\tfrac{1}{2}\,\mathrm{trd}(v_s E_{i,i} v_t E_{j,j}) = -\tfrac{1}{2}\delta_{i,j}\,\mathrm{trd}_D(v_s v_t).$$

Finally, let $s, t \in \{0, 1, 2, 3\}$ and let $i, j, k, l \in \{1, \dots, n\}$ with $i < j$ and $k < l$. We obtain

$$B(b_{s,i,j}, b_{t,k,l}) = \tfrac{1}{2}\,\mathrm{trd}_D(\bar{v}_s v_t + v_s \bar{v}_t)\delta_{i,k}\delta_{j,l} = \mathrm{trd}_D(v_s \bar{v}_t)\delta_{i,k}\delta_{j,l}.$$

Summing up we obtain a formula for the modulus factor

$$(3) \quad m(B)_v^2 = \left|\tfrac{1}{8}\det(\mathrm{trd}(v_s v_t))_{s,t=1,2,3}\right|_v^n \cdot \left|\det(\mathrm{trd}(v_s \bar{v}_t))_{s,t=0,1,2,3}\right|_v^{n(n-1)/2}$$

Since the elements $\bar{v}_0, \bar{v}_1, \bar{v}_2, \bar{v}_3$ form an $\mathcal{O}_v$-basis of Δ as well, we see that the second term $|\det(\mathrm{trd}(v_s \bar{v}_t))_{s,t=0,1,2,3}|_v$ is the valuation of the discriminant of Δ. It is known that the discriminant of Δ is $\mathfrak{p}_v^2$ [Reiner 2003, Theorem (14.9)].

To calculate the first term in (3) we consider $w_0 := 1$ and we define $w_s = v_s$ for $s = 1, 2, 3$. Note that w_0, w_1, w_2, w_3 is in general not an $\mathbb{O}_v$-basis of Δ since $\mathrm{trd}_D(1) = 2$ need not be a unit in $\mathbb{O}_v$. We can write

$$w_0 = 1 = r_0 v_0 + r_1 v_1 + r_2 v_2 + r_3 v_3$$

for certain r_0, r_1, r_2, r_3 in $\mathbb{O}_v$. Applying the reduced trace we get $2 = \mathrm{trd}_D(1) = r_0$. Furthermore, this implies that the matrix $(\mathrm{trd}(w_s w_t))_{s,t=0,1,2,3}$ can be written as a product of matrices

$$\begin{pmatrix} 2 & r_1 & r_2 & r_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} (\mathrm{trd}(v_i v_j))_{i,j=0,1,2,3} \begin{pmatrix} 2 & 0 & 0 & 0 \\ r_1 & 1 & 0 & 0 \\ r_2 & 0 & 1 & 0 \\ r_3 & 0 & 0 & 1 \end{pmatrix}.$$

Note that

$$(\mathrm{trd}(w_s w_t))_{s,t=0,1,2,3} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & \mathrm{trd}(v_1 v_1) & \mathrm{trd}(v_1 v_2) & \mathrm{trd}(v_1 v_3) \\ 0 & \mathrm{trd}(v_2 v_1) & \mathrm{trd}(v_2 v_2) & \mathrm{trd}(v_2 v_3) \\ 0 & \mathrm{trd}(v_3 v_1) & \mathrm{trd}(v_3 v_2) & \mathrm{trd}(v_3 v_3) \end{pmatrix}.$$

We deduce that $|\det(\mathrm{trd}(v_s v_t))_{s,t=1,2,3}|_v = |2|_v \, \mathrm{N}(\mathfrak{p})^{-2}$. In total the local modulus factor is

$$m(B)_v = |2|_v^{-n} \, \mathrm{N}(\mathfrak{p})^{-n-n(n-1)/2} = |2|_v^{-n} \, \mathrm{N}(\mathfrak{p})^{-n(n+1)/2}. \qquad \square$$

Corollary 5.20. *Let $\gamma \in Z^1(\tau^*, G(\mathbb{O}))$ be a cocycle. The global modulus factor $m(B)$ for the group $G(\gamma)$ with respect to the form B defined in Section 5C4 is*

$$m(B) = 2^{n[F:\mathbb{Q}]}(-1)^{rn(n+1)/2} \Delta_{\mathrm{rd}}(D)^{-n(n+1)/2},$$

where $\Delta_{\mathrm{rd}}(D)$ denotes the signed reduced discriminant of D (see the definition on page 390).

Proof. By Proposition 5.18, Proposition 5.19, and an application of the product formula we obtain

$$m(B) = \prod_{v \in V_f} |2|_v^{-n} \prod_{\mathfrak{p} \in \mathrm{Ram}_f(D)} \mathrm{N}(\mathfrak{p})^{-n(n+1)/2} = 2^{n[F:\mathbb{Q}]} \prod_{\mathfrak{p} \in \mathrm{Ram}_f(D)} \mathrm{N}(\mathfrak{p})^{-n(n+1)/2}. \qquad \square$$

5C6. *The Euler characteristic of the fixed point groups.* Let $\gamma \in Z^1(\tau^*, G(\mathbb{O}))$ be a cocycle. We are now able to compute the Euler characteristic of torsion-free arithmetic subgroups of $G(\gamma)$. In Theorem 5.21 we give a precise formula for principal congruence subgroups. More general subgroups can be treated analogously.

For Theorem 5.21 the number field F need not be totally real. Let $\mathfrak{a} \subset \mathcal{O}$ be a proper ideal. For a finite place $v \in V_f$ we define $K_v(\gamma, \mathfrak{a})$ to be the kernel of the reduction morphism $G(\gamma)(\mathcal{O}_v) \to G(\gamma)(\mathcal{O}_v/\mathfrak{a}\mathcal{O}_v)$. Note that $K_v(\gamma, \mathfrak{a}) = G(\gamma)(\mathcal{O}_v)$ for almost all places. The group

$$K_f(\gamma, \mathfrak{a}) := \prod_{v \in V_f} K_v(\gamma, \mathfrak{a})$$

is an open compact subgroup of the locally compact group $G(\gamma)(\mathbb{A}_f)$. This subgroup is given by a local datum (U, α) (see Section 3B3). Let $v \in V_f$ be a finite place and let $\mathfrak{p}$ be the associated prime ideal. Let $e = \nu_{\mathfrak{p}}(\mathfrak{a})$ be the exponent of $\mathfrak{p}$ in $\mathfrak{a}$. We have $\alpha_v = 1$ and $U_v = G(\gamma)(\mathcal{O}/\mathfrak{p})$ if $e = 0$, otherwise $\alpha_v = e$ and $U_v = \{1\} \subseteq G(\gamma)(\mathcal{O}/\mathfrak{p}^e)$.

Let $G(\gamma)_\infty = \prod_{v \in V_\infty} G(\gamma)(F_v)$ and let $K(\gamma)_\infty \subseteq G(\gamma)_\infty$ be a maximal compact subgroup. For every real ramified place $v \in \mathrm{Ram}_\infty(D)$ we denote the local signature of the class of γ in $H^1(\tau^*, G(F_v))$ by (p_v, q_v) (see the definition on page 393).

Theorem 5.21. *Assume that $G(\gamma)(F)$ acts freely on $K(\gamma)_\infty K_f(\gamma, \mathfrak{a}) \backslash G(\gamma)(\mathbb{A})$. The Euler characteristic of the double quotient space*

$$S(\mathfrak{a}) := K(\gamma)_\infty K_f(\gamma, \mathfrak{a}) \backslash G(\gamma)(\mathbb{A}) / G(\gamma)(F)$$

is nonzero if and only if F is totally real. In this case the following formula holds:

$$\chi(S(\mathfrak{a})) = 2^{-nr}\, \mathrm{N}(\mathfrak{a})^{n(2n+1)} \Delta_{\mathrm{rd}}(D)^{n(n+1)/2} \prod_{v \in \mathrm{Ram}_\infty(D)} \binom{n}{p_v} \prod_{j=1}^{n} M(j, \mathfrak{a}, D),$$

where $M(j, \mathfrak{a}, D)$ is defined as

$$M(j, \mathfrak{a}, D) := \zeta_F(1-2j) \prod_{\mathfrak{p} \mid \mathfrak{a}} \bigl(1 - \mathrm{N}(\mathfrak{p})^{-2j}\bigr) \prod_{\substack{\mathfrak{p} \in \mathrm{Ram}_f(D) \\ \mathfrak{p} \nmid \mathfrak{a}}} \bigl(1 + (-\mathrm{N}(\mathfrak{p}))^{-j}\bigr).$$

Here r is the number of real places of F where D is ramified. The sign of $\chi(S(\mathfrak{a}))$ is $(-1)^{sn(n+1)/2}$, where s denotes the number of real places where D splits.

Proof. It follows from the remark on page 384 that the Euler characteristic vanishes whenever F has a complex place. Therefore we may assume that F is totally real. We want to apply the adelic Euler characteristic formula (Theorem 3.3). We know that $G(\gamma)$ is a smooth group scheme over $\mathcal{O}$ (see Lemma 5.11). Further $G \times_{\mathcal{O}} F$ is an inner form of the symplectic group, and is thus a semisimple and simply connected algebraic group of dimension $d = n(2n+1)$ (see the remark on page 401). Note further that by assumption $G(\gamma)(F)$ acts freely on $K(\gamma)_\infty K_f(\gamma, \mathfrak{a}) \backslash G(\gamma)(\mathbb{A})$.

Moreover, we observe that $\dim X(\gamma)$ is even (see Section 5C2) and that the complexified Lie algebras $\mathfrak{k}(\gamma)_\infty \otimes \mathbb{C}$ and $\mathfrak{g}(\gamma)_{\infty,\mathbb{C}}$ have equal rank (see Section 5C3). We conclude that the Euler characteristic does not vanish and Theorem 3.3 applies.

We fix the nondegenerate bilinear form $B : \mathfrak{g}(\gamma)_F \times \mathfrak{g}(\gamma)_F \to F$ defined, as above, by $B(x, y) := -\frac{1}{2}\operatorname{trd}_A(xy)$. It is easy to see that the compact dual group $G(\gamma)_u$ of $G(\gamma)_\infty$ is isomorphic to $\mathrm{Sp}(n)^{[F:\mathbb{Q}]}$. Note further that B is given by the same formula on each factor of the compact dual group. Therefore the volume is

$$\mathrm{vol}_B(G(\gamma)_u) = \left(\prod_{j=1}^{n} \frac{(2\pi)^{2j}}{2\cdot(2j-1)!}\right)^{[F:\mathbb{Q}]}$$

according to Proposition 5.15. Using the global modulus factor, calculated in Corollary 5.20, and the quotient of the orders of the involved Weyl groups, derived in Section 5C3, the adelic formula yields

$$\begin{aligned}(4)\quad \chi(S(\mathfrak{a})) = (-1)^{[F:\mathbb{Q}]n(n+1)/2}|d_F|^{d/2}2^{ns} \prod_{v\in \mathrm{Ram}_\infty(D)} \binom{n}{p_v} \\ \cdot\left(\prod_{j=1}^{n} \frac{2\cdot(2j-1)!}{(2\pi)^{2j}}\right)^{[F:\mathbb{Q}]} 2^{-n[F:\mathbb{Q}]}\Delta_{\mathrm{rd}}(D)^{n(n+1)/2} \prod_{\mathfrak{p}\in V_f} \frac{\mathrm{N}(\mathfrak{p})^{d\alpha_\mathfrak{p}}}{|U_\mathfrak{p}|}.\end{aligned}$$

Here s denotes the number of real places of F which split D. The only terms that can be negative are $(-1)^{[F:\mathbb{Q}]n(n+1)/2}$ and the signed reduced discriminant. Consequently, the sign of the Euler characteristic is $(-1)^{sn(n+1)/2}$.

Let $v \in V_f$ be a finite place with associated prime ideal $\mathfrak{p}$ and consider $\dfrac{\mathrm{N}(\mathfrak{p})^{d\alpha_\mathfrak{p}}}{|U_\mathfrak{p}|}$.

Case (a): D splits at v and $\mathfrak{p}$ does not divide $\mathfrak{a}$. In this case $\alpha_\mathfrak{p} = 1$ and $U_\mathfrak{p} = G(\gamma)(\mathbb{O}/\mathfrak{p})$. Since $G(\gamma)$ is isomorphic to Sp_n over $\mathbb{O}_v$ (see Corollary 5.17), there is an isomorphism of finite groups $G(\gamma)(\mathbb{O}/\mathfrak{p}) \cong \mathrm{Sp}_n(\mathbb{O}/\mathfrak{p})$. From [Wilson 2009, Section 3.5] we deduce that

$$\left|G(\gamma)(\mathbb{O}/\mathfrak{p})\right| = \mathrm{N}(\mathfrak{p})^d \prod_{j=1}^{n} \left(1-\mathrm{N}(\mathfrak{p})^{-2j}\right).$$

Case (b): D is ramified at v and $\mathfrak{p}$ does not divide $\mathfrak{a}$. In this situation we have $\alpha_\mathfrak{p} = 1$ and $U_\mathfrak{p} = G(\gamma)(\mathbb{O}/\mathfrak{p})$. Let $k = \mathbb{O}/\mathfrak{p}$ be the finite residue class field and let ℓ/k be the unique quadratic extension. It is an easy exercise to show that $G(\gamma)(\mathbb{O}/\mathfrak{p})$ is isomorphic to a semidirect product $U(\ell/k) \ltimes \mathrm{Sym}_n(\ell)$, where $U(\ell/k)$ denotes the unitary group of the quadratic extension ℓ/k and $\mathrm{Sym}_n(\ell)$ denotes the abelian group of symmetric $(n\times n)$-matrices with entries in ℓ. Therefore (using [Wilson 2009, Section 3.6]) we get

$$|G(\gamma)(\mathbb{O}/\mathfrak{p})| = \mathrm{N}(\mathfrak{p})^d \prod_{j=1}^{n} \left(1-(-\mathrm{N}(\mathfrak{p}))^{-j}\right).$$

Case (c): $\mathfrak{p}$ divides $\mathfrak{a}$. In this case $\alpha_v = \nu_{\mathfrak{p}}(\mathfrak{a})$ and $|U_{\mathfrak{p}}| = 1$. Consequently,

$$\frac{\mathrm{N}(\mathfrak{p})^{d\alpha_{\mathfrak{p}}}}{|U_{\mathfrak{p}}|} = N(\mathfrak{p})^{d\nu_{\mathfrak{p}}(\mathfrak{a})}.$$

The product of these terms is

$$\prod_{\mathfrak{p}\in V_f} \frac{\mathrm{N}(\mathfrak{p})^{d\alpha_{\mathfrak{p}}}}{|U_{\mathfrak{p}}|} = \mathrm{N}(\mathfrak{a})^d \prod_{j=1}^{n} \Bigl(\zeta_F(2j) \prod_{\mathfrak{p}|\mathfrak{a}} \bigl(1-\mathrm{N}(\mathfrak{p})^{-2j}\bigr) \prod_{\substack{\mathfrak{p}\in \mathrm{Ram}_f(D)\\ \mathfrak{p}\nmid\mathfrak{a}}} \bigl(1+(-\mathrm{N}(\mathfrak{p}))^{-j}\bigr)\Bigr).$$

Here ζ_F denotes the zeta function of the number field F.

Note that $d = n(2n+1) = \sum_{j=1}^{n} 4j-1$ and so $|d_F|^{d/2} = \prod_{j=1}^{n} |d_F|^{(4j-1)/2}$. The functional equation of the zeta function of the totally real number field F [Weil 1995, Chapter VII §6, Theorem 3] yields

$$\zeta_F(2j)|d_F|^{(4j-1)/2}\Bigl(\frac{2\cdot(2j-1)!}{(2\pi)^{2j}}\Bigr)^{[F:\mathbb{Q}]} = (-1)^{j[F:\mathbb{Q}]}\zeta_F(1-2j)$$

for every integer $j \geq 1$. Using this we see that

$$|d_F|^{d/2}\Bigl(\prod_{j=1}^{n} \frac{2\cdot(2j-1)!}{(2\pi)^{2j}}\Bigr)^{[F:\mathbb{Q}]} \prod_{j=1}^{n} \zeta_F(2j) = (-1)^{[F:\mathbb{Q}]n(n+1)/2} \prod_{j=1}^{n} \zeta_F(1-2j).$$

Substitute this into (4); then a simple calculation proves the claim. □

5D. ***Proof of the main theorem.*** The notation and assumptions are those of the introduction. As usual F denotes an algebraic number field and $\mathbb{O}$ denotes its ring of integers. Let D be a quaternion algebra defined over F and let $\Lambda_D \subseteq D$ be a maximal $\mathbb{O}$-order. Let $n \geq 1$ be an integer; we consider the central simple F-algebra $A = M_n(D)$ and the maximal $\mathbb{O}$-order $\Lambda = M_n(\Lambda_D)$. Further $G := \mathrm{SL}_\Lambda$ is the smooth $\mathbb{O}$-group scheme defined as the kernel of the reduced norm over the order Λ (see the first definition on page 376).

We say that the quaternion algebra D over F is *totally definite* if F is totally real and D ramifies at every real place of F.

The algebraic group $G \times_{\mathbb{O}} F$ has strong approximation since it is an F-simple, simply connected group and $G_\infty \cong \mathrm{SL}_{2n}(\mathbb{R})^s \times \mathrm{SL}_n(\mathbb{H})^r \times \mathrm{SL}_{2n}(\mathbb{C})^t$ is not compact. Since the group $\mathrm{SL}_1(\mathbb{H})$ is compact, we need the assumption that $n \geq 2$ if D is totally definite.

Let $K_\infty \subseteq G_\infty$ be a τ^*-stable maximal compact subgroup. Further, let K_f be the open compact subgroup of $G(\mathbb{A}_f)$, which satisfies $\Gamma(\mathfrak{a}) = K_f \cap G(F)$ (see Section 5A3). Since $\Gamma(\mathfrak{a})$ is torsion-free and τ^*-stable, we can apply Theorem 4.2

and we obtain

$$\mathscr{L}(\tau^*, \Gamma(\mathfrak{a}), W) = \sum_{\eta \in \mathscr{H}^1(\tau^*)} \chi(\vartheta^{-1}(\eta)) \operatorname{Tr}(\tau^* | W(\gamma_\eta)). \tag{5}$$

Here γ_η is any representative of the $H^1(\tau^*, G(F))$ component of η and

$$\vartheta : \left(K_\infty K_f \backslash G(\mathbb{A})/G(F)\right)^{\tau^*} \to \mathscr{H}^1(\tau^*)$$

is the surjective continuous map defined in Section 4.

By Theorem 5.10 the projection $\pi : \mathscr{H}^1(\tau^*) \to H^1(\tau^*, G(F))$ is injective and there is an exact sequence of pointed sets

$$1 \longrightarrow \mathscr{H}^1(\tau^*) \xrightarrow{\pi} H^1(\tau^*, G(F)) \xrightarrow{\mathrm{pf}_\tau} \{\pm 1\} \longrightarrow 1. \tag{6}$$

We deduce that, given a class $\eta \in \mathscr{H}^1(\tau^*)$, every representative $\gamma_\eta \in \pi(\eta)$ has pfaffian one, and hence they all describe the trivial class in $H^1(\tau^*, G(\bar{F}))$. Thus there is some $g \in G(\bar{F})$ such that $\gamma_\eta = g^{-1}\, {}^{\tau^*}\!g$. It follows that $\operatorname{Tr}(\tau^*|W(\gamma_\eta)) = \operatorname{Tr}(\tau^*|W)$ since $\tau^*|\gamma_\eta = \rho(g)^{-1} \circ \tau^* \circ \rho(g)$ on W.

As a next step we describe the fixed point components. Let $\eta \in \mathscr{H}^1(\tau^*)$. Using strong approximation we can choose representing cocycles k_η in $Z^1(\tau^*, K_\infty K_f)$ and γ_η in $Z^1(\tau^*, \Gamma(\mathfrak{a}))$, and an element $a_\infty \in G_\infty$ such that

$$\eta = ([k_\eta], [\gamma_\eta]) \quad \text{and} \quad {}^{\tau^*}\!a_\infty = k_\eta^{-1} a_\infty \gamma_\eta.$$

We write $k_\eta = k_\infty k_0$ with $k_\infty \in K_\infty$ and $k_0 \in K_f$. Note that $k_0 = \gamma_\eta$ considered as elements in $G(\mathbb{A}_f)$. By Lemma 4.1 there is a homeomorphism

$$\vartheta^{-1}(\eta) \xrightarrow{\simeq} (a_\infty^{-1} K_\infty^{\tau^*|k_\infty} a_\infty) K_f(\gamma_\eta, \mathfrak{a}) \backslash G(\gamma_\eta)(\mathbb{A})/G(\gamma_\eta)(F).$$

In fact $(a_\infty^{-1} K_\infty^{\tau^*|k_\infty} a_\infty)$ is a maximal compact subgroup of $G(\gamma_\eta)_\infty$.

Let $v \in \mathrm{Ram}_\infty(D)$ and let (p_v, q_v) denote the local signature of γ_η at v. By Theorem 5.21 the Euler characteristic of the fixed point component is zero if F has a complex place. If F is totally real, which we assume from now on, then

$$\chi(\vartheta^{-1}(\eta)) = 2^{-nr}\, \mathrm{N}(\mathfrak{a})^{n(2n+1)} \Delta_{\mathrm{rd}}(D)^{n(n+1)/2} \prod_{v \in \mathrm{Ram}_\infty(D)} \binom{n}{p_v} \prod_{j=1}^{n} M(j, \mathfrak{a}, D).$$

The short exact sequence (6), in combination with the Hasse principle and Lemma 5.9, shows that the map which takes cocycles to their local signatures induces a bijection

$$\mathscr{H}^1(\tau^*) \xrightarrow{\simeq} \prod_{v \in \mathrm{Ram}_\infty(D)} \{(p_v, q_v) \mid p_v + q_v = n \ \text{ and } \ q_v \ \text{ is even}\}.$$

The following identity can be easily verified:

$$\sum_{\eta\in\mathcal{H}^1(\tau^*)}\prod_{v\in\mathrm{Ram}_\infty(D)}\binom{n}{q_v}=\sum_{u_1,\dots,u_r=0}^{[\frac{n}{2}]}\prod_{i=1}^{r}\binom{n}{2u_i}=2^{r(n-1)}.$$

As a final step we substitute all results in formula (5) and observe:

$$\mathcal{L}(\tau^*,\Gamma(\mathfrak{a}),W)=2^{-r}\,\mathrm{N}(\mathfrak{a})^{n(2n+1)}\Delta_{\mathrm{rd}}(D)^{n(n+1)/2}\,\mathrm{Tr}(\tau^*|W)\prod_{j=1}^{n}M(j,\mathfrak{a},D).$$

Note that the Lefschetz number is nonzero precisely when F is totally real and $\mathrm{Tr}(\tau^*|W)$ does not vanish.

5E. ***The growth of the total Betti number.*** There are many recent results on the asymptotic behaviour of Betti numbers of arithmetic groups. Most of these results are upper bound results — a strong asymptotic upper bound was obtained by Calegari and Emerton [2009]. However, there are no strong lower bound results. It seems that the only available lower bound results are nonvanishing results for certain degrees in the cohomology. Indeed, there is a geometric method to construct cohomology classes in a given degree for cocompact arithmetic groups. This method originated from the work of Millson and Raghunathan [1981] and has been further elaborated by Rohlfs and Schwermer [1993]. Another result that can be interpreted as a result on lower bounds has been obtained by Venkataramana [2008]. In this last section we prove Corollary 1.1 to show that Lefschetz numbers provide asymptotic lower bounds for the total Betti number. The only remaining step is to relate the Lefschetz number to the index of the congruence subgroup $\Gamma(\mathfrak{a})$. Let F be a totally real number field. If D is totally definite we assume $n\geq 2$ such that $G=\mathrm{SL}_\Lambda$ has strong approximation.

Lemma 5.22. *The index* $[G(\mathbb{O}):\Gamma(\mathfrak{a})]$ *of* $\Gamma(\mathfrak{a})$ *in* $G(\mathbb{O})$ *is*

$$\mathrm{N}(\mathfrak{a})^{4n^2-1}\prod_{\substack{\mathfrak{p}\in\mathrm{Ram}_f(D)\\ \mathfrak{p}\nmid\mathfrak{a}}}\Bigl(\prod_{j=2}^{2n}\bigl(1-\mathrm{N}(\mathfrak{p})^{-j}\bigr)\Bigr)\prod_{\substack{\mathfrak{p}\in\mathrm{Ram}_f(D)\\ \mathfrak{p}\mid\mathfrak{a}}}\Bigl(\bigl(1+\mathrm{N}(\mathfrak{p})^{-1}\bigr)\prod_{j=2}^{n}\bigl(1-\mathrm{N}(\mathfrak{p})^{-2j}\bigr)\Bigr).$$

In particular, the term $[G(\mathbb{O}):\Gamma(\mathfrak{a})]N(\mathfrak{a})^{-4n^2+1}$ *is bounded from above and from below independent of* $\mathfrak{a}$,

$$\prod_{j=2}^{2n}\zeta_F(j)^{-1}\leq[G(\mathbb{O}):\Gamma(\mathfrak{a})]N(\mathfrak{a})^{-4n^2+1}$$
$$\leq\prod_{\mathfrak{p}\in\mathrm{Ram}_f(D)}\bigl(1+\mathrm{N}(\mathfrak{p})^{-1}\bigr).$$

Proof. Using the smoothness of the group scheme combined with strong approximation, there is a short exact sequence of groups

$$1 \longrightarrow \Gamma(\mathfrak{a}) \longrightarrow G(\mathbb{O}) \longrightarrow G(\mathbb{O}/\mathfrak{a}) \longrightarrow 1,$$

from which we deduce $[G(\mathbb{O}) : \Gamma(\mathfrak{a})] = \prod_{\mathfrak{p}|\mathfrak{a}} |G(\mathbb{O}_\mathfrak{p}/\mathfrak{a}\mathbb{O}_\mathfrak{p})|$. Let $\mathfrak{p}$ be a prime ideal which divides $\mathfrak{a}$, say $\nu_\mathfrak{p}(\mathfrak{a}) = e \geq 1$. Then $\mathbb{O}_\mathfrak{p}/\mathfrak{a}\mathbb{O}_\mathfrak{p} \cong \mathbb{O}_\mathfrak{p}/\mathfrak{p}^e\mathbb{O}_\mathfrak{p}$ and it follows from the smoothness of G that

$$|G(\mathbb{O}_\mathfrak{p}/\mathfrak{p}^e\mathbb{O}_\mathfrak{p})| = \mathrm{N}(\mathfrak{p})^{(e-1)d}|G(\mathbb{O}_\mathfrak{p}/\mathfrak{p}\mathbb{O}_\mathfrak{p})|,$$

where d is the dimension of the group $G \times_\mathbb{O} F$ (use [Oesterlé 1984, Section I.2.1]). The dimension of G is $d = 4n^2 - 1$.

If $\mathfrak{p} \in \mathrm{Ram}_f(D)$, then one can show that

$$|G(\mathbb{O}_\mathfrak{p}/\mathfrak{p}\mathbb{O}_\mathfrak{p})| = \mathrm{N}(\mathfrak{p})^{4n^2-1}\bigl(1 + \mathrm{N}(\mathfrak{p})^{-1}\bigr)\prod_{j=2}^{n}\bigl(1 - \mathrm{N}(\mathfrak{p})^{-2j}\bigr).$$

If otherwise $\mathfrak{p} \notin \mathrm{Ram}_f(D)$, then $G \times_\mathbb{O} \mathbb{O}_\mathfrak{p}$ is isomorphic to the special linear group SL_{2n}. We deduce that

$$|G(\mathbb{O}_\mathfrak{p}/\mathfrak{p}\mathbb{O}_\mathfrak{p})| = \mathrm{N}(\mathfrak{p})^{4n^2-1}\prod_{j=2}^{2n}\bigl(1 - \mathrm{N}(\mathfrak{p})^{-j}\bigr),$$

due to [Wilson 2009, Section 3.3.1]. Now the assertions can be readily verified. □

Proof of Corollary 1.1. Since $\Gamma_0(\mathfrak{a})$ is a subgroup of finite index in $\Gamma(\mathfrak{a})$, we obtain from [Serre 1979, Chapter VII, Proposition 6] that $b_i(\Gamma(\mathfrak{a})) \leq b_i(\Gamma_0(\mathfrak{a}))$. It follows directly from the main theorem that there is a positive real number $b > 0$, depending on F, D and n, such that

$$b\,\mathrm{N}(\mathfrak{a})^{n(2n+1)} \leq |\mathscr{L}(\tau^*, \Gamma(\mathfrak{a}), \mathbb{C})|$$

for every ideal $\mathfrak{a} \subseteq \mathbb{O}$ that makes $\Gamma(\mathfrak{a})$ torsion-free. Since $B(\Gamma(\mathfrak{a})) \geq |\mathscr{L}(\tau^*, \Gamma(\mathfrak{a}), \mathbb{C})|$, it follows from Lemma 5.22 that

$$B(\Gamma(\mathfrak{a})) \geq a[G(\mathbb{O}) : \Gamma(\mathfrak{a})]^{\frac{n(2n+1)}{4n^2-1}}$$

for some $a > 0$ depending on F, D and n. We obtain

$$\begin{aligned} B(\Gamma_0(\mathfrak{a})) \geq a[G(\mathbb{O}) : \Gamma(\mathfrak{a})]^{\frac{n(2n+1)}{4n^2-1}} &\geq a[G(\mathbb{O}) \cap \Gamma_0 : \Gamma_0(\mathfrak{a})]^{\frac{n(2n+1)}{4n^2-1}} \\ &= a\bigl([\Gamma_0 : G(\mathbb{O}) \cap \Gamma_0]^{-1}[\Gamma_0 : \Gamma_0(\mathfrak{a})]\bigr)^{\frac{n(2n+1)}{4n^2-1}}. \end{aligned}$$

We define $\kappa = a[\Gamma_0 : G(\mathbb{O}) \cap \Gamma_0]^{-\frac{n(2n+1)}{4n^2-1}}$. □

Acknowledgements. I would like to thank Professor J. Schwermer for his support during my thesis work, upon which this article is based.

References

[An and Wang 2008] J. An and Z. Wang, "Nonabelian cohomology with coefficients in Lie groups", *Trans. Amer. Math. Soc.* **360**:6 (2008), 3019–3040. MR 2009k:20116 Zbl 1194.22006

[Batyrev 1999] V. V. Batyrev, "Birational Calabi–Yau n-folds have equal Betti numbers", pp. 1–11 in *New trends in algebraic geometry* (Warwick, 1996), edited by K. Hulek et al., London Math. Soc. Lecture Note Ser. **264**, Cambridge Univ. Press, 1999. MR 2000i:14059 Zbl 0955.14028

[Borel 1963] A. Borel, "Some finiteness properties of adele groups over number fields", *Inst. Hautes Études Sci. Publ. Math.* 16 (1963), 5–30. MR 34 #2578 Zbl 0135.08902

[Borel and Serre 1973] A. Borel and J.-P. Serre, "Corners and arithmetic groups", *Comment. Math. Helv.* **48** (1973), 436–491. MR 52 #8337 Zbl 0274.22011

[Bourbaki 1963] N. Bourbaki, *Intégration. Chapitre* 7*: Mesure de Haar; Chapitre* 8*: Convolution et représentations*, Actualités Scientifiques et Industrielles **1306**, Hermann, Paris, 1963. MR 31 #3539 Zbl 0156.03204

[Calegari and Emerton 2009] F. Calegari and M. Emerton, "Bounds for multiplicities of unitary representations of cohomological type in spaces of cusp forms", *Ann. of Math.* (2) **170**:3 (2009), 1437–1446. MR 2011c:22032 Zbl 1195.22015

[Demazure and Gabriel 1970] M. Demazure and P. Gabriel, *Groupes algébriques, I: Géométrie algébrique, généralités, groupes commutatifs*, Masson, Paris, 1970. MR 46 #1800 Zbl 0203.23401

[Fainsilber and Morales 1999] L. Fainsilber and J. Morales, "An injectivity result for Hermitian forms over local orders", *Illinois J. Math.* **43**:2 (1999), 391–402. MR 2000g:11026 Zbl 0939.11020

[Grothendieck 1964] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, I", *Inst. Hautes Études Sci. Publ. Math.* 20 (1964), 5–259. MR 30 #3885 Zbl 0136.15901

[Harder 1971] G. Harder, "A Gauss–Bonnet formula for discrete arithmetically defined groups", *Ann. Sci. École Norm. Sup.* (4) **4** (1971), 409–455. MR 46 #8255 Zbl 0232.20088

[Harder 1975] G. Harder, "On the cohomology of SL(2, O)", pp. 139–150 in *Lie groups and their representations* (Budapest, 1971), edited by I. M. Gelfand, Halsted, New York, 1975. MR 54 #12977 Zbl 0395.57028

[Helgason 1978] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics **80**, Academic Press, New York, 1978. MR 80k:53081 Zbl 0451.53038

[Humphreys 1972] J. E. Humphreys, *Introduction to Lie algebras and representation theory*, Graduate Texts in Mathematics **9**, Springer, New York, 1972. MR 48 #2197 Zbl 0254.17004

[Katok 1992] S. Katok, *Fuchsian groups*, University of Chicago Press, Chicago, IL, 1992. MR 93d:20088 Zbl 0753.30001

[Kionke 2012] S. Kionke, *Lefschetz numbers of involutions on arithmetic subgroups of inner forms of the special linear group*, Ph.D. thesis, Universität Wien, Vienna, 2012.

[Kionke and Schwermer 2012] S. Kionke and J. Schwermer, "On the growth of the first Betti number of arithmetic hyperbolic 3-manifolds", preprint, 2012. arXiv 1204.3750

[Klingen 1962] H. Klingen, "Über die Werte der Dedekindschen Zetafunktion", *Math. Ann.* **145** (1962), 265–272. MR 24 #A3138 Zbl 0101.03002

[Knapp 2002] A. W. Knapp, *Lie groups beyond an introduction*, 2nd ed., Progress in Mathematics **140**, Birkhäuser, Boston, MA, 2002. MR 2003c:22001 Zbl 1075.22501

[Knus et al. 1998] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, American Mathematical Society Colloquium Publications **44**, Amer. Math. Soc., Providence, RI, 1998. MR 2000a:16031 Zbl 0955.16001

[Kottwitz 1988] R. E. Kottwitz, "Tamagawa numbers", *Ann. of Math.* (2) **127**:3 (1988), 629–646. MR 90e:11075 Zbl 0678.22012

[Krämer 1985] N. Krämer, *Beiträge zur Arithmetik imaginärquadratischer Zahlkörper*, Ph.D. thesis, Rheinische Friedrich-Wilhelms-Universität, Bonn, 1985.

[Lai 1991] K. F. Lai, "Lefschetz numbers and unitary groups", *Bull. Austral. Math. Soc.* **43**:2 (1991), 193–209. MR 92a:11063 Zbl 0724.11019

[Lee and Schwermer 1983] R. Lee and J. Schwermer, "The Lefschetz number of an involution on the space of harmonic cusp forms of SL_3", *Invent. Math.* **73**:2 (1983), 189–239. MR 84k:22016 Zbl 0525.10014

[Millson and Raghunathan 1981] J. J. Millson and M. S. Raghunathan, "Geometric construction of cohomology for arithmetic groups, I", *Proc. Indian Acad. Sci. Math. Sci.* **90**:2 (1981), 103–123. MR 83d:22008 Zbl 0524.22012

[Oesterlé 1984] J. Oesterlé, "Nombres de Tamagawa et groupes unipotents en caractéristique p", *Invent. Math.* **78**:1 (1984), 13–88. MR 86i:11016 Zbl 0542.20024

[Ono 1966] T. Ono, "On algebraic groups and discontinuous groups", *Nagoya Math. J.* **27** (1966), 279–322. MR 33 #7342 Zbl 0166.29802

[Platonov and Rapinchuk 1994] V. Platonov and A. Rapinchuk, *Algebraic groups and number theory*, Pure and Applied Mathematics **139**, Academic Press, Boston, MA, 1994. MR 95b:11039 Zbl 0841.20046

[Prasad 1989] G. Prasad, "Volumes of S-arithmetic quotients of semi-simple groups", *Inst. Hautes Études Sci. Publ. Math.* 69 (1989), 91–117. MR 91c:22023 Zbl 0695.22005

[Reiner 2003] I. Reiner, *Maximal orders*, London Mathematical Society Monographs (N.S.) **28**, Oxford University Press, 2003. MR 2004c:16026 Zbl 1024.16008

[Rohlfs 1978] J. Rohlfs, "Arithmetisch definierte Gruppen mit Galoisoperation", *Invent. Math.* **48**:2 (1978), 185–205. MR 80j:20043 Zbl 0391.14007

[Rohlfs 1981] J. Rohlfs, "The Lefschetz number of an involution on the space of classes of positive definite quadratic forms", *Comment. Math. Helv.* **56**:2 (1981), 272–296. MR 83a:10037 Zbl 0474.10019

[Rohlfs 1985] J. Rohlfs, "On the cuspidal cohomology of the Bianchi modular groups", *Math. Z.* **188**:2 (1985), 253–269. MR 86e:11042 Zbl 0535.20028

[Rohlfs 1990] J. Rohlfs, "Lefschetz numbers for arithmetic groups", pp. 303–313 in *Cohomology of arithmetic groups and automorphic forms* (Luminy-Marseille, 1989), edited by J.-P. Labesse and J. Schwermer, Lecture Notes in Math. **1447**, Springer, Berlin, 1990. MR 92d:11055 Zbl 0762.11023

[Rohlfs and Schwermer 1993] J. Rohlfs and J. Schwermer, "Intersection numbers of special cycles", *J. Amer. Math. Soc.* **6**:3 (1993), 755–778. MR 94a:11075 Zbl 0811.11039

[Rohlfs and Schwermer 1998] J. Rohlfs and J. Schwermer, "An arithmetic formula for a topological invariant of Siegel modular varieties", *Topology* **37**:1 (1998), 149–159. MR 98f:11044 Zbl 0926.11036

[Scharlau 1985] W. Scharlau, *Quadratic and Hermitian forms*, Grundlehren der Mathematischen Wissenschaften **270**, Springer, Berlin, 1985. MR 86k:11022 Zbl 0584.10010

[Sengün and Türkelli 2012] M. H. Sengün and S. Türkelli, “On the dimension of cohomology of Bianchi groups”, preprint, 2012. arXiv 1204.0470v2

[Serre 1971] J.-P. Serre, “Cohomologie des groupes discrets”, pp. 77–169 in *Prospects in mathematics* (Princeton, NJ, 1970), Ann. of Math. Studies **70**, Princeton Univ. Press, 1971. MR 52 #5876 Zbl 0235.22020

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Mathematics **67**, Springer, New York, 1979. MR 82e:12016 Zbl 0423.12016

[Serre 1994] J.-P. Serre, *Cohomologie galoisienne*, 5th ed., Lecture Notes in Mathematics **5**, Springer, Berlin, 1994. MR 96b:12010 Zbl 0812.12002

[Shimura 1963] G. Shimura, “Arithmetic of alternating forms and quaternion hermitian forms”, *J. Math. Soc. Japan* **15** (1963), 33–65. MR 26 #3694 Zbl 0121.28102

[Shimura 1971] G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan **11**, Princeton University Press, 1971. MR 47 #3318 Zbl 0221.10029

[Siegel 1969] C. L. Siegel, “Berechnung von Zetafunktionen an ganzzahligen Stellen”, *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1969** (1969), 87–102. MR 40 #5570 Zbl 0186.08804

[Venkataramana 2008] T. N. Venkataramana, “Virtual Betti numbers of compact locally symmetric spaces”, *Israel J. Math.* **166** (2008), 235–238. MR 2010a:22015 Zbl 1221.57040

[Weil 1995] A. Weil, *Basic number theory*, 2nd ed., Springer, Berlin, 1995. MR 96c:11002 Zbl 0823.11001

[Wilson 2009] R. A. Wilson, *The finite simple groups*, Graduate Texts in Mathematics **251**, Springer, London, 2009. MR 2011e:20018 Zbl 1203.20012

Received May 22, 2013.

STEFFEN KIONKE
MATHEMATISCHES INSTITUT
HEINRICH-HEINE-UNIVERSITÄT DÜSSELDORF
UNIVERSITÄTSSTR. 1
40225 DÜSSELDORF
GERMANY

steffen.kionke@uni-duesseldorf.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.415

CATEGORIFICATION OF A PARABOLIC HECKE MODULE VIA SHEAVES ON MOMENT GRAPHS

MARTINA LANINI

We investigate certain categories, associated by Fiebig with the geometric representation of a Coxeter system, via sheaves on Bruhat graphs. We modify Fiebig's definition of translation functors in order to extend it to the singular setting and use it to categorify a parabolic Hecke module. As an application we obtain a combinatorial description of indecomposable projective objects of (truncated) noncritical singular blocks of (a deformed version of) category $\mathbb{O}$, using indecomposable special modules over the structure algebra of the corresponding Bruhat graph.

1. Introduction

A typical problem in the representation theory of Kac–Moody algebras is to understand the composition series of standard objects in the corresponding category $\mathbb{O}$ of Bernstein, I. Gelfand and S. Gelfand [Bernstein et al. 1976]. In the case of a standard object lying in a regular block, this question is the core of the Kazhdan–Lusztig theory, and the answer is known to be given by the Kazhdan–Lusztig polynomials evaluated at the identity. If we consider a singular block, we only have to replace these polynomials by their parabolic analogue. In the case of a principal block, this fact was conjectured in [Kazhdan and Lusztig 1979] and proved in several steps in [Kazhdan and Lusztig 1980; Beilinson and Bernstein 1981; Brylinski and Kashiwara 1980]. A fundamental role in the proof of the Kazhdan–Lusztig conjecture was played by the geometric interpretation of the problem in terms of perverse sheaves and intersection cohomology complexes. In particular, one could study certain properties of the Hecke algebra in the category of equivariant perverse sheaves on the corresponding flag variety.

An alternative way to attack the Kazhdan–Lusztig conjecture is via Soergel bimodules, which provide a combinatorial realisation of projective objects in category $\mathbb{O}$. The combinatorial description of indecomposable projective objects we present in this paper is an analogue of the combinatorial construction of [Soergel 1990]

MSC2010: 17B67, 20C08.
Keywords: sheaves on moment graphs, parabolic Hecke module.

(introduced at first for finite-dimensional Lie algebras). The Soergel bimodule approach to the Kazhdan–Lusztig conjecture recently led to an algebraic proof of it [Elias and Williamson 2014].

The procedure of considering a complicated object, such as a category, in order to understand a simpler one is motivated by the fact that the extra structure can provide us with new tools and allow us to prove and hopefully generalise certain phenomena that are difficult to address directly.

Deodhar [1987] associated with any Coxeter system $(\mathcal{W}, \mathcal{S})$ and any subset of the set of simple reflections $J \subseteq \mathcal{S}$ the parabolic Hecke module $\boldsymbol{M}^J$. The aim of this paper is to give a categorification of this module for any J generating a finite subgroup.

We have followed the definition of categorification of $\boldsymbol{M}^J$ in [Mazorchuk and Stroppel 2005, Remark 7.8], which is actually a weak categorification. This could be strengthened to a proper categorification by presenting the result as a 2-representation of some 2-category (see [Mazorchuk 2012, Sections 1–3] for various levels of categorification and Remark 5.9 of this paper for a more precise statement). In [Mazorchuk and Stroppel 2008], the authors properly categorify induced cell modules (in the finite case), which is a huge step outside the parabolic Hecke module (the latter being just a special case).

If $\mathcal{W}$ is a Weyl group, there is a partial flag variety Y corresponding to the set J, equipped with an action of a maximal torus T, and as for the regular case, one possible categorification is given by the category of B-equivariant perverse sheaves on Y. Our goal is to describe a general categorification, which can be defined also in the case in which there is no geometry available. In order to do this, our main tools will be Bruhat moment graphs and sheaves on them. We will see how these objects come naturally into the picture.

Moment graphs appeared for the first time in [Goresky et al. 1998] as 1-skeletons of actions of tori on complex algebraic varieties. In particular, Goresky, Kottwitz and MacPherson were able to describe explicitly the equivariant cohomology of these varieties using only the data encoded in the underlying moment graphs. Inspired by this result, Braden and MacPherson [2001] could study the equivariant intersection cohomology of a complex algebraic variety equipped with a Whitney stratification, stable with respect to the torus action. In order to do so, they introduced the notion of sheaves on moment graphs and, in particular, of *canonical sheaves*. We will refer to this class of sheaves as *Braden–MacPherson sheaves*, or BMP sheaves.

Even if moment graphs arose originally from geometry, Fiebig [2008b] observed that it is possible to give an axiomatic definition of them. In particular, he associated a moment graph to any Coxeter datum $(\mathcal{W}, \mathcal{S}, J)$ as above and in the case of $J = \varnothing$, he used it to give an alternative construction of Soergel's category of bimodules associated to a reflection-faithful representation of $(\mathcal{W}, \mathcal{S})$ (see [Fiebig 2008b]).

(We refer the reader to [Williamson 2011] for the singular version of Soergel's bimodules.) The indecomposable objects of the category defined by Fiebig are precisely the BMP sheaves that, if $\mathcal{W}$ is a Weyl group, are related to the intersection cohomology complexes, the simple objects in the category of perverse sheaves. A fundamental step in Fiebig's realisation of this category were translation functors, whose definition we extend to the parabolic setting (see p. 426).

The paper is organised as follows:

In Section 2 we recall the definition of the parabolic Hecke module $\boldsymbol{M}^J$ and the fact that it is the unique free $\mathbb{Z}[v, v^{-1}]$-module having rank $|\mathcal{W}/\langle J\rangle|$ equipped with a certain structure of a module over the Hecke algebra $\boldsymbol{H}$. This structure is described in terms of the action of the Kazhdan–Lusztig basis elements $\underline{H}_s$, for $s \in \mathcal{S}$. Then by a *categorification* of $\boldsymbol{M}^J$ (as in [Mazorchuk and Stroppel 2005, Remark 7.8]) we mean a category $\mathscr{C}$, which is exact in the sense of Quillen [1973], together with an autoequivalence G and exact functors $\{F_s\}_{s\in\mathcal{S}}$ that provide the Grothendieck group $[\mathscr{C}]$ with the structure of a $\mathbb{Z}[v, v^{-1}]$-module and $\boldsymbol{H}$-module, such that there exists an isomorphism from $[\mathscr{C}]$ to the parabolic module, satisfying certain compatibility conditions with these functors coming from the defining properties of $\boldsymbol{M}^J$ (see definition on p. 420).

In the third section we introduce the objects we will be dealing with in the rest of the paper. In particular, we review basic concepts of the theory of moment graphs and sheaves on them.

Section 4 is about $\mathbb{Z}$-graded modules over $\mathscr{Z}^J$, the structure algebra of a parabolic Bruhat graph. In particular for any $s \in \mathcal{S}$, we define the translation functor ${}^s\theta$ and define the category $\mathscr{H}^J$ of *special* $\mathscr{Z}^J$*-modules*. By definition, this category is stable under the shift in degree that we denote by $\langle\cdot\rangle$ and under ${}^s\theta$ for all $s \in \mathcal{S}$.

In Section 5 we study certain subquotients of objects in $\mathscr{H}^J$, and this allows us to define an exact structure on $\mathscr{H}^J$ and hence to state our main theorem:

Theorem 5.8. *The category $\mathscr{H}^J$ special $\mathscr{Z}^J$-modules together with the shift in degree $\langle -1\rangle$ and (shifted) translation functors is a categorification of the parabolic Hecke module $\boldsymbol{M}^J$.*

Section 6 is devoted to the proof of this theorem. First we show that ${}^s\theta\circ\langle 1\rangle$ is an exact functor (Lemma 6.1). Secondly we define the character map $h^J : [\mathscr{H}^J] \to \boldsymbol{M}^J$ and prove that the functors $\langle -1\rangle$ and ${}^s\theta\circ\langle 1\rangle$, $s \in \mathcal{S}$, satisfy the desired compatibility condition (Proposition 6.2). We conclude then by showing that the character map is an isomorphism of $\mathbb{Z}[v, v^{-1}]$-modules (Lemma 6.3 and Lemma 6.6).

Section 7 is about the categorification of a certain injective map of $\boldsymbol{H}$-modules $i : \boldsymbol{M}^J \hookrightarrow \boldsymbol{H}$, which allows us to see the category $\mathscr{H}^J$ as a subcategory of $\mathscr{H}^\varnothing$. More precisely, we define an exact functor $I : \mathscr{H}^J \to \mathscr{H}^\varnothing$ such that the following diagram commutes:

$$\begin{tikzcd} {[\mathcal{H}^J]} \arrow[r, hook, "{[I]}"] \arrow[d, "h^J"'] & {[\mathcal{H}^\varnothing]} \arrow[d, "h^\varnothing"] \\ \boldsymbol{M}^J \arrow[r, hook, "i"'] & \boldsymbol{H} \end{tikzcd}$$

In order to construct and investigate the functor I, we give a realisation of $\mathcal{H}^J$ via BMP sheaves (Proposition 6.5) and then use Fiebig's idea [2008b] of interchanging global and local viewpoints.

In the last section we discuss briefly the relationship between $\mathcal{H}^J$ and noncritical blocks of an equivariant version of category $\mathcal{O}$ for symmetrisable Kac–Moody algebras. In particular, we show that the indecomposable projective objects of a truncated, noncritical block $\mathcal{O}_{R,\Lambda^{\leq \nu}}$ are combinatorially described by indecomposable modules in $\mathcal{H}^J$, with J depending on Λ (Proposition 8.3).

2. Hecke modules

Here we recall some classical constructions, following [Soergel 1997]. We close the section by defining the concept of categorification of the parabolic Hecke module $\boldsymbol{M}^J$.

Hecke algebra. The Hecke algebra associated to a Coxeter system $(\mathcal{W}, \mathcal{S})$ is nothing but a quantisation of the group ring $\mathbb{Z}[\mathcal{W}]$. Let $\leq$ be the Bruhat order on $\mathcal{W}$ and $\ell : \mathcal{W} \to \mathbb{Z}$ be the length function associated to $\mathcal{S}$. Denote by $\mathcal{L} := \mathbb{Z}[v, v^{-1}]$ the ring of Laurent polynomials in the variable v over $\mathbb{Z}$.

Definition 2.1. The Hecke algebra $\boldsymbol{H} = \boldsymbol{H}(\mathcal{W}, \mathcal{S})$ is the free $\mathcal{L}$-module having basis $\{H_x \mid x \in \mathcal{W}\}$, subject to the following relations, for $x \in \mathcal{W}$, $s \in \mathcal{S}$:

$$H_s H_x = \begin{cases} H_{sx} & \text{if } sx > s, \\ (v^{-1} - v) H_x + H_{sx} & \text{if } sx < x, \end{cases} \tag{1}$$

It is well known that this defines an associative $\mathcal{L}$-algebra [Humphreys 1990].

It is easy to verify that H_x is invertible for any $x \in \mathcal{W}$, and this allows us to define an involution on $\boldsymbol{H}$, that is, the unique ring homomorphism $\bar{\cdot} : \boldsymbol{H} \to \boldsymbol{H}$ such that $\bar{v} = v^{-1}$ and $\overline{H}_x = (H_{x^{-1}})^{-1}$.

Kazhdan and Lusztig [1979] showed the existence of another basis $\{\underline{H}_x\}$ for $\boldsymbol{H}$, the so-called *Kazhdan–Lusztig basis*, that they used to define complex representations of the Hecke algebra and hence of the Coxeter group. The entries of the change of basis matrix are given by a family of polynomials in $\mathbb{Z}[v]$, which are called *Kazhdan–Lusztig polynomials*. There are many different normalisations of this basis appearing in the literature. The one we adopt, following [Soergel 1997], is determined by Theorem 2.2 (see Remark 2.3).

Parabolic Hecke modules. Deodhar [1987] generalised this construction to the parabolic setting in the following way. Let $\mathcal{W}, \mathcal{S}$ and $\boldsymbol{H}$ be as above. Fix a subset $J \subseteq \mathcal{S}$ and denote by $\mathcal{W}_J = \langle J \rangle$ the subgroup of $\mathcal{W}$ generated by J. Since $(\mathcal{W}_J, J)$ is also a Coxeter system, it makes sense to consider its Hecke algebra $\boldsymbol{H}_J = \boldsymbol{H}(\mathcal{W}_J, J)$.

For any simple reflection $s \in \mathcal{S}$, the element H_s satisfies the quadratic relation $(H_s)^2 = (v^{-1} - v)H_s + H_e$, that is, $(H_s + v)(H_s - v^{-1}) = 0$. If $u \in \{v^{-1}, -v\}$, we may define a map of $\mathcal{L}$-modules $\varphi_u : \boldsymbol{H}_J \to \mathcal{L}$ by $H_s \mapsto u$. In this way, $\mathcal{L}$ is endowed with the structure of a $\boldsymbol{H}_J$-bimodule, which we denote by $\mathcal{L}(u)$.

The parabolic Hecke modules are then defined as $\boldsymbol{M}^J := \boldsymbol{H} \otimes_{\boldsymbol{H}_J} \mathcal{L}(v^{-1})$ and $\boldsymbol{N}^J := \boldsymbol{H} \otimes_{\boldsymbol{H}_J} \mathcal{L}(-v)$. As in the Hecke algebra case, it is possible to define an involutive automorphism of these modules. Namely,

$$\begin{aligned} \bar{\cdot} : \boldsymbol{H} \otimes_{\boldsymbol{H}_J} \mathcal{L}(u) &\to \boldsymbol{H} \otimes_{\boldsymbol{H}_J} \mathcal{L}(u), \\ H \otimes a &\mapsto \overline{H} \otimes \bar{a}. \end{aligned} \tag{2}$$

For $u \in \{v^{-1}, -v\}$, let $H_w^{J,u} := H_w \otimes 1 \in \mathcal{L}(u) \otimes_{\boldsymbol{H}_J} \boldsymbol{H}$. Denote by $\mathcal{W}^J$ the set of minimal length representatives of $\mathcal{W}/\mathcal{W}_J$.

Theorem 2.2 [Deodhar 1987].

(1) *For all $w \in \mathcal{W}^J$ there exists a unique element $\underline{H}_w^{J,v^{-1}} \in \boldsymbol{M}^J$ such that*

  (a) $\overline{\underline{H}}_w^{J,v^{-1}} = \underline{H}_w^{J,v^{-1}}$, *and*
  (b) $\underline{H}_w^{J,v^{-1}} = \sum_{y \in \mathcal{W}^J} m_{y,w}^J H_y^{J,v^{-1}}$,

  where the $m_{y,w}^J$ are such that $m_{w,w}^J = 1$ and $m_{y,w}^J \in v\mathbb{Z}[v]$ if $y \neq w$.

(2) *For all $w \in \mathcal{W}^J$ there exists a unique element $\underline{H}_w^{J,-v} \in \boldsymbol{N}^J$ such that*

  (a) $\overline{\underline{H}}_w^{J,-v} = \underline{H}_w^{J,-v}$, *and*
  (b) $\underline{H}_w^{J,-v} = \sum_{y \in \mathcal{W}^J} n_{y,w}^J H_y^{J,-v}$,

where the $n_{y,w}^J$ are such that $n_{w,w}^J = 1$ and $n_{y,w}^J \in v\mathbb{Z}[v]$ if $y \neq w$.

Remark 2.3. In the case $J = \varnothing$, the two parabolic modules coincide with the regular module: $\boldsymbol{M}^\varnothing = \boldsymbol{N}^\varnothing = \boldsymbol{H}$. Moreover $\underline{H}_w^{\varnothing,v^{-1}} = \underline{H}_w^{\varnothing,-v} = \underline{H}_w$ for all $w \in W$.

From now on, we will focus on the case $u = v^{-1}$; that is, we will deal only with $\boldsymbol{M}^J$. The action of the Hecke algebra $\boldsymbol{H}$ on $\boldsymbol{M}^J$ is defined as follows. Let $s \in \mathcal{S}$ be a simple reflection and let $x \in \mathcal{W}^J$; then we have (see [Soergel 1997, §3])

$$\underline{H}_s \cdot H_x^{J,v^{-1}} = \begin{cases} H_{sx}^{J,v^{-1}} + vH_x^{J,v^{-1}} & \text{if } sx \in \mathcal{W}^J,\ sx > x, \\ H_{sx}^{J,v^{-1}} + v^{-1}H_x^{J,v^{-1}} & \text{if } sx \in \mathcal{W}^J,\ sx < x, \\ (v+v^{-1})H_x^{J,v^{-1}} & \text{if } sx \notin \mathcal{W}^J. \end{cases} \tag{3}$$

Definition of the categorification of M^J. For any category $\mathscr{C}$ which is exact in the sense of [Quillen 1973], let us denote by $[\mathscr{C}]$ its Grothendieck group, that is, the abelian group with generators

$$[X] \quad \text{for } X \in \mathrm{Ob}(\mathscr{C})$$

and relations

$$[Y]=[X]+[Z] \qquad \text{for every exact sequence } 0 \to X \to Y \to Z \to 0.$$

For an exact endofunctor F on $\mathscr{C}$, denote by $[F]$ the induced endomorphism of $[\mathscr{C}]$.

By a *categorification* of M^J, we mean an exact category $\mathscr{C}$ together with an autoequivalence G and a family of exact endofunctors $\{F_s\}_{s\in\mathscr{S}}$ satisfying the following requirements:

(C1) $[\mathscr{C}]$ becomes an $\mathscr{L}$-module via $v^i \cdot [A] = [G^i A]$ for any $i \in \mathbb{Z}$, and there is an isomorphism $h^J : [\mathscr{C}] \xrightarrow{\sim} M^J$ of $\mathscr{L}$-modules.

(C2) For any simple reflection $s \in \mathscr{S}$, we have an isomorphism $GF_s \cong F_s G$ of functors.

(C3) For any simple reflection $s \in \mathscr{S}$, the following diagram commutes:

$$\begin{tikzcd} {[\mathscr{C}]} \arrow[r, "{[F_s]}"] \arrow[d, "h^J"'] & {[\mathscr{C}]} \arrow[d, "h^J"] \\ M^J \arrow[r, "\underline{H}_s \cdot"'] & M^J \end{tikzcd}$$

Remark 2.4. Our notion of M^J-categorification differs from the one in [Mazorchuk and Stroppel 2005, Remark 7.8]. Indeed we made the (weaker) requirement of $\mathscr{C}$ being exact instead of abelian. If we take the above categorification, restrict it to the additive category of projective objects and then abelianise it in the standard way, then this abelianisation is a 2-functor (see [Mazorchuk 2012, §3.3]) and will transform the above categorification into a categorification using abelian categories, in the spirit of [Mazorchuk and Stroppel 2005].

Remark 2.5. Williamson [2011] studied the 2-category of singular Soergel bimodules. A full tensor subcategory of it (${}^{\varnothing}\mathscr{B}^J$ in his notation) also provides a categorification of M^J.

The main goal of this paper is to construct such a categorification. In particular, we will generalise a categorification of the Hecke algebra obtained in [Fiebig 2011], which is known by results in [Fiebig 2008b] to be equivalent to the one via the bimodules of [Soergel 2007].

3. Sheaves on moment graphs

Definition 3.1 [Fiebig 2008b]. Let k be a field, V a finite-dimensional k-vector space, and $\mathbb{P}(V)$ the corresponding projective space. A *V-moment graph* is given by a tuple $(\mathcal{V}, \mathcal{E}, \trianglelefteq, l)$ satisfying these conditions:

(MG1) $(\mathcal{V}, \mathcal{E})$ is a graph with a set of vertices $\mathcal{V}$ and a set of edges $\mathcal{E}$.

(MG2) $\trianglelefteq$ is a partial order on $\mathcal{V}$ such that $x, y \in \mathcal{V}$ are comparable if they are linked by an edge.

(MG3) $l : \mathcal{E} \to \mathbb{P}(V)$ is a map, called the *label function*.

Remark 3.2. This is the traditional definition. We note that the fact that $\mathcal{V}$ is equipped with a partial order (similarly to the notion of quasihereditary algebra) is used only in the definition of Braden–MacPherson sheaves.

As in [Fiebig 2008b], we think of the order as giving each edge a direction: we write $E : x \to y \in \mathcal{E}$ if $x \leq y$. We write $x \,—\, y$ or $y \,—\, x$ if we want to ignore the order.

Bruhat graphs. Let $(\mathcal{W}, \mathcal{S})$ be a Coxeter system and denote by m_{st} the order of the product of two simple reflections $s, t \in \mathcal{S}$. Let V be the geometric representation of $(\mathcal{W}, \mathcal{S})$ (see [Humphreys 1990, §5.3]). Then V is a real vector space with basis indexed by the set of simple reflections $\Pi = \{\alpha_s\}_{s\in\mathcal{S}}$ and s acts on V by

$$s : v \mapsto v - 2\langle v, \alpha_s\rangle\alpha_s,$$

where $\langle\cdot,\cdot\rangle : V \times V \longrightarrow \mathbb{R}$ is the symmetric bilinear form given by

$$\langle\alpha_s, \alpha_t\rangle = \begin{cases} -\cos(\pi/m_{st}) & \text{if } m_{st} \neq \infty, \\ -1 & \text{if } m_{st} = \infty. \end{cases}$$

Consider a subset $J \subseteq \mathcal{S}$ and keep the same notation as in the previous section. Choose $\lambda \in V$ such that $\mathcal{W}_J = \mathrm{Stab}_{\mathcal{W}}(\lambda)$. Then $\mathcal{W}^J$ can be identified with the orbit $\mathcal{W} \cdot \lambda$ via $x \mapsto x(\lambda)$.

Recall that the set of reflections $\mathcal{T}$ of $\mathcal{W}$ is

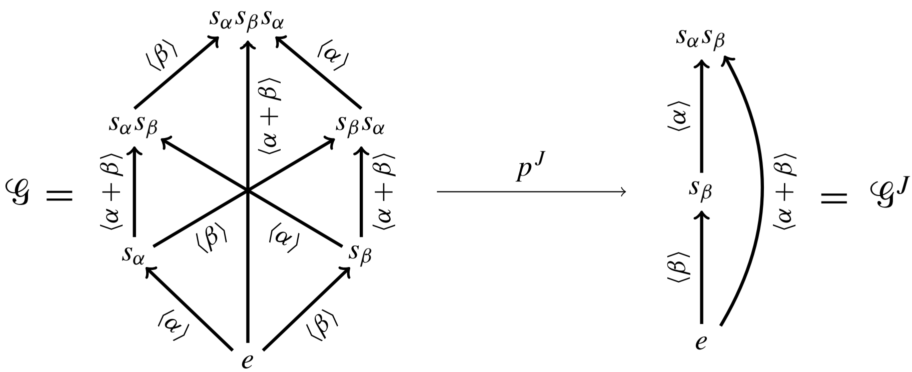
$$\mathcal{T} = \{wsw^{-1} \mid s \in \mathcal{S},\ w \in \mathcal{W}\}.$$

Definition 3.3 [Fiebig 2008b, §2.2]. The *Bruhat moment graph* $\mathcal{G}^J$ associated to the Coxeter datum $(\mathcal{W}, \mathcal{S}, J)$ is the following V-moment graph:

- The set of vertices is given by $\mathcal{W} \cdot \lambda \leftrightarrow \mathcal{W}^J$, and $x \to y$ is an edge if and only if $\ell(x) < \ell(y)$ and there exists a reflection $t \in \mathcal{T}$ such that $x(\lambda) = ty(\lambda)$, that is, $y = txw$ for some $w \in \mathcal{W}_J$ and $y \notin x\mathcal{W}_J$.
- The partial order W^J is the (induced) Bruhat order.
- $l(x \to txw)$ is given by the line generated by $x(\lambda) - tx(\lambda)$ in $\mathbb{P}(V)$.

Consider now two Bruhat moment graphs on V, namely $\mathcal{G} = \mathcal{G}(\mathcal{W}, \mathcal{S}, \varnothing)$ and $\mathcal{G}^J = \mathcal{G}(\mathcal{W}, \mathcal{S}, J)$. The *canonical quotient map* $p^J : \mathcal{G} \to \mathcal{G}^J$ is induced by the map $p^J_{\mathcal{V}} : x \to x^J$, with x^J a minimal length representative of the coset $x\mathcal{W}_J$.

Example 3.4. Let $\mathcal{W} = S_3$, the symmetric group on three letters. In this case we have $V = \mathbb{R}^2$, $\Pi = \{\alpha, \beta\}$, and the angle between the two roots is $2\pi/3$. If we fix $J = \{s_\alpha\}$, then p^J is as follows.



We have

$$p^J_{\mathcal{V}}(e) = p^J_{\mathcal{V}}(s_\alpha) = e,$$
$$p^J_{\mathcal{V}}(s_\beta) = p^J_{\mathcal{V}}(s_\beta s_\alpha) = s_\beta,$$
$$p^J_{\mathcal{V}}(s_\alpha s_\beta) = p^J_{\mathcal{V}}(s_\alpha s_\beta s_\alpha) = s_\alpha s_\beta.$$

Sheaves on a V-moment graph.

Conventions. For any finite-dimensional vector space V over the field k (with char $k \neq 2$), we denote by $S = \mathrm{Sym}(V)$ its symmetric algebra. Then S is a polynomial ring and we provide it with the grading induced by setting $S_{\{2\}} = V$. From now on, all the S-modules will be finitely generated and $\mathbb{Z}$-graded. Moreover we will consider only degree-zero morphisms between them. For a graded S-module $M = \bigoplus_i M_{\{i\}}$ and for $j \in \mathbb{Z}$, we denote by $M\langle j\rangle$ the $\mathbb{Z}$-graded S-module obtained from M by shifting the grading by j, that is, $(M\langle j\rangle)_{\{i\}} = M_{\{j+i\}}$.

Definition 3.5 [Braden and MacPherson 2001]. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \trianglelefteq, l)$ be a V-moment graph. A *sheaf $\mathcal{F}$ on $\mathcal{G}$* consists of $(\{\mathcal{F}^x\}, \{\mathcal{F}^E\}, \{\rho_{x,E}\})$ satisfying these conditions:

(SH1) $\mathcal{F}^x$ is an S-module for all $x \in \mathcal{V}$.

(SH2) $\mathcal{F}^E$ is an S-module such that $l(E) \cdot \mathcal{F}^E = \{0\}$ for all $E \in \mathcal{E}$.

(SH3) $\rho_{x,E} : \mathcal{F}^x \to \mathcal{F}^E$ is a homomorphism of S-modules, for all $x \in \mathcal{V}$, $E \in \mathcal{E}$ with x incident to the edge E.

Remark 3.6. We may consider the following topology on the space $\Gamma = \mathcal{V} \cup \mathcal{E}$ (see [Braden and MacPherson 2001, §1.3]). We say that a subset $O \subseteq \Gamma$ is open if whenever a vertex x is in O, then all edges adjacent to x are also in O. With this topology, the object in Definition 3.5 is a sheaf of S-modules on Γ in the usual sense. For our purposes, it will be sufficient to consider sheaves as purely combinatorial, algebraic objects.

Example 3.7 [Braden and MacPherson 2001, §1]. The *structure sheaf* $\mathcal{Z}$ of V-moment graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \trianglelefteq, l)$ is defined as follows:

- $\mathcal{Z}^x = S$ for all $x \in \mathcal{V}$.
- $\mathcal{Z}^E = S/l(E) \cdot S$ for all $E \in \mathcal{E}$.
- $\rho_{x,E} : S \to S/l(E) \cdot S$ is the canonical quotient map, for all $x \in \mathcal{V}$ and $E \in \mathcal{E}$ such that x is incident to the edge E.

Definition 3.8 [Fiebig 2009]. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \trianglelefteq, l)$ be a V-moment graph and let $\mathcal{F} = (\{\mathcal{F}^x\}, \{\mathcal{F}^E\}, \{\rho_{x,E}\})$, $\mathcal{F}' = (\{\mathcal{F}'^x\}, \{\mathcal{F}'^E\}, \{\rho'_{x,E}\})$ be two sheaves on $\mathcal{G}$. A *morphism* $\varphi : \mathcal{F} \longrightarrow \mathcal{F}'$ is given by the following data:

(MSH1) $\varphi^x : \mathcal{F}^x \to \mathcal{F}'^x$ is a homomorphism of S-modules, for all $x \in \mathcal{V}$.

(MSH2) $\varphi^E : \mathcal{F}^E \to \mathcal{F}'^E$ is a homomorphism of S-modules, for all $E \in \mathcal{E}$, and if $x \in \mathcal{V}$ is incident to the edge E, the following diagram commutes:

$$\begin{tikzcd}
\mathcal{F}^x \arrow[r, "\rho_{x,E}"] \arrow[d, "\varphi^x"'] & \mathcal{F}^E \arrow[d, "\varphi^E"] \\
\mathcal{F}'^x \arrow[r, "\rho'_{x,E}"'] & \mathcal{F}'^E
\end{tikzcd}$$

Definition 3.9. Let $\mathcal{G}$ be a V-moment graph. We denote by $\mathbf{Sh}(\mathcal{G})$ the category of sheaves on $\mathcal{G}$ and corresponding morphisms.

Remark 3.10. The category of sheaves on $\mathcal{G}$ is graded, with the shift of grading autoequivalence $\langle 1 \rangle : \mathbf{Sh}(\mathcal{G}) \to \mathbf{Sh}(\mathcal{G})$ given by

$$(\{\mathcal{F}^x\}, \{\mathcal{F}^E\}, \{\rho_{x,E}\}) \mapsto (\{\mathcal{F}^x \langle 1 \rangle\}, \{\mathcal{F}^E \langle 1 \rangle\}, \{\rho_{x,E} \circ \langle 1 \rangle\}).$$

Moreover $\mathbf{Sh}(\mathcal{G})$ is an additive category, with zero object $(\{0\}, \{0\}, \{0\})$, biproduct given by

$$(\{\mathcal{F}^x\}, \{\mathcal{F}^E\}, \{\rho_{x,E}\}) \oplus (\{\mathcal{F}'^x\}, \{\mathcal{F}'^E\}, \{\rho'_{x,E}\}) = (\{\mathcal{F}^x \oplus \mathcal{F}'^x\}, \{\mathcal{F}^E \oplus \mathcal{F}'^E\}, \{(\rho_{x,E}, \rho'_{x,E})\}),$$

and idempotent split.

Sections of a sheaf on a moment graph. Even if $\mathbf{Sh}(\mathcal{G})$ is not a category of sheaves in the usual sense, we may define the notion of sections following [Fiebig 2008a].

Definition 3.11. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \trianglelefteq, l)$ be a V-moment graph,

$$\mathcal{F} = (\{\mathcal{F}^x\}, \{\mathcal{F}^E\}, \{\rho_{x,E}\}) \in \mathbf{Sh}(\mathcal{G}),$$

and $\mathcal{I} \subseteq \mathcal{V}$. The *set of sections of* $\mathcal{F}$ *over* $\mathcal{I}$ is denoted by $\Gamma(\mathcal{I}, \mathcal{F})$ and defined as

$$\Gamma(\mathcal{I}, \mathcal{F}) := \left\{ (m_x) \in \prod_{x \in \mathcal{I}} \mathcal{F}^x \;\middle|\; \begin{array}{c} \rho_{x,E}(m_x) = \rho_{y,E}(m_y) \\ \text{for all } E : x \text{---} y \in \mathcal{E},\ x, y \in \mathcal{I} \end{array} \right\}.$$

We denote by $\Gamma(\mathcal{F}) := \Gamma(\mathcal{V}, \mathcal{F})$ the set of *global sections* of $\mathcal{F}$.

Example 3.12. A very important example is given by the set of global sections of the structure sheaf (see Example 3.7). In this case, we get the *structure algebra*:

$$\mathcal{Z} := \Gamma(\mathcal{Z}) = \left\{ (z_x)_{x \in \mathcal{V}} \in \prod_{x \in \mathcal{V}} S \;\middle|\; \begin{array}{c} z_x - z_y \in l(E) \cdot S \\ \text{for all } E : x \text{---} y \in \mathcal{E} \end{array} \right\}.$$

Remark 3.13. The algebra $\mathcal{Z}$ should be thought of as the centre of a noncritical block in the deformed category $\mathcal{O}$ (see [Fiebig 2003, Theorem 3.6]).

It is easy to check that $\mathcal{Z}$, equipped with componentwise addition and multiplication, is an algebra and that there is an action of S on it by diagonal multiplication. Moreover for any sheaf $\mathcal{F} \in \mathbf{Sh}(\mathcal{G})$, the structure algebra $\mathcal{Z}$ acts on the space $\Gamma(\mathcal{F})$ via componentwise multiplication, so Γ defines a functor from the category of sheaves on $\mathcal{G}$ to the category of $\mathcal{Z}$-modules:

$$\Gamma : \mathbf{Sh}(\mathcal{G}) \to \mathcal{Z}\text{-mod}. \tag{4}$$

BMP-sheaves. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \trianglelefteq, l)$ be a V-moment graph. For all $\mathcal{F} \in \mathbf{Sh}(\mathcal{G})$ and $x \in \mathcal{V}$, we set

$$\mathcal{E}_{\delta x} := \{E \in \mathcal{E} \mid \text{there is } y \in \mathcal{V} \text{ with } E : x \to y\},$$
$$\mathcal{V}_{\delta x} := \{y \in \mathcal{V} \mid \text{there is } E \in \mathcal{E}_{\delta x} \text{ with } E : x \to y\}.$$

Additionally for any $x \in \mathcal{V}$ denote $\{\triangleright x\} = \{y \in \mathcal{V} \mid y \triangleright x\}$ and define $\mathcal{F}^{\delta x}$ as the image of $\Gamma(\{\triangleright x\}, \mathcal{F})$ under the composition of the following functions:

$$u_x : \Gamma(\{\triangleright x\}, \mathcal{F}) \longrightarrow \bigoplus_{y \triangleright x} \mathcal{F}^y \longrightarrow \bigoplus_{y \in \mathcal{V}_{\delta x}} \mathcal{F}^y \xrightarrow{\oplus \rho_{y,E}} \bigoplus_{E \in \mathcal{E}_{\delta x}} \mathcal{F}^E.$$

Theorem 3.14 [Braden and MacPherson 2001]. *Suppose* $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \trianglelefteq, l)$ *is a* V*-moment graph and let* $w \in \mathcal{V}$*. There exists a unique up to isomorphism indecomposable sheaf* $\mathcal{B}(w)$ *on* $\mathcal{G}$ *with the following properties*:

(BMP1) *If* $x \in \mathcal{V}$, *then* $\mathcal{B}(w)^x \cong 0$, *unless* $x \trianglelefteq w$. *Moreover* $\mathcal{B}(w)^w \cong S$.

(BMP2) *If* $x, y \in \mathcal{V}$ *and* $E : x \to y \in \mathcal{E}$, *then the map* $\rho_{y,E} : \mathcal{B}(w)^y \to \mathcal{B}(w)^E$ *is surjective*, *with kernel* $l(E) \cdot \mathcal{B}(w)^y$.

(BMP3) *If* $x, y, w \in \mathcal{V}$, $x \neq w$ *and* $E : x \to y \in E$, *then* $\rho_{\delta x} := \bigoplus_{E \in \mathcal{E}_{\delta x}} \rho_{x,E} : \mathcal{B}(w)^x \to \mathcal{B}(w)^{\delta x}$ *is a projective cover in the category of graded* S-*modules.*

We call $\mathcal{B}(w)$ the *BMP sheaf.*

4. Modules over the structure algebra

Let $\mathcal{Z}$ be the structure algebra (see p. 424) of a regular Bruhat graph $\mathcal{G} = \mathcal{G}(\mathcal{W}, \varnothing)$ and denote by $\mathcal{Z}\text{-mod}^{\mathrm{f}}$ the category of $\mathbb{Z}$-graded $\mathcal{Z}$-modules that are torsion-free and finitely generated over S. Fiebig [2008b] defined translation functors on the category $\mathcal{Z}\text{-mod}^{\mathrm{f}}$. Using these, he defined inductively a full subcategory $\mathcal{H}$ of $\mathcal{Z}$-mod and proved that $\mathcal{H}$, in characteristic zero, is equivalent to a category of bimodules introduced by Soergel [2007]. In [Fiebig 2011], it is shown that $\mathcal{H}$ categorifies the Hecke algebra $\boldsymbol{H}$ (and the periodic module $\boldsymbol{M}$), using translation functors. The aim of this section is to define translation functors in the parabolic setting and to extend some results of [Fiebig 2011].

Let $\mathcal{W}$ be a Weyl group, let $\mathcal{S}$ be its set of simple reflections and let $J \subseteq \mathcal{S}$. Hereafter we will keep the notation we used in Section 2. Recall that for any $z \in \mathcal{W}$, there is a unique factorisation $x = x^J x_J$, with $x^J \in \mathcal{W}^J$, $x_J \in \mathcal{W}_J$ and $\ell(x) = \ell(x^J) + \ell(x_J)$ (see [Björner and Brenti 2005, Proposition 2.4.4]).

In [Fiebig 2008b], for all $s \in \mathcal{S}$, an involutive automorphism σ_s of the structure algebra of a regular Bruhat graph is given. In a similar way, we will define an involution ${}_s\sigma$ for a fixed simple reflection $s \in \mathcal{S}$ on the structure algebra $\mathcal{Z}^J$ of the parabolic Bruhat moment graph $\mathcal{G}^J$.

Let $x, y \in \mathcal{W}^J$. Notice that $l(x \text{ --- } y) = \alpha_t$ if and only if $l(sx \text{ --- } sy) = s(\alpha_t)$ because $sxw(sy)^{-1} = sxwy^{-1}s = sts$, for some $w \in \mathcal{W}_J$.

Denote by τ_s the automorphism of the symmetric algebra S induced by the mapping $\lambda \mapsto s(\lambda)$ for all $\lambda \in V$. For $(z_x)_{x \in \mathcal{W}^J} \in \mathcal{Z}^J$, we set ${}_s\sigma\big((z_x)_{x \in \mathcal{W}^J}\big) = (z'_x)_{x \in \mathcal{W}^J}$, where $z'_x := \tau_s(z_{(sx)^J})$. This is again an element of the structure algebra from what we have observed above.

Let us fix the following notation:

- ${}^s\mathcal{Z}^J$ denotes the space of invariants with respect to ${}_s\sigma$;
- ${}^{-s}\mathcal{Z}^J$ denotes the space of anti-invariants with respect to ${}_s\sigma$.

We denote by $\overline{\alpha}_s$ the element of $\mathcal{Z}^J$ whose components are all equal to α_s. We obtain the following decomposition of $\mathcal{Z}^J$ as a ${}^s\mathcal{Z}^J$-module:

Lemma 4.1.
$$\mathcal{Z}^J = {}^s\mathcal{Z}^J \oplus \overline{\alpha}_s \cdot {}^s\mathcal{Z}^J.$$

Proof. Because ${}_s\sigma$ is an involution, we get $\mathcal{Z}^J = {}^s\mathcal{Z}^J \oplus {}^{-s}\mathcal{Z}^J$. Since $\overline{\alpha}_s \in \mathcal{Z}^J$ and $s(\alpha_s) = -\alpha_s$, it follows that ${}_s\sigma(\overline{\alpha}_s) = -\overline{\alpha}_s$ and so $\overline{\alpha}_s \cdot {}^s\mathcal{Z}^J \subseteq {}^{-s}\mathcal{Z}^J$ and we now have to prove the other inclusion, that is, every element $z \in {}^{-s}\mathcal{Z}^J$ is divisible by $\overline{\alpha}_s$ in ${}^{-s}\mathcal{Z}^J$.

If $z = (z_x) \in {}^{-s}\mathcal{Z}^J$, then for all $x \in \mathcal{W}^J$,

$$z_x = -\tau_s(z_{(sx)^J}) \equiv -z_{(sx)^J} \pmod{\alpha_s}.$$

On the other hand,

$$z_x \equiv z_{(sx)^J} \pmod{\alpha_s}.$$

It follows that $2z_x \equiv 0 \pmod{\alpha_s}$, that is, α_s divides z_x in S.

It remains to verify that $z' := \overline{\alpha}_s{}^{-1} \cdot z \in \mathcal{Z}$, that is, $z'_x - z'_{(tx)^J} \equiv 0 \pmod{\alpha_t}$ for any $x \in \mathcal{W}^J$ and $t \in \mathcal{T}$. If $(tx)^J = (sx)^J$ there is nothing to prove since α_s divides $z'_x = z_{(sx)^J}$ and $z'_{(sx)^J} = z_x$ and hence also their difference. On the other hand, if $(tx)^J \neq (sx)^J$ we get the following:

$$\alpha_s \cdot (z'_x - z'_{(tx)^J}) = z_x - z_{(tx)^J} \equiv 0 \pmod{\alpha_t}.$$

Since α_s and α_t are linearly independent, $\alpha_s \not\equiv 0 \pmod{\alpha_t}$ and we obtain

$$z'_x - z'_{(tx)^J} \equiv 0 \pmod{\alpha_t}. \qquad \square$$

Translation functors and special modules. In order to define translation functors, we need an action of S on ${}^s\mathcal{Z}^J$ and $\mathcal{Z}^J$.

Lemma 4.2. *For any $\lambda \in V$ and any $x \in \mathcal{W}^J$, let us set*

$$c(\lambda)^J_x := \sum_{x_J \in \mathcal{W}_J} x x_J(\lambda). \tag{5}$$

Then $c(\lambda)^J := (c(\lambda)^J_x)_{x \in \mathcal{W}^J} \in {}^s\mathcal{Z}^J$.

Proof. First we prove that $c(\lambda)^J \in \mathcal{Z}^J$, that is, $c(\lambda)^J_x - c(\lambda)^J_{(tx)^J} \equiv 0 \pmod{\alpha_t}$. Since for any x_J there exists an element y_J such that $x x_J = t\,(tx)^J\,y_J$, we obtain

$$\begin{aligned}
\sum_{x_J \in \mathcal{W}_J} x x_J(\lambda) - \sum_{x_J \in \mathcal{W}_J} (tx)^J\, x_J(\lambda) &= \sum_{y_J \in \mathcal{W}_J} t\,(tx)^J\, y_J(\lambda) - \sum_{y_J \in \mathcal{W}_J} (tx)^J\, y_J(\lambda) \\
&= t\Big(\sum_{y_J \in \mathcal{W}_J} (tx)^J\, y_J(\lambda)\Big) - \sum_{y_J \in \mathcal{W}_J} (tx)^J\, y_J(\lambda) \\
&= \Big(\sum_{y_J \in \mathcal{W}_J} 2\big\langle (tx)^J\, y_J(\lambda), \alpha_t\big\rangle\Big)\alpha_t \\
&\equiv 0 \pmod{\alpha_t}.
\end{aligned}$$

To conclude it is left to show that $c(\lambda)^J$ is invariant with respect to ${}_s\sigma$. For any $x \in \mathcal{W}^J$, one has

$$\tau_s\big(c(\lambda)^J_x\big) = \tau_s\Big(\sum_{x_J \in \mathcal{W}_J} x x_J(\lambda)\Big) = \sum_{x_J \in \mathcal{W}_J} s x x_J(\lambda) = c(\lambda)^J_{sx}.$$

Hence we have ${}_s\sigma(c(\lambda)^J) = (\tau_s c(\lambda)^J_{sx})_{x\in\mathcal{W}^J} = c(\lambda)^J$. □

For any $x \in \mathcal{W}^J$, denote by η_x the map of (free) $S^{\mathcal{W}_J}$-modules $S \to S^{\mathcal{W}_J}$ induced by the map $\lambda \mapsto c(\lambda)^J_x$ for all $\lambda \in V$. Now by Lemma 4.2, the action of S on $\mathfrak{L}^J$ given by

$$p.(z_x)_{x\in\mathcal{W}^J} = (\eta_x(p) z_x), \quad p \in S\,,\ z \in \mathfrak{L}^J, \tag{6}$$

preserves ${}^s\mathfrak{L}^J$. Thus any $\mathfrak{L}^J$-module or ${}^s\mathfrak{L}^J$-module has an S-module structure as well. Suppose $\mathfrak{L}^J$-mod$^{\mathrm{f}}$, respectively, ${}^s\mathfrak{L}^J$-mod$^{\mathrm{f}}$, is the category of $\mathbb{Z}$-graded $\mathfrak{L}^J$-modules, respectively, ${}^s\mathfrak{L}^J$-modules, that are torsion-free and finitely generated over S, respectively, $S^{\mathcal{W}_J}$-modules.

The *translation on the wall* is the functor ${}^{s,\mathrm{on}}\theta : \mathfrak{L}^J$ -mod $\to {}^s\mathfrak{L}^J$-mod defined by the mapping $M \mapsto \mathrm{Res}^{{}^s\mathfrak{L}^J}_{\mathfrak{L}^J} M$.

The *translation out of the wall* is the functor ${}^{s,\mathrm{out}}\theta : {}^s\mathfrak{L}^J$ -mod $\to \mathfrak{L}^J$-mod defined by the mapping $N \mapsto \mathrm{Ind}^{{}^s\mathfrak{L}^J}_{\mathfrak{L}^J} N = \mathfrak{L}^J \otimes_{{}^s\mathfrak{L}^J} N$. Observe that this functor is well defined due to Lemma 4.1.

By composition, we get a functor ${}^s\theta^J := {}^{s,\mathrm{out}}\theta \circ {}^{s,\mathrm{on}}\theta : \mathfrak{L}^J$-mod$\to \mathfrak{L}^J$-mod that we call the *(left) translation functor*.

Remark 4.3. This construction is very similar to the one in [Soergel 1990], where translation functors are defined in the finite case for the coinvariant algebra.

Remark 4.4. One could consider the idempotent split additive tensor category generated by the translation functors we defined above and describe indecomposable projective objects. This would be useful in order to strengthen our main result to a proper categorification (see Remark 5.9). In this paper we are not going to investigate this category of translation functors but the one of special modules, defined on p. 428.

The following proposition describes the first properties of ${}^s\theta$:

Proposition 4.5. (1) *The functors from ${}^s\mathfrak{L}^J$-mod to $\mathfrak{L}^J$-mod mapping M to $\mathfrak{L}^J\langle 2\rangle \otimes_{{}^s\mathfrak{L}^J} M$ and* $\mathrm{Hom}_{{}^s\mathfrak{L}^J}(\mathfrak{L}^J, M)$, *respectively, are isomorphic.*

(2) *The functor ${}^s\theta = \mathfrak{L}^J \otimes_{{}^s\mathfrak{L}^J} - : \mathfrak{L}^J$-mod $\to \mathfrak{L}^J$-mod is self-adjoint up to a shift.*

Proof. (1) Let $M \in {}^s\mathcal{Z}^J$-mod. We must prove that $\mathcal{Z}^J\langle 2\rangle \otimes_{{}^s\mathcal{Z}^J} M \cong \mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, M)$ as $\mathcal{Z}^J$-modules.

First we show that $\mathcal{Z}^J\langle 2\rangle \cong \mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, {}^s\mathcal{Z}^J)$ as ${}^s\mathcal{Z}^J$-modules. By Lemma 4.1, $\{\bar{1}, \bar{\alpha}_s\}$ is a ${}^s\mathcal{Z}^J$-basis for $\mathcal{Z}^J$. Let $\bar{1}^*, \bar{\alpha}_s{}^* \in \mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, {}^s\mathcal{Z}^J)$ be the ${}^s\mathcal{Z}^J$-basis dual to $\bar{1}$ and $\bar{\alpha}_s$, that is,

$$\bar{1}^*(\bar{1}) = \bar{1}, \quad \bar{1}^*(\bar{\alpha}_s) = \bar{0}, \quad \bar{\alpha}_s{}^*(\bar{\alpha}_s) = \bar{1}, \quad \bar{\alpha}_s{}^*(\bar{1}) = \bar{0},$$

where $\bar{1} \in {}^s\mathcal{Z}^J$, respectively, $\bar{0} \in {}^s\mathcal{Z}^J$, is the section with 1, respectively, 0, in all entries. Since $\deg(1) - 2 = -2 = \deg(\bar{\alpha}_s{}^*)$ and $\deg(\bar{\alpha}_s) - 2 = 0 = \deg \bar{1}^*$, we have an isomorphism of ${}^s\mathcal{Z}^J$-modules $\mathcal{Z}^J\langle 2\rangle \cong \mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, {}^s\mathcal{Z}^J)$ defined by the mapping

$$\bar{1} \mapsto \bar{\alpha}_s{}^*, \quad \bar{\alpha}_s \mapsto \bar{1}^*.$$

Because $\mathcal{Z}^J$ is free of rank two over ${}^s\mathcal{Z}^J$,

$$\mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, M) \cong \mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, {}^s\mathcal{Z}^J) \otimes_{{}^s\mathcal{Z}^J} M$$

by the map

$$\varphi \mapsto \bar{\alpha}_s{}^* \otimes \varphi(\bar{\alpha}_s) + \bar{1}^* \otimes \varphi(\bar{1}).$$

This conclude the proof of (1).

(2) Since $\mathcal{Z}^J \otimes_{{}^s\mathcal{Z}^J} -$ and $\mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, -)$ are, respectively, left- and right-adjoint to the restriction functor, we obtain the following chain of isomorphisms for any pair $M, N \in \mathcal{Z}^J$:

$$\begin{aligned}
\mathrm{Hom}_{\mathcal{Z}^J}({}^s\theta M, N) &= \mathrm{Hom}_{\mathcal{Z}^J}\big(\mathcal{Z}^J \otimes_{{}^s\mathcal{Z}^J} (\mathrm{Res}^{{}^s\mathcal{Z}^J}_{\mathcal{Z}^J} M), N\big) \\
&\cong \mathrm{Hom}_{\mathcal{Z}^J}\big(\mathrm{Res}^{{}^s\mathcal{Z}^J}_{\mathcal{Z}^J} M, \mathrm{Res}^{{}^s\mathcal{Z}^J}_{\mathcal{Z}^J} N\big) \\
&\cong \mathrm{Hom}_{\mathcal{Z}^J}\big(M, \mathrm{Hom}_{{}^s\mathcal{Z}^J}(\mathcal{Z}^J, \mathrm{Res}^{{}^s\mathcal{Z}^J}_{\mathcal{Z}^J} N)\big) \\
&\cong \mathrm{Hom}_{\mathcal{Z}^J}\big(M, \mathcal{Z}^J\langle 2\rangle \otimes_{{}^s\mathcal{Z}^J} (\mathrm{Res}^{{}^s\mathcal{Z}^J}_{\mathcal{Z}^J} N)\big) \\
&= \mathrm{Hom}_{\mathcal{Z}^J}(M, {}^s\theta\langle 2\rangle N).
\end{aligned}$$

□

Parabolic special modules. As in [Fiebig 2008b], we define inductively a full subcategory of $\mathcal{Z}^J$-mod.

Let $B_e^J \in \mathcal{Z}^J$-mod be the free S-module of rank one on which $z = (z_x)_{x \in \mathcal{W}^J}$ acts via multiplication by z_e.

Definition 4.6. The category $\mathcal{H}^J$ of *special $\mathcal{Z}^J$-modules* is the full subcategory of $\mathcal{Z}^J$-mod$^{\mathrm{f}}$ whose objects are isomorphic to a direct summand of a direct sum of modules of the form ${}^{s_{i_1}}\theta \circ \cdots \circ {}^{s_{i_r}}\theta(B_e^J)\langle n\rangle$, where $s_{i_1}, \ldots, s_{i_r} \in \mathcal{S}$ and $n \in \mathbb{Z}$.

The category ${}^s\mathcal{H}^J$ of *special ${}^s\mathcal{Z}^J$-modules* is the full subcategory of ${}^s\mathcal{Z}^J$-mod$^{\mathrm{f}}$ whose objects are isomorphic to a direct summand of ${}^{s,\mathrm{on}}\theta(M)$ for some $M \in \mathcal{H}^J$.

Finiteness of special modules. Let Ω be a finite subset of $\mathcal{W}^J$ and denote by $\mathcal{Z}^J(\Omega)$ the sections of the structure sheaf over Ω, that is,

$$\mathcal{Z}^J(\Omega) = \left\{ (z_x) \in \prod_{x\in\Omega} S \;\middle|\; \begin{array}{c} z_x \equiv z_y \pmod{\alpha_t} \\ \text{if there is } w \in \mathcal{W}_J \text{ s.t. } y\,w\,x^{-1} = t \in \mathcal{T} \end{array} \right\}.$$

If $\Omega \subseteq \mathcal{W}^J$ is s-invariant, that is, $s\Omega = \Omega$, we may restrict ${}_s\sigma$ to it. We denote by ${}^s\mathcal{Z}^J(\Omega) \subseteq \mathcal{Z}^J(\Omega)$ the space of invariants and using Lemma 4.1, we get a decomposition $\mathcal{Z}^J(\Omega) = {}^s\mathcal{Z}^J(\Omega) \oplus \overline{\alpha}_s \cdot {}^s\mathcal{Z}^J(\Omega)$.

In the following lemma we prove the *finiteness* of special $\mathcal{Z}^J$-modules as Fiebig [2011] does for special $\mathcal{Z}$-modules.

Lemma 4.7. (1) *Let $M \in \mathcal{H}^J$. Then there exists a finite subset $\Omega \subset \mathcal{W}^J$ such that $\mathcal{Z}^J$ acts on M via the canonical map $\mathcal{Z}^J \to \mathcal{Z}^J(\Omega)$.*

(2) *Let $s \in \mathcal{S}$ and let N be an object in ${}^s\mathcal{H}^J$. Then there exists a finite s-invariant subset $\Omega \subset \mathcal{W}^J$ such that ${}^s\mathcal{Z}^J$ acts on N via the canonical map ${}^s\mathcal{Z}^J \to {}^s\mathcal{Z}^J(\Omega)$.*

Proof. We prove (1) by induction. It holds clearly for B_e, since $\mathcal{Z}^J$ acts on it via the map $\mathcal{Z}^J \to \mathcal{Z}^J(\{e\})$. Now we have to show that if the claim is true for $M \in \mathcal{H}^J$, then it holds also for ${}^s\theta(M)$. Suppose $\mathcal{Z}^J$ acts via the map $\mathcal{Z}^J \to \mathcal{Z}^J(\Omega)$ over M. Observe that we may assume Ω s-invariant since we can just replace it by $\Omega \cup s\Omega$, which is still finite. In this way the ${}^s\mathcal{Z}^J$-action on ${}^s\theta M$ factors via ${}^s\mathcal{Z}^J \to {}^s\mathcal{Z}^J(\Omega)$ and so we obtain ${}^s\theta M := \mathcal{Z}^J \otimes_{{}^s\mathcal{Z}^J} M = \mathcal{Z}^J(\Omega) \otimes_{{}^s\mathcal{Z}^J(\Omega)} M$.

Claim (2) follows directly from claim (1). $\square$

5. Modules with Verma flag and statement of the main result

We recall some notation from [Fiebig 2008a]. Let Q be the quotient field of S and let A be an S-module. Then we denote by $A_Q = A \otimes_S Q$. Let us assume $\mathcal{G}$ to be such that for any $M \in \mathcal{Z}\text{-mod}^{\mathrm{f}}$ there is a canonical decomposition

$$M_Q = \bigoplus_{x\in\mathcal{V}} M_Q^x \tag{7}$$

and so a canonical inclusion $M \subseteq \bigoplus_{x\in\mathcal{V}} M_Q^x$. For all subsets of the set of vertices $\Omega \subseteq \mathcal{V}$, we may define

$$M_\Omega := M \cap \bigoplus_{x\in\Omega} M_Q^x,$$
$$M^\Omega := M/M_{\mathcal{V}\setminus\Omega} = \mathrm{im}\Big(M \to M_Q \to \bigoplus_{x\in\Omega} M_Q^x\Big).$$

For all $x \in \mathcal{V}$, we set

$$M_{[x]} := \ker\big(M^{\{\trianglerighteq x\}} \to M^{\{\triangleright x\}}\big)$$

and if $x \lhd y$ and $[x, y] = \{x, y\}$, we denote

$$M_{[x,y]} := \ker\left(M^{\{\unrhd x\}} \to M^{\{\unrhd x\}\setminus\{x,y\}}\right).$$

Remark 5.1. In [Fiebig 2008a], the module $M_{[x]}$ is denoted by $M^{[x]}$. The notation we are adopting in this paper is the one from [Fiebig 2011].

Modules with a Verma flag. From now on, let $\mathcal{G}$ be a Bruhat moment graph. In [Fiebig 2008a] it is shown that in this case any $M \in \mathcal{Z}\text{-mod}^{\mathrm{f}}$ admits a decomposition like (7) and hence the modules $M_{[x]}$ are well defined for any $x \in \mathcal{V}$.

Let $\mathcal{V}$ denote the full subcategory of $\mathcal{Z}\text{-mod}^{\mathrm{f}}$ whose objects admit a Verma flag, that is, $M \in \mathcal{V}$ if and only if M^{Ω} is a graded free S-module for any $\Omega \subseteq \mathcal{V}$ upwardly closed with respect to the partial order in the set of vertices. In our hypotheses this condition is equivalent to $M_{[x]}$ being a graded free S-module for any $x \in \mathcal{V}$ (see [Fiebig 2008a, Lemma 4.7]).

Exact structure. Now we want to equip the category $\mathcal{V}$ with an exact structure.

Definition 5.2. A sequence $L \to M \to N$ in $\mathcal{V}$ is called *short exact* if

$$0 \to L_{[x]} \to M_{[x]} \to N_{[x]} \to 0$$

is a short exact sequence of S-modules for any $x \in \mathcal{V}$.

Remark 5.3. This is not the original definition of exact structure Fiebig [2008a] gave, which was on the whole category $\mathcal{Z}\text{-mod}^{\mathrm{f}}$, but it is known to be equivalent to it if we only consider the category $\mathcal{V}$, that is, precisely the one we are dealing with (see [Fiebig 2008b, Lemma 2.12]).

Decomposition and subquotients of modules on $\mathcal{Z}^{\mathbf{J}}$. To show that $\mathcal{H}^J$ categorifies the parabolic Hecke algebra, we will use a description of the action of ${}^s\theta$ on the subquotients $M_{[x]}$, for $x \in \mathcal{V}$ (Lemma 5.6). As a stepping-stone we prove an easy combinatorial consequence (Lemma 5.5) of the so-called lifting lemma:

Lemma 5.4 (lifting lemma [Humphreys 1990, Lemma 7.4]). *Let $s \in \mathcal{S}$ and $v, u \in \mathcal{W}$ be such that $vs < v$ and $u < v$.*

(1) *If $us < u$, then $us < vs$.*

(2) *If $us > u$, then $us \le v$ and $u \le vs$.*

Thus in both cases, $us \le v$.

Lemma 5.5. *Let $x \in \mathcal{W}^J$ and $t \in \mathcal{S}$. If $tx \notin \mathcal{W}^J$, then $(tx)^J = x$.*

Proof. If $tx \notin \mathcal{W}^J$, then there exists a simple reflection $r \in J$ such that $txr < tx$ and since $x \in \mathcal{W}^J$, $xr > x$. Using (the left version of) Lemma 5.4(1) with $s = t$, $v = xr$ and $u = tx$, we get $txr < x$. Applying Lemma 5.4(1) with $s = r$, $v = x$ and $u = txr$ it follows $tx > x$. Finally from Lemma 5.4(2) we obtain $txr \le x$, which together with $x < xr$, gives $txr = x$. $\square$

Lemma 5.6. *Let $s \in \mathcal{S}$ and $x \in \mathcal{W}^J$; then*

$$({}^s\theta M)_{[x]} \cong \begin{cases} M_{[x]}\langle -2\rangle \oplus M_{[sx]}\langle -2\rangle & \text{if } sx \in \mathcal{W}^J, sx > x, \\ M_{[x]} \oplus M_{[sx]} & \text{if } sx \in \mathcal{W}^J, sx < x, \\ M_{[x]}\langle -2\rangle \oplus M_{[x]} & \text{if } sx \notin \mathcal{W}^J. \end{cases}$$

Proof. By Lemma 5.5, if $sx \notin \mathcal{W}^J$, then $(sx)^J = x$ and $M_{[x]} \in {}^s\mathcal{Z}^J$-mod; so by Lemma 4.1, we get $\mathcal{Z}^J \otimes_{{}^s\mathcal{Z}^J} M_{[x]} = M_{[x]}\langle -2\rangle \oplus M_{[x]}$.

If $x \neq sx$, we have a short exact sequence of S-modules

$$0 \to M_{[x]} \to M_{[x,sx]} \to M_{[sx]} \to 0.$$

By Lemma 4.1, the finitely generated free S-module $\mathcal{Z}^J$ is flat over ${}^s\mathcal{Z}^J$, which is a finitely generated free $S^{\mathcal{W}_J}$-module. Hence ${}^s\theta M_{[x,sx]} = \mathcal{Z}^J \otimes_{{}^s\mathcal{Z}^J} M_{x,sx} = ({}^s\theta M)_{[x,sx]} = {}^s\theta M_{[x]} \oplus {}^s\theta M_{[sx]}$. Also ${}^s\theta M_{[x,sx]} = \mathcal{Z}^J(\{x, sx\}) \otimes_{{}^s\mathcal{Z}^J(\{x,sx\})} M_{[x,sx]}$. The two isomorphisms follow keeping in mind that $\mathcal{Z}^J(\{x, sx\})_{[x]} \cong S\langle -2\rangle$ if $x < sx$, while $\mathcal{Z}^J(\{x, sx\})_{[x]} \cong S$ if $x > sx$. □

Using induction, we obtain the following corollary:

Corollary 5.7. *Let $M \in \mathcal{H}^J$; then for any $x \in \mathcal{W}^J$, $M_{[x]}$ is a finitely generated torsion-free S-module and hence $M \in \mathcal{V}$.*

In this way we get an exact structure on $\mathcal{H}^J$ as well and we are finally able to state the main result of this paper:

Theorem 5.8. *The category $\mathcal{H}^J$ together with the shift in degree $\langle -1\rangle$ and (shifted) translation functors is a categorification of the parabolic Hecke module $\boldsymbol{M}^J$.*

Remark 5.9. Theorem 5.8 could be strengthen to a proper categorification by presenting the result as a 2-representation of a 2-category. The 2-category to be considered is the one generated by the translation functors we defined on p. 426, and the 2-representation to look at is given by the action of these functors on the category $\mathcal{H}^J$ we constructed on p. 428. The question of describing indecomposable 1-morphisms in this category, which we are not going to address in this paper, seems to be very interesting.

Remark 5.10. It follows from [Elias and Williamson 2014] that the results of [Mazorchuk and Stroppel 2005] transfer to all Coxeter systems.

6. Proof of the categorification theorem

The proof of Theorem 5.8 consists of several steps:

(1) We show that the functor ${}^s\theta \circ \langle 1\rangle$ is exact (Lemma 6.1).

(2) We define the character map $h^J : [\mathcal{H}^J] \to \boldsymbol{M}^J$ (p. 432).

(3) We observe that the map $[\langle -1\rangle] : [\mathcal{H}^J] \to [\mathcal{H}^J]$ provides $[\mathcal{H}^J]$ with a structure of $\mathcal{L}$-module and that h^J is a map of $\mathcal{L}$-modules (p. 432).

(4) Via explicit calculations, we prove that the functors ${}^s\theta \circ \langle 1\rangle$, $s \in \mathcal{S}$, satisfy (C3), that is, the maps they induce on $[\mathcal{H}^J]$ commute with h^J (Proposition 6.2).

(5) We demonstrate that the character map is surjective by choosing a certain basis for $\boldsymbol{M}^J$ and showing that every element of this basis has a preimage in $[\mathcal{H}^J]$ under h^J (Lemma 6.3).

(6) We prove that the character map is surjective (Lemma 6.6) using a description of indecomposable special modules in terms of Braden–MacPherson sheaves (Proposition 6.5).

This concludes the proof since (C2), that is, $\langle -1\rangle \circ ({}^s\theta \circ \langle 1\rangle) \cong ({}^s\theta \circ \langle 1\rangle) \circ \langle -1\rangle$ for any $s \in \mathcal{S}$, is trivially satisfied.

We start by proving the exactness of shifted translation functors.

Lemma 6.1. *For any $s \in \mathcal{S}$ the functor ${}^s\theta \circ \langle 1\rangle : \mathcal{H}^J \to \mathcal{H}^J$ is exact.*

Proof. Let $L \to M \to N$ be an exact sequence; then for any $x \in \mathcal{V}$

$$0 \to L_{[x]} \to M_{[x]} \to M_{[x]} \to 0$$

is a short exact sequence of S-modules. In particular,

$$0 \to L_{[sx]} \to M_{[sx]} \to N_{[sx]} \to 0$$

is short exact as well. The claim follows immediately from Lemma 5.6 and the fact that finite direct sums and shifts preserve exactness. □

Character maps. Let A be a $\mathbb{Z}$-graded, free and finitely generated S-module; then $A \cong \bigoplus_{i=1}^n S\langle k_i\rangle$ for some $k_i \in \mathbb{Z}$. We can associate to A its *graded rank*, that is, the following Laurent polynomial:

$$\underline{\mathrm{rk}}\, A := \sum_{i=1}^{n} v^{-k_i} \in \mathcal{L}.$$

This is well defined because the k_i are uniquely determined, up to reordering.

Let $M \in \mathcal{H}^J$. By Corollary 5.7, we may define a map $h^J : [\mathcal{H}^J] \to \boldsymbol{M}^J$ by

$$h^J([M]) := \sum_{x \in \mathcal{W}^J} v^{\ell(x)}\, \underline{\mathrm{rk}}\, M_{[x]}\, H_x^{J,v^{-1}} \in \boldsymbol{M}^J.$$

The Grothendieck group $[\mathcal{H}^J]$ is equipped with a structure of $\mathcal{L}$-module via $v^i[M] = [M\langle -i\rangle]$. Observe that for any $M \in \mathcal{H}^J$

$$h^J(v[M]) = h^J([M\langle -1\rangle]) = vh^J([M])$$

and so h^J is a map of $\mathcal{L}$-modules.

Proposition 6.2. *For each $M \in \mathcal{H}^J$, $s \in \mathcal{S}$ we have $h^J([{}^s\theta M\langle 1\rangle]) = \underline{H}_s \cdot h^J([M])$; that is, the following diagram is commutative*:

$$\begin{tikzcd} {[\mathcal{H}^J]} \arrow[r, "{[{}^s\theta\circ\langle 1\rangle]}"] \arrow[d, "h^J"'] & {[\mathcal{H}^J]} \arrow[d, "h^J"] \\ \boldsymbol{M}^J \arrow[r, "\underline{H}_s\cdot"'] & \boldsymbol{M}^J \end{tikzcd}$$

Proof. By Lemma 5.6, for any $x \in \mathcal{W}^J$ we have

$$\underline{\mathrm{rk}}({}^s\theta M)_{[x]} = \begin{cases} v^2(\underline{\mathrm{rk}}\, M_{[x]} + \underline{\mathrm{rk}}\, M_{[sx]}) & \text{if } sx \in \mathcal{W}^J,\, sx > x, \\ \underline{\mathrm{rk}}\, M_{[x]} + \underline{\mathrm{rk}}\, M_{[sx]} & \text{if } sx \in \mathcal{W}^J,\, sx < x, \\ (v^2+1)\,\underline{\mathrm{rk}}\, M_{[x]} & \text{if } sx \notin \mathcal{W}^J. \end{cases}$$

Then

$$\begin{aligned} h^J([{}^s\theta M\langle 1\rangle]) &= \sum_{x\in\mathcal{W}^J} v^{\ell(x)-1}\,\underline{\mathrm{rk}}({}^s\theta M)_{[x]} H_x^{J,v^{-1}} \\ &= \sum_{\substack{x\in\mathcal{W}^J,\, sx\in\mathcal{W}^J\\ sx>x}} v^{\ell(x)+1}(\underline{\mathrm{rk}}\, M_{[x]} + \underline{\mathrm{rk}}\, M_{[sx]}) H_x^{J,v^{-1}} \\ &\quad + \sum_{\substack{x\in\mathcal{W}^J,\, sx\in\mathcal{W}^J\\ sx<x}} v^{\ell(x)-1}(\underline{\mathrm{rk}}\, M_{[x]} + \underline{\mathrm{rk}}\, M_{[sx]}) H_x^{J,v^{-1}} \\ &\quad + \sum_{x\in\mathcal{W}^J,\, sx\notin\mathcal{W}^J} (v^{\ell(x)+1} + v^{\ell(x)-1})\,\underline{\mathrm{rk}}\, M_{[x]} H_x^{J,v^{-1}}. \end{aligned}$$

Finally

$$\begin{aligned} \underline{H}_s \cdot h^J([M]) &= \sum_{x\in\mathcal{W}^J} v^{\ell(x)}(\underline{\mathrm{rk}}\, M_{[x]})\underline{H}_s \cdot H_x^{J,v^{-1}} \\ &= \sum_{\substack{x\in\mathcal{W}^J,sx\in\mathcal{W}^J\\ sx>x}} v^{\ell(x)}(\underline{\mathrm{rk}}\, M_{[x]})(H_{sx}^{J,v^{-1}} + vH_x^{J,v^{-1}}) \\ &\quad + \sum_{\substack{x\in\mathcal{W}^J,\, sx\in\mathcal{W}^J\\ sx<x}} v^{\ell(x)}(\underline{\mathrm{rk}}\, M_{[x]})(H_{sx}^{J,v^{-1}} + v^{-1}H_x^{J,v^{-1}}) \\ &\quad + \sum_{x\in\mathcal{W}^J,\, sx\notin\mathcal{W}^J} v^{\ell(x)}\,\underline{\mathrm{rk}}\, M_{[x]}(v+v^{-1})H_x^{J,v^{-1}} \\ &= \sum_{\substack{x\in\mathcal{W}^J,\, sx\in\mathcal{W}^J\\ sx>x}} \big[(v^{\ell(x)}v\,\underline{\mathrm{rk}}\, M_{[x]}) + (v^{\ell(sx)}\,\underline{\mathrm{rk}}\, M_{[sx]})\big] H_x^{J,v^{-1}} \\ &\quad + \sum_{\substack{x\in\mathcal{W}^J,\, sx\in\mathcal{W}^J\\ sx<x}} \big[(v^{\ell(x)}v^{-1}\,\underline{\mathrm{rk}}\, M_{[x]}) + (v^{\ell(sx)}\,\underline{\mathrm{rk}}\, M_{[sx]})\big] H_x^{J,v^{-1}} \\ &\quad + \sum_{x\in\mathcal{W}^J,\, sx\notin\mathcal{W}^J} (v^{\ell(x)+1} + v^{\ell(x)-1})\,\underline{\mathrm{rk}}\, M_{[x]} H_x^{J,v^{-1}} \end{aligned}$$

$$= h^J([{}^s\theta M\langle 1\rangle]). \qquad \square$$

The character map is an isomorphism. In order to prove that $(\mathcal{H}^J, \langle -1\rangle, \{{}^s\theta \circ \langle 1\rangle\})$ is a categorification of $\boldsymbol{M}^J$, the only step left is to show that h^J is an isomorphism.

Lemma 6.3. *The map $h^J : [\mathcal{H}^J] \to \boldsymbol{M}^J$ is surjective.*

Proof. We start by defining a basis of $\boldsymbol{M}^J$. Let us set $\underline{\widetilde{H}}_e^{J,v^{-1}} = \underline{H}_e^{J,v^{-1}}$. For any $x \in \mathcal{W}^J$ with $\ell(x) = r > 0$, let us fix a reduced $x = s_{i_1} \dots s_{i_r}$, with $s_{i_1}, \dots, s_{i_r} \in \mathcal{S}$, and denote

$$\underline{\widetilde{H}}_x^{J,v^{-1}} = \underline{H}_{s_1}^{J,v^{-1}} \cdots \underline{H}_{s_r}^{J,v^{-1}}.$$

From Theorem 2.2, it follows that

$$\underline{\widetilde{H}}_x^{J,v^{-1}} = H_x^{J,v^{-1}} + \sum_{\substack{y \in \mathcal{W}^J \\ y<x}} p_y H_y^{J,v^{-1}}, \quad \text{with } p_z \in \mathbb{Z}[v, v^{-1}]. \tag{8}$$

Since $\{H_x^{J,v^{-1}}\}_{x\in\mathcal{W}^J}$ is a basis of $\boldsymbol{M}^J$ as a $\mathbb{Z}[v, v^{-1}]$-module, $\{\underline{\widetilde{H}}_x^{J,v^{-1}}\}_{x\in\mathcal{W}^J}$ is also a basis for $\boldsymbol{M}^J$ and it is enough to show that, for any $x \in \mathcal{W}^J$, there exists an object $H \in \mathcal{H}^J$ such that $h^J([H]) = \underline{\widetilde{H}}_x^{J,v^{-1}}$.

By definition, $h^J(B_e^J) = M_e = \underline{H}_e^{J,v^{-1}}$. By applying Proposition 6.2, we obtain

$$\begin{aligned} h^J\left({}^{s_{i_1}}\theta \circ \cdots \circ {}^{s_{i_r}}\theta B_e^J\langle n\rangle\right) &= (\underline{H}_{s_1}^{J,v^{-1}} \cdots \underline{H}_{s_r}^{J,v^{-1}}) M_e \\ &= \underline{H}_{s_1}^{J,v^{-1}} \cdots \underline{H}_{s_r}^{J,v^{-1}} = \underline{\widetilde{H}}_x^{J,v^{-1}}. \end{aligned}$$

This concludes the proof of the lemma. $\square$

Proposition 6.5 will allow us to see any element in $\mathcal{H}^J$ as the space of global sections of some BMP sheaf on $\mathcal{G}^J$. From now on, we will denote by $B^J(w)$ the space of global sections of the indecomposable BMP sheaf $\mathcal{B}^J(w) \in \mathbf{Sh}(\mathcal{G}^J)$. Let us recall a fundamental characterisation of $B^J(w)$.

Theorem 6.4 [Fiebig 2008b, Theorem 5.2]. *For any $w \in \mathcal{G}^J$, the module $B^J(w) \in \mathcal{V}$ is indecomposable and projective. Moreover every indecomposable projective object in $\mathcal{V}$ is isomorphic to $B^J(w)\langle k\rangle$ for a unique $w \in \mathcal{G}^J$ and a unique $k \in \mathbb{Z}$.*

Proposition 6.5. *A module $M \in \mathcal{Z}^J$-mod^f is an indecomposable special module if and only if there exist a BMP sheaf $\mathcal{B} \in \boldsymbol{Sh}(\mathcal{G}^J)$ and $k \in \mathbb{Z}$ such that $M \cong \Gamma(\mathcal{B}\langle k\rangle)$ as $\mathcal{Z}^J$-modules.*

Proof. By induction, from the exactness of ${}^s\theta^J$, it follows that the objects of $\mathcal{H}^J$ are all projective, and then by Theorem 6.4, any $M \in \mathcal{H}^J$ may be identified (up to a shift) with the space of global sections of a BMP sheaf on $\mathcal{G}^J$.

We now want to show that for any $x \in \mathcal{W}_J$, $B^J(x) \in \mathcal{H}^J$. We prove the claim by induction on $\#\operatorname{supp}(M)$, where $\operatorname{supp}(M) = \{x \in \mathcal{W}^J \mid M^x \neq 0\}$. Clearly $B_e \cong B^J(e)$.

The statement follows straightforwardly, once we prove that if $sx > x$, then ${}^s\theta^J(B^J(x)) = B^J(sx) \oplus B$.

First we show that $\mathrm{supp}({}^s\theta^J(B^J(x))) \subseteq \{\leq sx\}$, that is, $({}^s\theta^J B^J(x))^y = 0$ for all $y \notin \{\leq sx\} \cap \mathcal{W}^J$. From Lemma 5.6, it follows easily that $({}^s\theta^J(B^J(x)))_{[y]} = 0$ for all $y \notin \{\leq sx\} \cap \mathcal{W}^J$.

Let us observe that as ${}^s\theta^J B^J(x) \in \mathcal{H}^J$, from what we have proved above, there exist $w_1, \dots, w_r \in \mathcal{W}^J$ and $k_1, \dots, k_r$ such that ${}^s\theta^J(B^J(x)) = \bigoplus_{i=1}^r B^J(w_i)\langle k_i\rangle$, and for any $y \in \mathcal{W}^J$,

$$\left(\bigoplus_{i=1}^r B^J(w_i)\langle k_i\rangle\right)_{[y]} = \bigoplus_{i=1}^r B^J(w_i)_{[y]}\langle k_i\rangle.$$

So, in particular, for all $y \notin \{\leq sx\} \cap \mathcal{W}^J$,

$$\begin{aligned} 0 &= B^J(w_i)_{[y]} \\ &= \ker\bigl(\rho_{\delta y} : \mathcal{B}^J(w_i)^y \to \mathcal{B}^J(w_i)^{\delta y}\bigr). \end{aligned}$$

This implies $\mathcal{B}^J(w_i)^y = B^J(w_i)^y = 0$ for all $i = 1, \dots r$, and so

$${}^s\theta^J(B^J(x)) = \bigoplus_{i=1}^r B^J(w_i)\langle k_i\rangle,$$

where $w_i \in \{\leq sx\}$ for all $i = 1, \dots, r$.

It is left to show that there exists at least one $i \in \{1, \dots, r\}$ such that $w_i = sx$. By applying once again Lemma 5.6, we obtain $({}^s\theta^J(B^J(x)))^{sx} = ({}^s\theta^J(B^J(x)))_{[sx]} \cong S$ and hence the statement. □

Lemma 6.6. *The map $h^J : [\mathcal{H}^J] \to \boldsymbol{M}^J$ is injective.*

Proof. By Theorem 6.4 and Proposition 6.5 we know that $\{[B^J(w)]\}_{w\in\mathcal{W}^J}$ is a $\mathbb{Z}[v, v^{-1}]$-basis of $[\mathcal{H}^J]$ and so every element $Y \in [\mathcal{H}^J]$ can be written as $Y = \sum a_w[B^J(w)]$, with $a_x \in \mathbb{Z}[v, v^{-1}]$. Let us suppose $Y \in \ker(h^J)$. Then

$$\begin{aligned} 0 = h^J(Y) &= \sum_{w\in\mathcal{W}^J} a_w \sum_{x\in\mathcal{W}^J} v^{\ell(x)}\, \underline{\mathrm{rk}}\, B^J(w)_{[x]}\; H_x^{J,v^{-1}} \\ &= \sum_{x\in\mathcal{W}^J}\Bigl(\sum_{w\in\mathcal{W}^J} v^{\ell(x)} a_w\, \underline{\mathrm{rk}}\, B^J(w)_{[x]}\Bigr)\; H_x^{J,v^{-1}}. \end{aligned}$$

Since the elements $H_x^{J,v^{-1}}$ are linearly independent, it follows that

$$\sum_{w\in\mathcal{W}^J} v^{\ell(x)} a_w\, \underline{\mathrm{rk}}\, B^J(w)_{[x]} = 0 \quad \text{for any } x \in \mathcal{W}^J.$$

If it were the case that $Y \neq 0$, then we would find a maximal element $\overline{w}$ such that

$a_{\overline{w}} \neq 0$. By (BMP1), we obtain $B^J(w)_{[\overline{w}]} = 0$ for all $w < \overline{w}$ and $B^J(\overline{w})_{[\overline{w}]} \cong S$. Then

$$0 = \sum_{w \in \mathcal{W}^J} v^{\ell(x)} a_w \ \underline{\mathrm{rk}}\, B^J(w)_{[\overline{w}]} = v^{\ell(x)} a_{\overline{w}}\, \underline{\mathrm{rk}}\, B^J(\overline{w})_{[\overline{w}]} = v^{\ell(x)} a_{\overline{w}}\, \underline{\mathrm{rk}}\, S = v^{\ell(x)} a_{\overline{w}}.$$

The chain of equalities above gives us a contradiction since we assumed $a_{\overline{w}} \neq 0$. $\square$

This concludes the proof of Theorem 5.8.

7. The functor I

In this section we define an exact functor $I : \mathcal{H}^J \to \mathcal{H}^{\varnothing}$ such that the following diagram commutes:

$$\begin{tikzcd} {[\mathcal{H}^J]} \arrow[r, hook, "{[I]}"] \arrow[d, "h^J"'] & {[\mathcal{H}^{\varnothing}]} \arrow[d, "h^{\varnothing}"] \\ \boldsymbol{M}^J \arrow[r, hook, "i"'] & \boldsymbol{H} \end{tikzcd}$$

where $i : \boldsymbol{M}^J \hookrightarrow \boldsymbol{H}$ is the map of $\mathcal{L}$-modules given by

$$H_x^{J,v^{-1}} \mapsto \sum_{z \in \mathcal{W}_J} v^{\ell(w_J) - \ell(z)} H_{xz}, \tag{9}$$

with w_J the longest element of $\mathcal{W}_J$.

The map i is interesting since it gives us a way to see the parabolic Hecke module $\boldsymbol{M}^J$ as submodule of $\boldsymbol{H}$, and hence its categorification tells us that we can think about the $\mathcal{H}^J$ as a subcategory of $\mathcal{H}^{\varnothing}$.

We construct the functor I by using a localisation-globalisation procedure. More precisely, we first map the elements of $\mathcal{H}^J$ to certain sheaves on $\mathcal{G}^J$, then apply a pullback functor mapping them to sheaves on $\mathcal{G}$, and finally we take global sections of the latter. A priori it is not clear that we obtain an object in $\mathcal{H}^{\varnothing}$. This fact is shown in Lemma 7.3. We then demonstrate the exactness of I (Proposition 7.5) and the commutativity of diagram 7 (Proposition 7.7) by a study of the subquotients involved in the definition of the character map. The realisation of special modules in terms of Braden–MacPherson sheaves given in the previous section (Proposition 6.5) plays a crucial role in the proof of any of the above results.

Construction of the functor I. The definition of I involves Fiebig's localisation functor $\mathcal{L}$ [2008a, §3.3], which allows us to see objects of $\mathcal{Z}^J$-mod as sheaves on the parabolic Bruhat moment graph $\mathcal{G}^J$.

Let us assume $\mathcal{G}$ to be such that for any $M \in \mathcal{Z}$-mod$^{\mathrm{f}}$ there is a canonical decomposition like the one in (7). Let $\mathcal{Z}$ be the corresponding structure algebra and $M \in \mathcal{Z}\text{-mod}^{\mathrm{f}}$. For any vertex $x \in \mathcal{V}$, we set

$$\mathcal{L}(M)^x = M^x. \tag{10}$$

For any edge $E : x \text{ --- } y$, let us consider $\mathcal{Z}(E) = \{(z_x, z_x) \in S \oplus S \mid z_x - z_y \in l(E)S\}$ and $M(E) := \mathcal{Z}(E) \cdot M^{x,y}$. For $m = (m_x, m_y) \in M(E)$, let us set $\pi_x((m)) = m_x$, $\pi_y((m)) = m_y$. Then we get $\mathcal{L}(M)^E$ as the pushout in the following diagram of S-modules:

$$\begin{tikzcd} M(E) \arrow[r, "\pi_x"] \arrow[d, "\pi_y"'] & M^x \arrow[d, "\rho_{x,E}"'] \\ M^y \arrow[r, "\rho_{y,E}"] & \mathcal{L}(M)^E \end{tikzcd}$$

This provides us also with the restriction maps $\rho_{x,E}$ and $\rho_{y,E}$.

It is not hard to verify (see [Fiebig 2008a, §3.3]) that this is a well defined functor

$$\mathcal{L} : \mathcal{Z}\text{-mod}^{\mathrm{f}} \to \mathbf{Sh}(\mathcal{G}). \tag{11}$$

Moreover the localisation functor $\mathcal{L}$ turns out to be left-adjoint to Γ (see [Fiebig 2008a, Theorem 3.5]). Let

- $\mathcal{Z}\text{-mod}^{\mathrm{loc}}$ be the full subcategory of $\mathcal{Z}\text{-mod}^{\mathrm{f}}$ whose objects are the elements M such that there is an isomorphism $\Gamma \circ \mathcal{L}(M) \cong M$, and
- $\mathbf{Sh}(\mathcal{G})^{\mathrm{glob}}$ be the full subcategory of $\mathbf{Sh}(\mathcal{G})$ whose objects are the elements $\mathcal{F}$ such that there is an isomorphism $\mathcal{L} \circ \Gamma(\mathcal{F}) \cong \mathcal{F}$.

Remark 7.1. In general, for a given a sheaf $\mathcal{F}$, one has $(\mathcal{L} \circ \Gamma(\mathcal{F}))^x = \Gamma(\mathcal{F})^x \subseteq \mathcal{F}^x$. If we consider a BMP sheaf $\mathcal{B}$, then by property (BMP3), $\Gamma(\mathcal{B})^x = \mathcal{B}^x$ for any vertex $x \in \mathcal{V}$ and $\mathcal{L}(\Gamma(\mathcal{B}))^E \cong \mathcal{B}^E$ for any edge $E \in \mathcal{E}$. Therefore $\mathcal{L} \circ \Gamma(\mathcal{B}) \cong \mathcal{B}$ and $\mathcal{B} \in \mathbf{Sh}(\mathcal{G})^{\mathrm{glob}}$.

Thus the functors $\mathcal{L}$ and Γ induce two inverse equivalences:

$$\mathcal{Z}\text{-mod}^{\mathrm{loc}} \longleftrightarrow \mathbf{Sh}(\mathcal{G})^{\mathrm{glob}}\ .$$

Let us focus again on the Bruhat case and consider the functor

$$p^{J,*} : \mathbf{Sh}(\mathcal{G}^J) \to \mathbf{Sh}(\mathcal{G})$$

defined as follows:

- $(p^{J,*}\mathcal{F})^x := \mathcal{F}^{x^J}$ for all $x \in \mathcal{W}$.
- for all $E : x \text{ --- } y \in \mathcal{E}$,

$$(p^{J,*}\mathcal{F})^E = \begin{cases} \mathcal{F}^{f_{\mathcal{V}}(x)} / l(E)\mathcal{F}^{f_{\mathcal{V}}(x)} & \text{if } x^J = y^J, \\ \mathcal{F}^{f_{\mathcal{E}}(E)} & \text{otherwise}\ . \end{cases}$$

- for all $x \in \mathcal{W}$ and $E \in \mathcal{E}$ such that $E : x \text{---} y$,

$$(p^{J,*}\rho)_{x,E} = \begin{cases} \text{canonical quotient map} & \text{if } x^J = y^J, \\ \rho_{f_{\mathcal{V}}(x), f_{\mathcal{E}}(E)} & \text{otherwise .} \end{cases}$$

Finally we set $I := \langle -\ell(w_J) \rangle \circ \Gamma \circ p^{J,*} \circ \mathcal{L}$.

To prove that the functor I maps $\mathcal{H}^J$ to $\mathcal{H}$, we need to recall the moment graph analogue of a theorem by Deodhar relating parabolic Kazhdan–Lusztig polynomials and regular ones. The following is a reformulation of Theorem 6.1 of [Lanini 2012]:

Theorem 7.2. *Let $J \subseteq \mathcal{S}$ be such that W_J is finite, with longest element w_J. Let $w \in \mathcal{W}^J$; then $p^{J,*}(\mathcal{B}^J(w)) \cong \mathcal{B}^{\varnothing}(ww_J)$ as sheaves on $\mathcal{G} = \mathcal{G}(\mathcal{W}, \varnothing)$.*

Lemma 7.3. *The functor I maps $\mathcal{H}^J$ to $\mathcal{H}$.*

Proof. Let $M \in \mathcal{H}^J$; then, by Proposition 6.5, there exist $w_1, \dots w_r \in \mathcal{W}^J$ and $m_1, \dots m_r \in \mathbb{Z}$ such that $M = \bigoplus_{i=1}^r B^J(w_i)\langle m_i \rangle$. Then

$$I(M) = I\left(\bigoplus_{i=1}^r B^J(w_i)\langle m_i \rangle\right) = \bigoplus_{i=1}^r \Gamma \circ p^{J,*} \circ \mathcal{L}(B^J(w_i))\langle m_i - \ell(w_J) \rangle.$$

By Remark 7.1, $\mathcal{L}(B^J(w_i)) \cong \mathcal{B}^J(w_i)$ for any i and, by Theorem 7.2, we conclude that

$$I(M) \cong \bigoplus_{i=1}^r B^{\varnothing}(w_i w_J)\langle m_i - \ell(w_J) \rangle. \qquad \square$$

Exactness of I.

Lemma 7.4. *Let $w \in \mathcal{W}^J$. Then for all $x \in \mathcal{W}$,*

$$(\Gamma \circ p^{J,*}\mathcal{B}^J(w))_{[x]} = \left(\prod_{\substack{y \in \mathcal{V}^{\delta x}, \\ y \in x\mathcal{W}_J}} \alpha_y\right) B^J(w)_{[x^J]},$$

where α_y denotes the label of $x \to y$.

Proof. For $z \in \mathcal{W}^J$ and E an edge of $\mathcal{G}^J = \mathcal{G}(\mathcal{W}, J)$, let us denote by $\rho_{z,E}$ the corresponding restriction map. Then we have

$$\begin{aligned}(\Gamma \circ p^{J,*}\mathcal{B}^J(w))_{[x]} &= \bigcap_{y \in \mathcal{V}^{\delta x}} \ker\left((p^{*,J}\rho)_{x,x \to y}\right) \\ &= \left(\bigcap_{\substack{y \in \mathcal{V}^{\delta x} \\ y \notin x\mathcal{W}_J}} \ker(\rho_{x^J, x^J \to y^J})\right) \cap \left(\bigcap_{\substack{y \in \mathcal{V}^{\delta x}, \\ y \in x\mathcal{W}_J}} \ker \pi_{x,x \to y}\right),\end{aligned}$$

where $\pi_{x,x\to y} : \mathcal{B}^J(w)^{x^J} \to \mathcal{B}^J(w)^{x^J}/\alpha_y \mathcal{B}^J(w)^{x^J}$ is the canonical quotient map and α_y is a generator of $l(x \to y)$.

Let us observe that by definition,

$$\bigcap_{\substack{y\in\mathcal{V}^{\delta x}\\ y\notin x\mathcal{W}_J}} \ker(\rho_{x^J,x^J\to y^J}) = B^J(w)_{[x^J]}.$$

Moreover since there is at most one edge adjacent to x labelled by a multiple of α_y, the labels of such edges are pairwise linearly independent and we get

$$\bigcap_{\substack{y\in\mathcal{V}^{\delta x},\\ y\in x\mathcal{W}_J}} \ker\pi_{x,x\to y} = \prod_{\substack{y\in\mathcal{V}^{\delta x},\\ y\in x\mathcal{W}_J}} \alpha_y\cdot\mathcal{B}^J(w)^{x^J}.$$

It follows that

$$(\Gamma\circ p^{J,*}\mathcal{B}^J(w))_{[x]} = \Big(\prod_{\substack{y\in\mathcal{V}^{\delta x},\\ y\in x\mathcal{W}_J}} \alpha_y\Big)\mathcal{B}^J(w)_{[x^J]}.$$

This concludes the proof of the lemma. □

Proposition 7.5. *The functor I is exact with respect to the exact structure in Section 5.*

Proof. Let us take $M, N\in\mathcal{H}^J$, with $M=\bigoplus_k B^J(w_k)\langle m_k\rangle$ and $N=\bigoplus_l B^J(w_l)\langle n_l\rangle$.

Let us consider the map $f: L\to M$ and the induced maps $f_{[x^J]}: M_{[x^J]}\to N_{[x^J]}$ for any $x^J\in\mathcal{W}^J$. Thanks to Lemma 7.4, it is easy to describe $I(f)_{[x]}$. Namely if

$$\prod_{\substack{y\in\mathcal{V}^{\delta x}\\ y\in x\mathcal{W}_J}} \alpha_y = \alpha_{i_1}\cdots\alpha_{i_r},$$

we obtain

$$\begin{array}{cccc} I(f)\ : & I(M)_{[x]} & \longrightarrow & I(N)_{[x]},\\ & (\alpha_{i_1}\cdots\alpha_{i_r})m & \longmapsto & (\alpha_{i_1}\cdots\alpha_{i_r})f_{[x]}(m). \end{array}$$

It is clear that if $0\to L_{[x]}\to M_{[x]}\to N_{[x]}\to 0$ is a short exact sequence of S-modules, then $0\to(IL)_{[x]}\to(IM)_{[x]}\to(IN)_{[x]}\to 0$ is also exact. □

Commutativity of the diagram. The last step missing is the commutativity of Diagram 7. Before proving it, we need the following preliminary lemma:

Lemma 7.6. *Let $w\in\mathcal{W}^J$ and let w_J be the longest element of $\mathcal{W}_J$. There is an isomorphism $B^\varnothing(ww_J)_{[x]}\cong B^J(w)_{[x^J]}\langle 2\ell(x)-2\ell(x^J)-2\ell(w_J)\rangle$ of graded S-modules.*

Proof. By Theorem 7.2, $\mathcal{B}^\varnothing(ww_J)\cong p^{J,*}\mathcal{B}^J(w_J)$ as sheaves on $\mathcal{G}=\mathcal{G}(\mathcal{W},\varnothing)$. It follows that for any $x\in\mathcal{W}$, $B^\varnothing(ww_J)_{[x]}\cong(\Gamma\circ p^{J,*}\mathcal{B}^J(w_J))_{[x]}$ as graded

S-modules and then, by Lemma 7.4, we obtain

$$\begin{aligned} B^{\varnothing}(ww_J)_{[x]} &\cong \Bigg(\prod_{\substack{y\in\mathcal{V}^{\delta x},\\ y\in x\mathcal{W}_J}} \alpha_y\Bigg) B^J(w)_{[x^J]} \\ &\cong B^J(w)_{[x^J]}\big\langle 2\cdot\#\{y\in\mathcal{V}^{\delta x},\ y\in x\mathcal{W}_J\}\big\rangle. \end{aligned}$$

Let $x'=(x^J)^{-1}x\in W_J$. Now if $\mathcal{T}_J$ is the set of reflections of $\mathcal{W}_J$,

$$\begin{aligned} \#\{y\in\mathcal{V}^{\delta x},\ y\in x\mathcal{W}_J\} &= \#\{z\in\mathcal{W}_J \mid \text{there exists } t\in\mathcal{T}_J \text{ s.t. } z=x't \text{ and } x'<z\} \\ &= \ell(w_J)-\ell(x') = \ell(w_J)-\ell(x)+\ell(x^J). \end{aligned}$$ □

Finally we are able to prove the following proposition, which enables us to embed $\mathcal{H}^J$ in $\mathcal{H}$.

Proposition 7.7. *The following diagram is commutative*:

$$\begin{tikzcd} {[\mathcal{H}^J]} \arrow[r, hook, "{[I]}"] \arrow[d, "h^J"'] & {[\mathcal{H}^{\varnothing}]} \arrow[d, "h^{\varnothing}"] \\ \boldsymbol{M}^J \arrow[r, hook, "i"'] & \boldsymbol{H} \end{tikzcd}$$

Proof. As $I\big(\bigoplus_{i\in I} B^J(w_i)\big)=\bigoplus I\big(B^J(w_i)\big)$, it is enough to prove the statement for the module $B^J(w)$. In this case, we have:

$$\begin{aligned} I(B^J(w)) &= \langle-\ell(w_J)\rangle\circ\Gamma\circ p^{J,*}\circ\mathcal{L}(B^J(w)) \\ &= \langle-\ell(w_J)\rangle\circ\Gamma\circ p^{J,*}(\mathcal{B}^J(w)) \\ &\cong \langle-\ell(w_J)\rangle\circ\Gamma(\mathcal{B}^{\varnothing}(ww_J)) \\ &= B(ww_J)\langle-\ell(w_J)\rangle. \end{aligned}$$

Thus if $B^J(w)_{[x^J]}=\bigoplus_{i\in I_{x^J}} S\langle k_i\rangle$, we get

$$\begin{aligned} h^{\varnothing}\circ[I]([B^J(w)]) &= h^{\varnothing}(B^{\varnothing}(ww_J)\langle\ell(w_J)\rangle) \\ &= \sum_{x\in\mathcal{W}} v^{-\ell(w_J)+\ell(x)}\, \underline{\mathrm{rk}}\, B^{\varnothing}(ww_J)_{[x]} H_x \\ \text{(by Lemma 7.6)}\quad &= \sum_{x\in\mathcal{W}} v^{\ell(w_J)+\ell(x)}\, \underline{\mathrm{rk}}\big(B^J(w)_{[x^J]}\langle 2\ell(x_J)-2\ell(w_J)\rangle\big) H_x \\ &= \sum_{x\in\mathcal{W}} v^{-\ell(w_J)+\ell(x)}\Bigg(\sum_{i\in I_{x^J}} v^{-2\ell(x_J)+2\ell(w_J)-k_i}\Bigg) H_x \\ &= \sum_{x\in\mathcal{W}} v^{\ell(w_J)+\ell(x)}\Bigg(\sum_{i\in I_{x^J}} v^{-2\ell(x_J)-k_i}\Bigg) H_x, \end{aligned}$$

where $H_x = H_x^{\varnothing, v^{-1}}$. On the other hand, we have

$$\begin{aligned} i \circ h^J([B^J(w)]) &= i\Bigg(\sum_{x^J \in \mathcal{W}^J} v^{\ell(x^J)} \,\underline{\mathrm{rk}}\, B^J(w)_{[x^J]} H_{x^J}^{J,v^{-1}}\Bigg) \\ &= \sum_{x^J \in \mathcal{W}^J} \Bigg[v^{\ell(x^J)} \Bigg(\sum_{i \in I_{x^J}} v^{-k_i} \Bigg) i(H_{x^J}^{J,v^{-1}}) \Bigg] \\ &= \sum_{x^J \in \mathcal{W}^J} \Bigg[v^{\ell(x^J)} \Bigg(\sum_{i \in I_{x^J}} v^{-k_i} \Bigg) \Bigg(\sum_{x_J \in \mathcal{W}_J} v^{\ell(w_J) - l(x_J)} H_{x^J x_J} \Bigg) \Bigg] \\ &= \sum_{x^J \in \mathcal{W}^J} \sum_{x_J \in \mathcal{W}_J} \Bigg(\sum_{i \in I_{x^J}} v^{\ell(x^J) - k_i + \ell(w_J) - \ell(x_J)} \Bigg) H_{x^J x_J} \\ &= \sum_{x \in \mathcal{W}} v^{\ell(w_J) + \ell(x)} \Bigg(\sum_{i \in I_{x^J}} v^{-2\ell(x_J) - k_i} \Bigg) H_x. \end{aligned}$$ □

8. Connection with the equivariant category $\mathcal{O}$

In this section, we briefly discuss the connection of our results with noncritical blocks in an equivariant version of category $\mathcal{O}$.

Let $\mathfrak{g}$ be a complex symmetrisable Kac–Moody algebra and $\mathfrak{b} \supseteq \mathfrak{h}$ its Borel and Cartan subalgebras. The Weyl group $\mathcal{W}$ of $\mathfrak{g}$ naturally acts on $\mathfrak{h}^\star$, and we can consider equivalence classes $\Lambda \in \mathfrak{h}^\star/\sim$. An element $\lambda \in \mathfrak{h}^\star$ is noncritical if $2(\lambda + \rho, \beta) \notin \mathbb{Z}(\beta, \beta)$ for any imaginary root β, and an orbit Λ is noncritical if any $\lambda \in \Lambda$ is noncritical.

Fix a noncritical orbit Λ and a weight $\lambda_0 \in \Lambda$. As in Definition 3.3, we can look at the $\mathcal{W}$-orbit of λ_0, which gives us a Bruhat moment graph on $\mathfrak{h}^\star$. We want to discuss the representation theoretic content of $\mathcal{H}^J$, where J is in this case given by the set of simple reflections generating $\mathrm{Stab}_{\mathcal{W}}\, \lambda_0$. Denote by $\mathcal{G}(\Lambda)$ such a graph.

Let $S = S(\mathfrak{h})$ be the symmetric algebra of $\mathfrak{h}$, $R = S_{(\mathfrak{h})}$ be its localisation at $0 \in \mathfrak{h}^\star$ and $\tau : S \to R$ be the canonical map. For any $\mu \in \mathfrak{h}^\star$ and any $(\mathfrak{g}$-$R)$-bimodule M, we define its *μ-weight space* as

$$M_\mu = \{ m \in M \mid H \,.\, m = (\lambda(H) + \tau(H))m \;\text{ for any } H \in \mathfrak{h} \}.$$

If $\mathfrak{g}$-mod-R denotes the category of $(\mathfrak{g}$-$R)$-bimodules, then the equivariant version of category $\mathcal{O}$ we want to study is

$$\mathcal{O}_R = \left\{ M \in \mathfrak{g}\text{-mod-}R \;\middle|\; \begin{array}{l} M \text{ is locally finite as a } (\mathfrak{b}\text{-}R)\text{-bimodule}, \\ M = \bigoplus_{\mu \in \mathfrak{h}^\star} M_\mu \end{array} \right\}.$$

For any $\mu \in \mathfrak{h}^\star$ let us consider the $(\mathfrak{h}$-$R)$-bimodule R_μ free of rank one over R on which $\mathfrak{h}$ acts via the character $\mu + \tau$. The projection $\mathfrak{b} \to \mathfrak{h}$ allows us to consider R_μ as a $(\mathfrak{b}$-$R)$-bimodule and we can now induce to obtain the *equivariant Verma module* of weight μ: $M_R(\mu) = U(\mathfrak{g}) \otimes_{U(\mathfrak{b})} R_\mu$, where $U(\mathfrak{g})$ and $U(\mathfrak{b})$ are the enveloping algebras of $\mathfrak{g}$ and $\mathfrak{b}$, respectively.

Let $\mathcal{M}_R$ be the full subcategory of $\mathcal{O}_R$ whose objects admit a finite filtration with subquotients isomorphic to equivariant Verma modules. Since $\mathcal{O}_R$ is abelian and $\mathcal{M}_R$ is closed under extensions in $\mathcal{O}_R$, the category $\mathcal{M}_R$ inherits an exact structure.

For an equivalence class $\Lambda \in \mathfrak{h}^\star/\sim$, let $\mathcal{O}_{R,\Lambda}$, respectively, $\mathcal{M}_{R,\Lambda}$, be the full subcategory of $\mathcal{O}_R$, respectively, $\mathcal{M}_R$, consisting of all objects M such that the highest weight of every simple subquotient of M lies in Λ. Then there are block decompositions, according to the following two results.

Proposition 8.1 [Fiebig 2003, Proposition 2.8; Fiebig 2008a, Theorem 6.1]. *The functors*

$$\prod_{\Lambda\in\mathfrak{h}^\star/\sim} \mathcal{O}_{R,\Lambda} \to \mathcal{O}_R, \quad \{M_\Lambda\} \mapsto \bigoplus_{\Lambda\in\mathfrak{h}^\star/\sim} M_\Lambda$$

and

$$\prod_{\Lambda\in\mathfrak{h}^\star/\sim} \mathcal{M}_{R,\Lambda} \to \mathcal{M}_R, \quad \{M_\Lambda\} \mapsto \bigoplus_{\Lambda\in\mathfrak{h}^\star/\sim} M_\Lambda$$

are equivalences of categories.

Now it is important to notice that we could have substituted S by the local algebra R in the constructions and definitions we have considered, and all the results of this paper would have still worked. Let us denote by $\mathcal{Z}_R$ the R-version of the structure algebra of $\mathcal{G}(\Lambda)$ and by $\mathcal{V}_{R,\Lambda}$ the category of $\mathcal{Z}_R$-modules admitting a Verma flag. The main result of [Fiebig 2008a] is the following one:

Theorem 8.2 [Fiebig 2008a, Theorem 7.1]. *There is an equivalence of exact categories*

$$\mathbb{V} : \mathcal{M}_{R,\Lambda} \to \mathcal{V}_{R,\Lambda}.$$

Projective objects. For $\nu \in \Lambda$, let $\Lambda^{\leq\nu} := \{\lambda \in \Lambda \mid \lambda \leq \nu\}$. We want to consider a truncated version of $\mathcal{M}_{R,\Lambda}$:

$$\mathcal{M}_{R,\Lambda^{\leq\nu}} = \left\{ M \in \mathcal{M}_{R,\Lambda} \;\middle|\; \big(M : M_R(\mu)\big) \neq 0 \text{ only if } \mu \in \Lambda^{\leq\nu} \right\}.$$

As a reference for the truncated category $\mathcal{O}$, we address the reader to [Rocha-Caridi and Wallach 1982], where it was introduced.

Denote by $\mathcal{V}_{R,\Lambda^{\leq\nu}}$ the category of sheaves on the moment graph $\mathcal{G}(\Lambda)^{\leq\nu}$, obtained by restricting the set of vertices of $\mathcal{G}(\Lambda)$ to $\Lambda^{\leq\nu}$. By [Fiebig 2006, Proposition 3.11], the functor $\mathbb{V}$ restricts to a functor $\mathbb{V}^{\leq\nu} : \mathcal{M}_{R,\Lambda^{\leq\nu}} \to \mathcal{V}_{R,\Lambda^{\leq\nu}}$, which is also an equivalence of categories.

Let $\mathcal{H}^J_R$ denote the R-version of the category of special modules, and let $\mathcal{H}^J_{R,\Lambda^{\leq\nu}}$ be the subcategory of $\mathcal{H}^J_R$ consisting of modules having support on $\mathcal{G}(\Lambda)^{\leq\nu}$. From Theorem 6.4, a module $M \in \mathcal{V}_{R,\Lambda^{\leq\nu}}$ is indecomposable and projective if and only if there exist a $w \in \Lambda^{\leq\nu}$ and a $k \in \mathbb{Z}$ such that $M \cong B^J(w)\langle k\rangle$ and, by Proposition 6.5, there exists one and only one indecomposable $M \in \mathcal{H}^J_{R,\Lambda^{\leq\nu}}$ isomorphic to $B^J(w)$. In summary:

Proposition 8.3. *Let $P \in \mathcal{M}_{R,\Lambda^{\leq\nu}}$. Then P is indecomposable, projective if and only if $\mathbb{V}P$ is an indecomposable special module.*

For λ_0 regular, that is, $\mathrm{Stab}_{\mathcal{W}}\, \lambda_0 = \{e\}$, this was already proven in [Fiebig 2008b] and used in [Fiebig 2011], where the interchange between local and global descriptions of the image of the projective modules under $\mathbb{V}$ played a fundamental role.

Acknowledgements

I would like to acknowledge Peter Fiebig for useful discussions and Winston Fairbairn for helpful conversations. I owe many thanks to Michael Ehrig and Ben Salisbury for their careful proofreading.

References

[Beilinson and Bernstein 1981] A. Beilinson and J. Bernstein, "Localisation de g-modules", *C. R. Acad. Sci. Paris Sér. I Math.* **292**:1 (1981), 15–18. MR 82k:14015 Zbl 0476.14019

[Bernstein et al. 1976] I. N. Bernstein, I. M. Gelfand, and S. I. Gelfand, "A certain 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 53 #10880 Zbl 0353.18013

[Björner and Brenti 2005] A. Björner and F. Brenti, *Combinatorics of Coxeter groups*, Graduate Texts in Mathematics **231**, Springer, New York, 2005. MR 2006d:05001 Zbl 1110.05001

[Braden and MacPherson 2001] T. Braden and R. MacPherson, "From moment graphs to intersection cohomology", *Math. Ann.* **321**:3 (2001), 533–551. MR 2003g:14030 Zbl 1077.14522

[Brylinski and Kashiwara 1980] J.-L. Brylinski and M. Kashiwara, "Démonstration de la conjecture de Kazhdan–Lusztig sur les modules de Verma", *C. R. Acad. Sci. Paris Sér. A-B* **291**:6 (1980), A373–A376. MR 81k:17004 Zbl 0457.22012

[Deodhar 1987] V. V. Deodhar, "On some geometric aspects of Bruhat orderings, II: The parabolic analogue of Kazhdan–Lusztig polynomials", *J. Algebra* **111**:2 (1987), 483–506. MR 89a:20054 Zbl 0656.22007

[Elias and Williamson 2014] B. Elias and G. Williamson, "The Hodge theory of Soergel biomodules", *Ann. of Math.* **180**:3 (2014), 1089–1136. arXiv 1212.0791

[Fiebig 2003] P. Fiebig, "Centers and translation functors for the category $\mathcal{O}$ over Kac–Moody algebras", *Math. Z.* **243**:4 (2003), 689–717. MR 2004c:17051 Zbl 1021.17007

[Fiebig 2006] P. Fiebig, "The combinatorics of category $\mathcal{O}$ over symmetrizable Kac–Moody algebras", *Transform. Groups* **11**:1 (2006), 29–49. MR 2006k:17040 Zbl 1122.17016

[Fiebig 2008a] P. Fiebig, "Sheaves on moment graphs and a localization of Verma flags", *Adv. Math.* **217**:2 (2008), 683–712. MR 2008m:17044 Zbl 1140.14044

[Fiebig 2008b] P. Fiebig, "The combinatorics of Coxeter categories", *Trans. Amer. Math. Soc.* **360**:8 (2008), 4211–4233. MR 2009g:20087 Zbl 1160.20032

[Fiebig 2009] P. Fiebig, "Moment graphs in representation theory and geometry", Lecture script, 2009, http://tinyurl.com/Fiebig2009.

[Fiebig 2011] P. Fiebig, "Sheaves on affine Schubert varieties, modular representations, and Lusztig's conjecture", *J. Amer. Math. Soc.* **24**:1 (2011), 133–181. MR 2012a:20072 Zbl 1270.20053

[Goresky et al. 1998] M. Goresky, R. Kottwitz, and R. MacPherson, "Equivariant cohomology, Koszul duality, and the localization theorem", *Invent. Math.* **131**:1 (1998), 25–83. MR 99c:55009 Zbl 0897.22009

[Humphreys 1990] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics **29**, Cambridge University Press, 1990. MR 92h:20002 Zbl 0725.20028

[Kazhdan and Lusztig 1979] D. Kazhdan and G. Lusztig, "Representations of Coxeter groups and Hecke algebras", *Invent. Math.* **53**:2 (1979), 165–184. MR 81j:20066 Zbl 0499.20035

[Kazhdan and Lusztig 1980] D. Kazhdan and G. Lusztig, "Schubert varieties and Poincaré duality", pp. 185–203 in *Geometry of the Laplace operator* (Honolulu, HI, 1979), edited by R. Osserman and A. Weinstein, Proc. Sympos. Pure Math. **36**, Amer. Math. Soc., Providence, R.I., 1980. MR 84g:14054 Zbl 0461.14015

[Lanini 2012] M. Lanini, "Kazhdan–Lusztig combinatorics in the moment graph setting", *J. Algebra* **370** (2012), 152–170. MR 2966832 Zbl 06162673

[Mazorchuk 2012] V. Mazorchuk, *Lectures on algebraic categorification*, Eur. Math. Soc., Zürich, 2012. MR 2918217 Zbl 1238.18001

[Mazorchuk and Stroppel 2005] V. Mazorchuk and C. Stroppel, "Translation and shuffling of projectively presentable modules and a categorification of a parabolic Hecke module", *Trans. Amer. Math. Soc.* **357**:7 (2005), 2939–2973. MR 2006g:17012 Zbl 1095.17001

[Mazorchuk and Stroppel 2008] V. Mazorchuk and C. Stroppel, "Categorification of (induced) cell modules and the rough structure of generalised Verma modules", *Adv. Math.* **219**:4 (2008), 1363–1426. MR 2010a:20014 Zbl 1234.17007

[Quillen 1973] D. Quillen, "Higher algebraic K-theory, I", pp. 85–147 in *Algebraic K-theory, I: Higher K-theories* (Seattle, WA, 1972), edited by H. Bass, Lecture Notes in Math. **341**, Springer, Berlin, 1973. MR 49 #2895 Zbl 0292.18004

[Rocha-Caridi and Wallach 1982] A. Rocha-Caridi and N. R. Wallach, "Projective modules over graded Lie algebras, I", *Math. Z.* **180**:2 (1982), 151–177. MR 83h:17018 Zbl 0467.17006

[Soergel 1990] W. Soergel, "Kategorie $\mathbb{O}$, perverse Garben und Moduln über den Koinvarianten zur Weylgruppe", *J. Amer. Math. Soc.* **3**:2 (1990), 421–445. MR 91e:17007 Zbl 0747.17008

[Soergel 1997] W. Soergel, "Kazhdan–Lusztig polynomials and a combinatoric[s] for tilting modules", *Represent. Theory* **1** (1997), 83–114. MR 98d:17026 Zbl 0886.05123

[Soergel 2007] W. Soergel, "Kazhdan-Lusztig–Polynome und unzerlegbare Bimoduln über Polynomringen", *J. Inst. Math. Jussieu* **6**:3 (2007), 501–525. MR 2009c:20009 Zbl 1192.20004

[Williamson 2011] G. Williamson, "Singular Soergel bimodules", *Int. Math. Res. Not.* **2011**:20 (2011), 4555–4632. MR 2844932 Zbl 1236.18009

Received June 2, 2013. Revised July 25, 2013.

MARTINA LANINI
DEPARTMENT MATHEMATIK
FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
CAUERSTR. 11
91058 ERLANGEN
GERMANY
lanini@math.fau.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.445

UNITARY REPRESENTATIONS OF GL(n, K) DISTINGUISHED BY A GALOIS INVOLUTION FOR A p-ADIC FIELD K

NADIR MATRINGE

Let F be a p-adic field and K a quadratic extension of F. Using Tadić's classification of the unitary dual of GL(n, K), we give the list of irreducible unitary representations of this group distinguished by GL(n, F) in terms of distinguished discrete series. It is known that a generalised Steinberg representation $St(\rho, k)$ is distinguished if and only if the cuspidal representation ρ is η^{k-1}-distinguished for η, the character of F^* with kernel consisting of the norms of K^*. This actually gives a classification of distinguished unitary representations in terms of distinguished cuspidal representations.

Introduction

In the present work, for F a p-adic field and K a quadratic extension of F, smooth and complex unitary (which will be synonymous with unitarisable for us), we study representations of GL(n, K) which admit on their space a nonzero invariant linear form under GL(n, F). These unitary representations are called GL(n, F)-distinguished (or simply distinguished) and are conjectured to be the unitary part of the image of a functorial lift, in the Langlands' program, from $U(n, K/F)$ to GL(n, K).

Distinguished generic representations of GL(n, K) have been classified in [Matringe 2011b], in terms of distinguished quasi-discrete series, using Zelevinsky's classification of generic representations. Here we do the same for distinguished irreducible unitary representations using Tadić's classification of irreducible unitary representations. Our main result (Theorem 2.13) is similar to the main result of [Matringe 2011b]. However, to extend the result from generic unitary to irreducible unitary representations, we use different techniques. Our main tools are the Bernstein–Zelevinksy derivative functors, and we apply ideas from [Bernstein 1984]. For instance, the building blocks for unitary representations (the so-called Speh representations) are not parabolically induced; hence one needs new methods to deal with these representations. That is what we do in the second part of Section 2 to obtain a definitive statement in Corollary 2.9, which we state here as a theorem.

MSC2010: primary 22E50; secondary 22E35.
Keywords: distinguished representations of p-adic groups.

Theorem. *Let k and m be two positive integers, and let n be equal to km. If Δ is a discrete series of* $\mathrm{GL}(m, K)$ *and $u(\Delta, k)$ is the corresponding Speh representation of* $\mathrm{GL}(n, K)$ (*see Definition 1.3*), *then $u(\Delta, k)$ is distinguished if and only if Δ is.*

One direction is given by the fact that if π is a distinguished irreducible unitary representation, then it is also the case of its highest shifted derivative (see Proposition 2.4). The other direction is a nontrivial generalisation of the following simple observation: if ρ is a distinguished cuspidal representation of $\mathrm{GL}(n, K)$, then it is known that the parabolically induced representation $\nu^{1/2}\rho \times \nu^{-1/2}\rho$ is distinguished, and it is also known that its irreducible submodule $St(\rho, 2)$ is not distinguished. Hence its quotient $u(\rho, 2)$, which is a Speh representation of $\mathrm{GL}(2n, K)$, is distinguished. The case of general irreducible unitary representations of $\mathrm{GL}(n, K)$ distinguished by $\mathrm{GL}(n, F)$ is treated in the third part of Section 2.We obtain the main result of the paper in Theorem 2.13. Denoting by σ the nontrivial element of the Galois group of K over F, by $\pi^\vee$ the smooth contragredient of a representation π of $\mathrm{GL}(n, K)$, and by π^σ the representation $\pi \circ \sigma$, its statement is as follows.

Theorem. *Let n be a positive integer and π an irreducible unitary representation of* $\mathrm{GL}(n, K)$. *By Tadić's classification* (*see Theorem 1.9*), *the representation π is a commutative product* (*in the sense of normalised parabolic induction*) *of representations of the form $u(\Delta, k)$ for $k > 0$ and Δ a discrete series, and representations of the form $\pi(u(\Delta, k), \alpha)$* (*see Definition 1.8*) *for Δ and k as before and α an element of* $(0, 1/2)$. *Then the representation π is distinguished if and only if $\pi^\vee$ is isomorphic to π^σ and the Speh representations $u(\Delta, k)$ occurring in the product π with odd multiplicity are distinguished.*

1. Preliminaries

Basic facts and notations. First, in the following, we fix a nonarchimedean local field F of characteristic 0 and an algebraic closure $\bar{F}$ of F. We denote by K a quadratic extension of F in $\bar{F}$. We denote by $\mathfrak{O}_F$ and $\mathfrak{P}_F$ the ring of integers of F and the unique maximal ideal of F respectively. We similarly define $\mathfrak{O}_K$ and $\mathfrak{P}_K$. We denote by $|\cdot|_F$ and $|\cdot|_K$ the normalised absolute values, which satisfy $|x|_K = |x|_F^2$ for x in F. We fix a nontrivial smooth character θ of K which is trivial on F. We denote by σ the nontrivial element of the Galois group $\mathrm{Gal}_F(K)$ of K over F and by η the quadratic character of F^*, whose kernel is the set of norms of K^*. For n and $m \geq 1$, we denote by $\mathcal{M}_{n,m}$ the space of matrices $\mathcal{M}(n, m, K)$, by $\mathcal{M}_n$ the algebra $\mathcal{M}_{n,n}$, and by G_n the group of invertible elements in $\mathcal{M}_n$. We will denote by G_0 the trivial group. If m belongs to $\mathcal{M}_n$, we denote by m^σ the matrix obtained from m by applying σ to each entry. If S is a subset of $\mathcal{M}_n$, we denote by S^σ the subset of S consisting of elements fixed by σ. For

$m \in \mathcal{M}_n$, we denote by $|m|_K$ or $\nu_K(m)$ the real number $|\det m|_K$, and we define similarly $|m|_F$ or $\nu_F(m)$ for m in $\mathcal{M}_n^\sigma$.

When G is a closed subgroup of G_n, we denote by $\mathrm{Alg}(G)$ the category of smooth complex G-modules. If (π, V) belongs to $\mathrm{Alg}(G)$, H is a closed subgroup of G, and χ is a character of H, we denote by $V(H, \chi)$ the subspace of V generated by vectors of the form $\pi(h)v - \chi(h)v$ for h in H and v in V. This space is stable under the action of the subgroup $N_G(\chi)$ of the normalizer $N_G(H)$ of H in G, which fixes χ. We denote by δ_H the positive character of $N_G(H)$ such that if μ is a right Haar measure on H and int is the action of $N_G(H)$ on smooth functions f with compact support in H, given by $(\mathrm{int}(n)f)(h) = f(n^{-1}hn)$, then $\mu \circ \mathrm{int}(n) = \delta_H(n)\mu$ for n in $N_G(H)$. The space $V(H, \chi)$ is $N_G(\chi)$-stable. Thus, if L is a closed subgroup of $N_G(\chi)$ and δ' is a (smooth) character of L (which will be a normalising character dual to that of normalised induction later), the quotient $V_{H,\chi} = V/V(H, \chi)$ (which we simply denote by V_H when χ is trivial) becomes a smooth L-module for the (normalised) action $l.(v + V(H, \chi)) = \delta'(l)\pi(l)v + V(H, \chi)$ of L on $V_{H,\chi}$. If (ρ, W) belongs to $\mathrm{Alg}(H)$, we define the objects

$$(\mathrm{ind}_H^G(\rho), V_c = \mathrm{ind}_H^G(W)) \quad \text{and} \quad (\mathrm{Ind}_H^G(\rho), V = \mathrm{Ind}_H^G(W))$$

of $\mathrm{Alg}(G)$ as follows. The space V is the space $\mathscr{C}^\infty(H\backslash G, \rho)$ of smooth functions from G to W fixed under right translation by the elements of a compact open subgroup U_f of G, and satisfying $f(hg) = \rho(h)f(g)$ for all h in H and g in G. The space V_c is the subspace $\mathscr{C}_c^\infty(H\backslash G, \rho)$ of V consisting of functions with support compact mod H. In both cases, the action of G is by right translation on the functions. By definition, the real part $\mathrm{Re}(\chi)$ of a character χ of F^* is the real number r such that $|\chi(t)|_{\mathbb{C}} = |t|^r$, where $|z|_{\mathbb{C}} = \sqrt{z\bar{z}}$ for z in $\mathbb{C}$.

***Irreducible representations of* GL(*n*).** We will only consider smooth representations of G_n and its closed subgroups. We denote by A_n the maximal torus of diagonal matrices in G_n. It will sometimes be useful to parametrise A_n with simple roots, that is, to write an element $t = \mathrm{diag}(t_1, \dots, t_n)$ of A_n as $t = z_1 \cdots z_n$, where $z_n = t_n I_n$ and $z_i = \mathrm{diag}((t_i/t_{i+1})I_i, I_{n-i})$ belongs to the centre of G_i embedded in G_n, which we denote Z_i. For $z_i = \mathrm{diag}(t_i I_i, I_{n-i})$ in Z_i, we denote t_i by $t(z_i)$. If $n \geq 1$, let $\bar{n} = (n_1, \dots, n_t)$ be a partition of n of length t (i.e., an ordered set of t positive integers whose sum is n). We denote by $M_{\bar{n}}$ the Levi subgroup of G_n of matrices of the form $\mathrm{diag}(g_1, \dots, g_t)$ with each g_i in G_{n_i}, by $N_{\bar{n}}$ the unipotent subgroup of matrices of the form

$$\begin{pmatrix} I_{n_1} & \star & \star \\ & \ddots & \star \\ & & I_{n_t} \end{pmatrix},$$

and by $P_{\bar{n}}$ the standard parabolic subgroup $M_{\bar{n}}N_{\bar{n}}$ (where $M_{\bar{n}}$ normalises $N_{\bar{n}}$). Note that $M_{(1,\dots,1)}$ is equal to A_n, and we set $N_{(1,\dots,1)} = N_n$. For each i, let π_i be a smooth representation of G_{n_i}. Then the tensor product $\pi_1 \otimes \cdots \otimes \pi_t$ is a representation of $M_{\bar{n}}$, which can be considered as a representation of $P_{\bar{n}}$ that is trivial on $N_{\bar{n}}$. We will use the product notation

$$\pi_1 \times \cdots \times \pi_t = \mathrm{Ind}_{P_{\bar{n}}}^{G_n}(\delta_{P_{\bar{n}}}^{1/2} \pi_1 \otimes \cdots \otimes \pi_t)$$

for the normalised parabolic induction.

We say that an irreducible representation (ρ, V) of G_n is cuspidal if the Jacquet module $V_{N_{\bar{n}}}$ is zero when $\bar{n}$ is a proper partition of n. Suppose that $\bar{n} = (m, \dots, m)$ is a partition of n of length l and that ρ is a cuspidal representation of G_m. Let a and b be two integers with $a \leq b$ such that $b - a + 1 = l$. Then [Zelevinsky 1980, Theorem 9.3] implies that the G_n-module $\nu_K^a \rho \times \cdots \times \nu_K^b \rho$ has a unique irreducible quotient, which we denote by $\Delta(\rho, b, a)$. We call it a segment or a quasi-discrete series of G_n. If, in addition, a quasi-discrete series is unitary (which amounts to saying that its central character is unitary), we will call it a discrete series or a unitary segment. We will sometimes write $St(\rho, l) = \Delta(\rho, (l-1)/2, -(l-1)/2)$.

We end this section with a word about induced representations of Langlands' type and their quotients.

Definition 1.1. Let $\Delta_1, \dots, \Delta_t$ be segments of $G_{n_1}, \dots, G_{n_t}$ respectively, and suppose that the central characters satisfy the relation $\mathrm{Re}(c_{\Delta_i}) \geq \mathrm{Re}(c_{\Delta_{i+1}})$. Let $n = n_1 + \cdots + n_t$. Then the representation $\Delta_1 \times \cdots \times \Delta_t$ of G_n is said to be induced of Langlands' type.

The following result is well known and can be found in [Rodier 1982].

Proposition 1.2. *Let $\pi = \Delta_1 \times \cdots \times \Delta_t$ be an induced representation of Langlands' type as above. Then π has a unique irreducible quotient, which we denote by $L(\Delta_1, \dots, \Delta_t)$. If $\Delta'_1, \dots, \Delta'_s$ are other segments with $\mathrm{Re}(c_{\Delta'_j}) \geq \mathrm{Re}(c_{\Delta'_{j+1}})$ such that $L(\Delta_1, \dots, \Delta_t) = L(\Delta'_1, \dots, \Delta'_s)$, then we have the equality of nonordered sets $\{\Delta_1, \dots, \Delta_t\} = \{\Delta'_1, \dots, \Delta'_s\}$.*

A particular class of Langlands' quotients is the class of Speh representations, which are the building blocks of the unitary dual of G_n in Tadić's classification.

Definition 1.3. Let k and m be two positive integers, and set $n = km$. If Δ is a segment of G_m, we denote by $u(\Delta, k)$ the representation $L(\nu_K^{(k-1)/2}\Delta, \dots, \nu_K^{(1-k)/2}\Delta)$ of G_n.

We now recall some basic facts about Bernstein–Zelevinksy derivatives.

Derivatives. We define a character of N_n, denoted again by θ, by the formula $\theta(m) = \theta(\sum_{i=1}^{n-1} m_{i,i+1})$. For $n \geq 2$, we denote by U_n the group of matrices of the form $\begin{pmatrix} I_{n-1} & v \\ & 1 \end{pmatrix}$. For $n > k \geq 1$, the group G_k embeds naturally in G_n and is given by matrices of the form $\mathrm{diag}(g, I_{n-k})$. We denote by P_n the mirabolic subgroup $G_{n-1}U_n$ of G_n for $n \geq 2$, and we set $P_1 = \{1_{G_1}\}$. If one sees P_{n-1} as a subgroup of G_{n-1} itself embedded in G_n, then P_{n-1} is the normaliser of $\theta_{|U_n}$ in G_{n-1} (i.e., if $g \in G_{n-1}$, then $\theta(g^{-1}ug)$ for all $u \in U_n$ if and only if $g \in P_{n-1}$). We define the following functors:

- The functor Φ^+ from $\mathrm{Alg}(P_{k-1})$ to $\mathrm{Alg}(P_k)$ such that for π in $\mathrm{Alg}(P_{k-1})$, one has $\Phi^+\pi = \mathrm{ind}_{P_{k-1}U_k}^{P_k}(\delta_{U_k}^{1/2}\pi \otimes \theta)$.
- The functor $\hat{\Phi}^+$ from $\mathrm{Alg}(P_{k-1})$ to $\mathrm{Alg}(P_k)$ such that for π in $\mathrm{Alg}(P_{k-1})$, one has $\hat{\Phi}^+\pi = \mathrm{Ind}_{P_{k-1}U_k}^{P_k}(\delta_{U_k}^{1/2}\pi \otimes \theta)$.
- The functor Φ^- from $\mathrm{Alg}(P_k)$ to $\mathrm{Alg}(P_{k-1})$ such that if (π, V) is a smooth P_k-module, $\Phi^- V = V_{U_k,\theta}$, and P_{k-1} acts on $\Phi^-(V)$ by $\Phi^-\pi(p)(v+V(U_k,\theta)) = \delta_{U_k}(p)^{-1/2}\pi(p)(v+V(U_k,\theta))$.
- The functor Ψ^- from $\mathrm{Alg}(P_k)$ to $\mathrm{Alg}(G_{k-1})$ such that if (π, V) is a smooth P_k-module, $\Psi^- V = V_{U_k,1}$, and G_{k-1} acts on $\Psi^-(V)$ by $\Psi^-\pi(g)(v+V(U_k,1)) = \delta_{U_k}(g)^{-1/2}\pi(g)(v+V(U_k,1))$.
- The functor Ψ^+ from $\mathrm{Alg}(G_{k-1})$ to $\mathrm{Alg}(P_k)$ such that for π in $\mathrm{Alg}(G_{k-1})$, one has $\Psi^+\pi = \mathrm{ind}_{G_{k-1}U_k}^{P_k}(\delta_{U_k}^{1/2}\pi \otimes 1) = \delta_{U_k}^{1/2}\pi \otimes 1$.

If τ is a representation of P_n (or a representation of G_n, which we consider as a P_n-module by restriction), the representation $\tau^{(k)}$ of G_{n-k} will be defined as $\Psi^-(\Phi^-)^{k-1}\tau$ and will be called the k-th derivative of τ. It is shown in [Bernstein and Zelevinsky 1977, Section 3.5] that these representations give a natural filtration of any P_n-module.

Lemma 1.4. *If τ is an object of $\mathrm{Alg}(P_n)$, then τ has a natural filtration of P_n-modules $0 \subset \tau_n \subset \cdots \subset \tau_1 = \tau$, where $\tau_k = \Phi^{+k-1}\Phi^{-k-1}\tau$. Moreover, the quotient τ_k/τ_{k+1} is isomorphic to $(\Phi^+)^{k-1}\Psi^+\tau^{(k)}$ as a P_n-module.*

It is shown in [Zelevinsky 1980, Section 8] that if π is an irreducible representation of G_n, then its highest derivative π^-, which is the derivative $\pi^{(k)}$ for $k \leq n$ that is maximal for the condition $\pi^{(k)} \neq 0$, is an irreducible representation of G_{n-k}. The following lemma is an immediate consequence of [Bernstein and Zelevinsky 1977, Lemma 4.5].

Lemma 1.5. *Let π_i be an irreducible representation of G_{n_i} for positive integers $n_1, \dots, n_t$. Then the highest derivative of $\pi_1 \times \cdots \times \pi_t$ is the representation $\pi_1^- \times \cdots \times \pi_t^-$.*

As we study unitary representations, we will need some further properties of these derivatives, which are extracted from [Bernstein 1984]. First, as in this reference, we introduce the following definition.

Definition 1.6. Let τ be a P_n-module. We denote by $\tau^{[k]}$ the representation $\nu_K^{1/2}\tau^{(k)}$ of G_{n-k} and call it the k-th shifted derivative of τ. We denote by $\tau^{[-]}$ the highest shifted derivative of τ.

We then recall the following consequence of the unitarisability criterion given in [ibid., Section 7.3].

Proposition 1.7. *If π is an irreducible unitary representation of G_n with highest derivative $\pi^{(h)}$, then $\pi^{[h]}$ is unitary and the central characters of the irreducible subquotients of $\pi^{[k]}$ all have positive real parts for* $0 < k < h$.

***Unitary representations of* GL(n).** We now recall results from [Tadić 1986] about the classification of irreducible unitary representations of G_n.

Definition 1.8. For $\alpha \in \mathbb{R}$, $m > 0$, $k > 0$, and Δ a segment of G_m, we denote by $\pi(u(\Delta,k),\alpha)$ the representation $\nu_K^{\alpha}u(\Delta,k)\times\nu_K^{-\alpha}u(\Delta,k)$ of G_n for $n = 2mk$.

Theorem 1.9 [Tadić 1986, Theorem D]. *Let π be an irreducible unitary representation of G_n. Then there is a partition $(n_1,\dots,n_t)$ of n and representations π_i of G_{n_i}, each of which is either of the form $\pi(u(\Delta_i,k_i),\alpha_i)$ for Δ_i a unitary segment*, $k_i \geq 1$, *and* $0<\alpha_i<1/2$ *or of the form $u(\Delta_i,k_i)$ for Δ_i a unitary segment and* $k_i \geq 1$, *such that $\pi = \pi_1\times\cdots\times\pi_t$. Moreover*, *the representation π is equal to $\pi'_1\times\cdots\times\pi'_s$ for representations π'_j of the same type as the representations π_i if and only if $\{\pi_1,\dots,\pi_t\} = \{\pi'_1,\dots,\pi'_s\}$ as nonordered sets.*

If all the representations π_i in the above theorem are such that $k_i = 1$, we say that π is a *generic* unitary representation of G_n.

We will also need the description of the composition series of the so-called end of complementary series, which is proved in [Tadić 1987a] (see [Badulescu 2011, Theorem 2] for a quick proof). If Δ is the segment $St(\rho,l)$ for $l \geq 1$, we write $\Delta_+ = St(\rho,l+1)$ and $\Delta_- = St(\rho,l-1)$, where $St(\rho,0)$ is $\mathbf{1}_{G_0}$ by convention.

Theorem 1.10. *Let m be a positive integer*, *Δ a segment of G_m*, *$k \geq 2$ an integer*, *and $n = 2mk$. The representation $\pi(u(\Delta,k),1/2)$ of G_n is of length* 2, *and its irreducible subquotients are $u(\Delta_-,k)\times u(\Delta_+,k)$ and $u(\Delta,k-1)\times u(\Delta,k+1)$.*

Finally, we recall the formula which gives the highest shifted derivative of a Speh representation, from [Tadić 1987b, Section 6.1] (see [Offen and Sayag 2008, (3.3)] for the proof).

Proposition 1.11. *Let $m > 0$ and $k > 1$ be two integers*, *and let Δ be a segment of G_m. The highest shifted derivative of the representation $u(\Delta,k)$ is equal to $u(\Delta,k)^{[m]} = u(\Delta,k-1)$. The highest* (*shifted*) *derivative of Δ is equal to $\mathbf{1}_{G_0}$.*

Distinguished representations of $\mathbf{GL}(n)$**.** In this paragraph, we recall results from [Matringe 2011b]. First, we introduce some notations and definitions.

Definition 1.12. Let G be a closed subgroup of G_n, H a closed subgroup of G, and χ a character of H. We say that a representation π in $\mathrm{Alg}(G)$ is (H, χ)-distinguished if the space $\mathrm{Hom}_H(\pi, \chi)$ is nonzero. If H is clear, we say χ-distinguished instead of (H, χ)-distinguished, and if χ is trivial, we say H-distinguished (or distinguished if H is clear). If $G = G_n$ and $H = G_n^\sigma$, we will sometimes say (σ, χ)-distinguished instead of (H, χ)-distinguished, and if χ is trivial, we will simply say σ-distinguished.

We recall the following general facts from [Flicker 1991] about σ-distinguished representations of G_n. We denote by π^σ the representation $g \mapsto \pi(g^\sigma)$ for π a representation of G_n.

Proposition 1.13. *Let $n \geq 1$ be an integer and π be an irreducible representation of G_n. If π is σ-distinguished, then $\pi^\vee = \pi^\sigma$ and* $\mathrm{Hom}_{G_n^\sigma}(\pi, \mathbf{1})$ *is of dimension* 1.

We now introduce the class of σ-induced irreducible unitary representations of G_n. They will turn out to be the σ-distinguished irreducible unitary representations of G_n.

Definition 1.14. For $n \geq 1$, let π be an irreducible unitary representation

$$\pi = u(\Delta_1, k_1) \times \cdots \times u(\Delta_s, k_s) \times \pi\bigl(u(\Delta_{s+1}, k_{s+1}), \alpha_{s+1}\bigr) \times \cdots \times \pi\bigl(u(\Delta_t, k_t), \alpha_t\bigr)$$

of G_n with unitary segments Δ_i, positive integers k_i, and $\alpha_i \in (0, 1/2)$. The representation π is said to be σ-induced if it satisfies $\pi^\vee = \pi^\sigma$ and if for every $i \leq s$ such that $u(\Delta_i, k_i)$ occurs with odd multiplicity in the product π, the segment Δ_i is σ-distinguished.

Remark 1.15. Maybe the preceding definition is not completely transparent to the reader. Let us try to explain what σ-induced irreducible unitary representations look like. Let

$$\pi = u(\Delta_1, k_1) \times \cdots \times u(\Delta_t, k_s) \times \pi\bigl(u(\Delta_{s+1}, k_{s+1}), \alpha_{k_{s+1}}\bigr) \times \cdots \times \pi\bigl(u(\Delta_t, k_t), \alpha_{k_t}\bigr)$$

be an irreducible unitary representation of G_n. First, if one has $\pi^\vee = \pi^\sigma$ (call this relation σ-self-duality), then it means the two following things:

(a) For i between 1 and s, either $u(\Delta_i, k_i)$ is σ-self-dual or, if this relation is not satisfied, there exists $j \neq i$ between 1 and s such that $u(\Delta_j, k_j)^\vee = u(\Delta_i, k_i)^\sigma$.

(b) For i between $s+1$ and t, either $\pi(u(\Delta_i, k_i), \alpha_i)$ is σ-self-dual or, if this relation is not satisfied, there exists $j \neq i$ between $s+1$ and t such that $\pi(u(\Delta_j, k_j), \alpha_j)^\vee = \pi(u(\Delta_i, k_i), \alpha_i)^\sigma$.

In (a) above, if you have $u(\Delta_i, k_i)^\vee = u(\Delta_i, k_i)^\sigma$ which occurs with multiplicity at least 2, that is, if there is $j \neq i$ between 1 and s such that $u(\Delta_j, k_j) = u(\Delta_i, k_i)$, then one has $u(\Delta_j, k_j)^\vee = u(\Delta_i, k_i)^\sigma$. Hence (a) can also be stated as:

(a′) $u(\Delta_1,k_1)\times\cdots\times u(\Delta_s,k_s)$ is a product of representations of the form $u(\Delta_i,k_i)\times (u(\Delta_i,k_i)^\vee)^\sigma$ and of σ-self dual representations $u(\Delta_j,k_j)$ which occur with odd multiplicity.

Now in (b), if $\pi(u(\Delta_i,k_i),\alpha_i)$ is σ-self dual, it is equal to

$$\nu_K^{\alpha_i}u(\Delta_i,k_i)\times\big((\nu_K^{\alpha_i}u(\Delta_i,k_i))^\vee\big)^\sigma$$

(because $\Delta_i^\vee$ must be equal to Δ_i^σ). All in all, π is σ-self dual if and only if it is a product of representations of the form

$$\nu_K^{\alpha}u(\Delta,k)\times\big((\nu_K^{\alpha}u(\Delta,k))^\vee\big)^\sigma$$

for $0\le\alpha<1/2$, Δ a discrete series, and k a positive integer (we allow here α to be equal to zero in order to take in account representations $u(\Delta_i,k_i)\times(u(\Delta_i,k_i)^\vee)^\sigma$ occurring in (a′)), of representations of the form

$$\pi(u(\Delta,k),\alpha)\times\big(\pi(u(\Delta,k),\alpha)^\vee\big)^\sigma$$

for α in $(0,1/2)$ and Δ and k as above, and of representations of the form $u(\Delta',k')$ (Δ' unitary and $k'>0$) occurring with odd multiplicity and which are σ-self dual. In this situation, π is σ-induced if and only if these representations $u(\Delta',k')$ are such that Δ' is σ-distinguished.

Theorem 5.2 of [Matringe 2011b] then classifies distinguished generic representations.

Theorem 1.16. *For $n\ge1$, a generic unitary representation of G_n is σ-distinguished if and only if it is σ-induced.*

We also recall [Matringe 2009, Corollary 3.1] about distinction of discrete series.

Proposition 1.17. *Let ρ be a cuspidal representation of G_r for $r\ge1$ and $\Delta=St(\rho,l)$ for $l\ge1$. The segment Δ of G_{lr} is σ-distinguished if and only if ρ is (σ,η^{l-1})-distinguished.*

Finally, [Anandavardhanan et al. 2004, Corollary 1.6] says that the segment Δ above cannot be σ-distinguished and (σ,η)-distinguished at the same time. This has the following immediate corollary.

Corollary 1.18. *Let Δ be a segment of G_n for $n\ge2$. Then Δ is σ-distinguished if and only if Δ_+ is (σ,η)-distinguished. In particular, if Δ is distinguished,then Δ_+ is not.*

2. Distinguished unitary representations

We will first prove the convergence of integrals defining invariant linear forms.

Asymptotics in the degenerate Kirillov model. We denote by $N_{n,h}$ the group of matrices $h(a,n) = \left(\begin{smallmatrix} a & x \\ 0 & n \end{smallmatrix}\right)$ with a in G_{n-h}, n in N_h, and x in $\mathcal{M}_{n-h,h}$. It is proved in [Zelevinsky 1980, Section 5] that any irreducible representation π of G_n has a "degenerate Kirillov model" (which is just the standard Kirillov model in the nondegenerate case). This means that the restriction of π to P_n embeds as a unique P_n-submodule $K(\pi,\theta)$ of $(\hat{\Phi}^+)^{h-1}\Psi^+(\pi^{(h)})$, where $\pi^{(h)} = \pi^-$. The space $K(\pi,\theta)$ consists of smooth functions W from P_n to $V_{\pi^{(h)}}$ which are fixed under right translation by an open subgroup U_W and satisfy the relation

$$W(h(a,n)p) = |a|_K^{h/2}\theta(n)\pi^{(h)}(a)W(p)$$

for $h(a,n)$ in $N_{n,h}$ and p in P_n. It can be handy to identify such a function with a map from P_n to $V_{\pi^{[h]}}$ which satisfies the relation

$$W(h(a,n)p) = |a|_K^{(h-1)/2}\theta(n)\pi^{[h]}(a)W(p) \tag{1}$$

for $h(a,n)$ in $N_{n,h}$ and p in P_n.

We now give an asymptotic expansion of the elements of $K(\pi,\theta)$ in terms of the exponents of π. The proof, which is omitted, is an easy adaptation of the proof of [Matringe 2011a, Theorem 2.1]. We write $\mathscr{C}_c^\infty(F,V)$ for the space of smooth functions with compact support from F to a complex vector space V.

Theorem 2.1. *Let π be an irreducible representation of G_n for $n \geq 2$. Let $\pi^{(h)}$ be the highest derivative of π, and let W belong to $K(\pi,\theta)$. We suppose that we have $h \geq 2$, and we denote by $(c_{k,i_k})_{1\leq k\leq r_k}$ the family of central characters of the irreducible subquotients of $\pi^{(k)}$. In this situation, the restriction $W(z_{n-h+1}\cdots z_{n-1})$ of W to the torus $Z_{n-h+1}\cdots Z_{n-1}$ is a linear combination of functions of the form*

$$\prod_{k=n-h+1}^{n-1} [c_{i_k,k}\delta_{U_{k+1}}^{1/2}\cdots\delta_{U_n}^{1/2}](z_k)v_F(z_k)^{m_k}\phi_k(t(z_k))$$

for i_k between 1 and r_k, nonnegative integers m_k, and functions ϕ_k in $\mathscr{C}_c^\infty(F, V_{\pi^{(h)}})$.

From this, we deduce the convergence of the following integrals, which we will need later.

Proposition 2.2. *Let π be an irreducible unitary representation of G_n for $n \geq 1$. Let $\pi^{(h)}$ be the highest derivative of π, and let W belong to $K(\pi,\theta)$. We suppose that there is a nonzero G_{n-h}^σ-invariant linear form L on the space of $\pi^{[h]}$, and for every element W of $K(\pi,\theta)$, we define the map $f_{L,W} = L\circ W$. Then for all W in $K(\pi,\theta)$, the integral*

$$\Lambda(W) = \int_{N_{n,h}^\sigma\backslash P_n^\sigma} f_{L,W}(p)\,dp$$

is absolutely convergent and Λ defines a nonzero P_n^σ-invariant linear form on V_π.

Proof. If h equals 1, then $\Lambda(W)$ is equal to $L(W(I_n))$ up to normalisation, and the result is obvious. For $h \geq 2$, first, thanks to Relation (1), the restriction of the map $f_{L,W}$ to P_n^σ satisfies the relation

$$f_{L,W}(h(a,n)p) = |a|_F^{h-1} f_{L,W}(p)$$

for p in P_n^σ and $h(a,n)$ in $N_{n,h}^\sigma$. We notice that $|a|_F^{h-1}$ is indeed equal to

$$\frac{\delta_{N_{n,h}^\sigma}}{\delta_{P_n^\sigma}}(h(a,n)) = \frac{|a|_F^h}{|a|_F}.$$

Actually, the integral $\Lambda(W)$ is equal to

$$\int_{N_{n-1,h}^\sigma \backslash G_{n-1}^\sigma} f_{L,W}(p)\,dp.$$

Hence, thanks to the Iwasawa decomposition, the integral Λ will converge absolutely for any W in $K(\pi,\theta)$ if and only if the following integral does as well:

$$\int_{Z_{n-h+1}\dots Z_{n-1}} f_{L,W}(z_{n-h+1}\dots z_n)\,\delta^{-1}_{N_{n-1,h}^\sigma}(z_{n-h+1}\dots z_{n-1})\,d^*z_{n-h+1}\dots d^*z_{n-1}$$

for any W in $K(\pi,\theta)$. As $\delta_{N_{n-1,h}^\sigma}(z_{n-h+1}\dots z_{n-1})$ is equal to the product

$$\prod_{k=n-h+1}^{n-1} \delta_{U_{k+1}^\sigma}\dots\delta_{U_{n-1}^\sigma}(z_k) = \prod_{k=n-h+1}^{n-1} \delta_{U_{k+1}}^{1/2}\dots\delta_{U_{n-1}}^{1/2}(z_k)$$

for the z_i in Z_i^σ, we obtain that the integral

$$\int_{Z_{n-h+1}\dots Z_{n-1}} |f_{L,W}(z_{n-h+1}\dots z_n)|\delta^{-1}_{N_{n-1,h}^\sigma}(z_{n-h+1}\dots z_{n-1})\,d^*z_{n-h+1}\dots d^*z_{n-1}$$

is majorized by a sum of integrals of the form

$$\prod_{k=n-h+1}^{n-1} \int_{Z_k} c_{i_k,k}\,\delta_{U_n}^{1/2}(z_k)v_F(z_k)^{m_k} f_k(t(z_k))\,d^*z_k$$
$$= \prod_{k=n-h+1}^{n-1} \int_{Z_k} c_{i_k,k}\,\delta_{U_{k+1}}^{1/2}(z_k)v_F(z_k)^{m_k} f_k(t(z_k))\,d^*z_k$$

for functions $f_k = L\circ\phi_k$ in $\mathscr{C}_c^\infty(F)$, thanks to Theorem 2.1. These last integrals are convergent, as, according to Proposition 1.7, the real part $\mathrm{Re}(c_{i_k,k}\delta_{U_{k+1}}^{1/2})$ is positive. This concludes the proof of the convergence. To show that Λ is nonzero, we just need to remember that π contains as a P_n-submodule the space $(\Phi^+)^{h-1}(\Psi^+(\pi^{(h)}))$ and the restriction to P_n^σ of elements of $(\Phi^+)^{h-1}(\Psi^+(\pi^{(h)}))$ is surjective on the space

$$\mathscr{C}_c^\infty\left(N_{n,h}^\sigma\backslash P_n^\sigma, \frac{\delta_{N_{n,h}^\sigma}}{\delta_{P_n^\sigma}}\pi^{[k]}\otimes\mathbf{1}\right). \qquad \square$$

The case of Speh representations. The aim of this section is to prove that a representation $u(\Delta, k)$ is σ-distinguished if and only if Δ is, independently of k. Oddly enough, the trickiest part is to prove that when Δ is σ-distinguished, so is $u(\Delta, k)$. We first recall, as a lemma, [Kable 2004, Proposition 1], which is the key ingredient of the proof of the functional equation of the local Asai L-function.

Lemma 2.3. *Let τ be a representation of P_n for $n \geq 1$. Then the space* $\mathrm{Hom}_{P_n^\sigma}(\tau, \mathbf{1})$ *is isomorphic to* $\mathrm{Hom}_{P_{n+1}^\sigma}(\Phi^+(\tau), \mathbf{1})$.

This implies the following generalisation of [Anandavardhanan et al. 2004, Theorem 1.1]:

Proposition 2.4. *Let π be an irreducible unitary representation of G_n for $n \geq 1$. The representation π is P_n^σ-distinguished if and only if its highest shifted derivative $\pi^{[-]}$ is σ-distinguished.*

Proof. One implication follows from Proposition 2.2. For the other one, we first notice that by the definition of Ψ^+, if π' is a representation of G_k for $k \geq 0$, then the space $\mathrm{Hom}_{P_{k+1}^\sigma}(\Psi^+(\pi'), \mathbf{1})$ is isomorphic to $\mathrm{Hom}_{G_k^\sigma}(\nu^{1/2}\pi', \mathbf{1})$. Hence, thanks to Lemma 2.3, the space $\mathrm{Hom}_{P_{k+l}^\sigma}((\Phi^+)^{l-1}\Psi^+(\tau), \mathbf{1})$ is isomorphic to $\mathrm{Hom}_{G_k^\sigma}(\nu^{1/2}\pi', \mathbf{1})$. Now, if τ is an irreducible unitary representation of G_n, let h be the integer such that $\pi^- = \pi^{(h)}$. The restriction of π to P_n has a filtration with factors $(\Phi^+)^{k-1}\Psi^+(\pi^{(k)})$ for k between 1 and h, according to Lemma 1.4. If L is a nonzero P_n^σ-invariant linear form on π, it must induce a nonzero element of $\mathrm{Hom}_{P_n^\sigma}((\Phi^+)^{k-1}\Psi^+\pi^{(k)}, \mathbf{1}) \simeq \mathrm{Hom}_{G_{n-k}^\sigma}(\pi^{[k]}, \mathbf{1})$ for some k in $\{1, \ldots, h\}$. But if the space $\mathrm{Hom}_{G_{n-k}^\sigma}(\pi^{[k]}, \mathbf{1})$ is nonzero, it implies that the central character of one of the irreducible subquotients of $\pi^{[k]}$ has real part equal to zero because F^* must act trivially on at least one irreducible subquotient of $\pi^{[k]}$. Hence, according to Proposition 1.7, this means that the space $\mathrm{Hom}_{P_n^\sigma}((\Phi^+)^{k-1}\Psi^+\pi^{(k)}, \mathbf{1})$ is reduced to zero for k between 1 and $h-1$ and that the space

$$\mathrm{Hom}_{P_n^\sigma}((\Phi^+)^{h-1}\Psi^+\pi^{(h)}, \mathbf{1}) \simeq \mathrm{Hom}_{G_{n-h}^\sigma}(\pi^{[h]}, \mathbf{1})$$

is nonzero. The result is thus proved. $\square$

The proof of the preceding proposition implicitly contains the following statement.

Proposition 2.5. *Let π be an irreducible unitary representation of G_n which is P_n^σ-distinguished. Then its highest shifted derivative $\pi^{[-]}$ is σ-distinguished, and the space* $\mathrm{Hom}_{P_n^\sigma}(\pi, \mathbf{1})$ *is of dimension* 1 *with basis equal to a certain linear form L. Moreover, the restriction of L to $\tau_0 = (\Phi^+)^{h-1}\Psi^+(\pi^-)$ is nonzero, and if τ is any P_n-submodule of π which is P_n^σ-distinguished, then τ contains τ_0 and the space* $\mathrm{Hom}_{P_n^\sigma}(\tau, \mathbf{1})$ *is spanned by the restriction $L_{|\tau}$.*

From this, we deduce a statement which will be used twice in a crucial way.

Proposition 2.6. *Let n_1 and n_2 be two positive integers and π_1 and π_2 be two irreducible unitary representations of G_{n_1} and G_{n_2} respectively. Suppose that π_1 is $G_{n_1}^\sigma$-distinguished and that π_2 is $P_{n_2}^\sigma$-distinguished. In this situation, if $\pi=\pi_1\times\pi_2$ is G_n^σ-distinguished, then π_2 is $G_{n_2}^\sigma$-distinguished.*

Proof. We write $\pi_1\times\pi_2$ as induced from the lower parabolic subgroup $P^-=P^-_{(n_1,n_2)}$ obtained by transposing $P_{(n_1,n_2)}$. It is thus the space $\mathscr{C}_c^\infty(P^-\backslash G_n,\delta_{P^-}^{1/2}\pi_1\otimes\pi_2)$. The double class P^-P_n, being open in G_n, contains

$$\tau=\mathscr{C}_c^\infty(P^-\backslash P^-P_n,\delta_{P^-}^{1/2}\pi_1\otimes\pi_2),$$

which is a P_n-submodule of π. Let L_1 be a basis of $\mathrm{Hom}_{G_{n_1}^\sigma}(\pi_1,\mathbf{1})$, L_2 be a basis of $\mathrm{Hom}_{P_{n_2}^\sigma}(\pi_2,\mathbf{1})$, and denote by λ the linear form $L_1\otimes L_2$ on $\pi_1\otimes\pi_2$. We now introduce the following linear form on τ:

$$L: f\mapsto \int_{P^-\cap P_n^\sigma\backslash P_n^\sigma}\lambda(f(p))\,dp.$$

It is well-defined because the restriction of f to P_n^σ has compact support modulo $P^-\cap P_n^\sigma$ because it satisfies $f(hp)=|a|_F^{-n_2}|b|_F^{n_1}f(p)$ for

$$h=\begin{pmatrix} a&0&0\\ x&b&y\\ 0&0&1\end{pmatrix}\in P^-\cap P_n^\sigma$$

written in blocks according to the partition $(n_1,n_2-1,1)$ of n and because of the relation

$$\frac{\delta_{P^-\cap P_n^\sigma}}{\delta_{P_n^\sigma}}(h)=\frac{|a|_F^{1-n_2}|b|_F^{1+n_1}}{|a|_F|b|_F}=|a|_F^{-n_2}|b|_F^{n_1}.$$

Let's now show that L is nonzero. For v_1 in V_{π_1} and v_2 in V_{π_2}, let U be a congruence subgroup of G_n such that $U\cap G_{n_1}$ fixes v_1 and $U\cap G_{n_2}$ fixes v_2. As U has an Iwahori decomposition with respect to P^-, the map defined by $f_{U,v_1,v_2}(p^-u)=\delta_{P^-}^{1/2}\pi_1\otimes\pi_2(p^-)(v_1\otimes v_2)$ for u in U, p^- in P^- and by zero outside P^-U belongs to V_π. Moreover, $L(f_{U,v_1,v_2})$ is a positive multiple of $L_1(v_1)L_2(v_2)$. In particular, L is nonzero. This implies that L belongs to $\mathrm{Hom}_{P_n^\sigma}(\tau,\mathbf{1})-\{0\}$. It remains to prove that π_2 is $G_{n_2}^\sigma$-distinguished. We are going to prove that L_2 is actually $G_{n_2}^\sigma$-invariant. By Proposition 2.5, as π is irreducible, unitary, and σ-distinguished, we know that $\mathrm{Hom}_{P_n^\sigma}(\pi,\mathbf{1})$ is one-dimensional, spanned by a linear form L'. Moreover, by the same proposition, up to multiplying L' by a scalar, the restriction of L' to τ is equal to L. Hence we denote L' by L. The fact that $\mathrm{Hom}_{P_n^\sigma}(\pi,\mathbf{1})$ is one-dimensional also implies that L is in fact G_n^σ-invariant. Now take h of the form $\mathrm{diag}(I_{n_1},b)$ with b in $G_{n_2}(\mathfrak{O}_K)$. We have $\rho(h)f_{U,v_1,v_2}=f_{U,v_1,\rho(b)v_2}$. Moreover, if b belongs to $G_{n_2}(\mathfrak{O}_K)^\sigma$, the relation $L(\rho(h)f_{U,v_1,v_2})=L(f_{U,v_1,v_2})$ implies the equality $L_1(v_1)L_2(\rho(b)v_2)=L_1(v_1)L_2(v_2)$. This implies that L_2 is $G_{n_2}(\mathfrak{O}_K)^\sigma$-invariant.

In particular, it is w_{n_2}-invariant, where w_{n_2} is the antidiagonal matrix with ones on the second diagonal. As L_2 is $P^\sigma_{n_2}$-invariant by hypothesis, it is $G^\sigma_{n_2}$-invariant because w_{n_2} and $P^\sigma_{n_2}$ span the group $G^\sigma_{n_2}$, and this concludes the proof. □

For Speh representations, we first obtain the following criterion of P^σ_n-distinction.

Proposition 2.7. *Let r be a positive integer, k be an integer ≥ 2, and $n = kr$. Let Δ be a discrete series of G_r. Then the representation $u(\Delta, k)$ is P^σ_n-distinguished if and only if $u(\Delta, k-1)$ is σ-distinguished.*

Proof. We recall from Proposition 1.11 that $u(\Delta, k)^{[-]}$ is equal to $u(\Delta, k-1)$. We then apply Proposition 2.4. □

Proposition 2.4 also has the following corollary.

Corollary 2.8. *Let $n_1, \dots, n_t$ and k be positive integers and Δ_i be a unitary segment of G_{n_i} for each i. If the product $u(\Delta_1, k) \times \cdots \times u(\Delta_t, k)$ is σ-distinguished, then the product $\Delta_1 \times \cdots \times \Delta_t$ is σ-distinguished as well.*

Proof. First, according to Theorem 1.9, the product $u(\Delta_1, k) \times \cdots \times u(\Delta_t, k)$ is unitary. According to Lemma 1.5 and Proposition 1.11, the highest shifted derivative of this product is $u(\Delta_1, k-1) \times \cdots \times u(\Delta_t, k-1)$. It is σ-distinguished according to Proposition 2.4. Hence, by induction, the product $\Delta_1 \times \cdots \times \Delta_t$ is σ-distinguished as well. □

In particular, if $u(\Delta, k)$ is σ-distinguished, then Δ is σ-distinguished. We are now able to prove the main result of this section.

Corollary 2.9. *Let k and m be two positive integers and Δ be a discrete series of G_m. The representation $u(\Delta, k)$ is σ-distinguished if and only if Δ is σ-distinguished.*

Proof. If $u(\Delta, k)$ is σ-distinguished, we already noticed that Δ is σ-distinguished as a consequence of Corollary 2.8. For the converse, we do an induction on k.

The case $k = 1$ is clear, so let's suppose that $u(\Delta, l)$ is σ-distinguished for $l \leq k$ with $k \geq 1$. We recall from Theorem 1.10 that $\nu^{1/2}u(\Delta, k) \times \nu^{-1/2}u(\Delta, k)$ is of length two and has $u(\Delta_-, k) \times u(\Delta_+, k)$ and $u(\Delta, k-1) \times u(\Delta, k+1)$ as irreducible subquotients. Now, as $u(\Delta, k)^\vee = u(\Delta, k)^\sigma$, according to the main theorem of [Blanc and Delorme 2008], the representation $\nu^{1/2}u(\Delta, k) \times \nu^{-1/2}u(\Delta, k)$ is σ-distinguished. But $u(\Delta_-, k) \times u(\Delta_+, k)$ can't be distinguished, otherwise $\Delta_- \times \Delta_+$ would be distinguished thanks to Corollary 2.8, and this would in turn imply that both Δ_- and Δ_+ are also distinguished according to Theorem 1.16, which contradicts Corollary 1.18. Hence, the representation $u(\Delta, k-1) \times u(\Delta, k+1)$ must be σ-distinguished. We recall that the representation $u(\Delta, k-1)$ is σ-distinguished by the induction hypothesis. As $u(\Delta, k)$ is σ-distinguished by hypothesis as well, the representation $u(\Delta, k+1)$ is $P^\sigma_{(k+1)m}$-distinguished by Proposition 2.7. Then, the representation $u(\Delta, k+1)$ is σ-distinguished according to Proposition 2.6, and this provides the induction step. $\square$

As a corollary, we obtain the following result.

Corollary 2.10. *Let k and m be positive integers. If Δ is a segment of G_m and $u(\Delta, k)^\vee$ is isomorphic to $u(\Delta, k)^\sigma$, then $u(\Delta, k)$ is either σ-distinguished or (σ, η)-distinguished and not both at the same time.*

Proof. The representation $u(\Delta, k)^\vee$ is isomorphic to $u(\Delta, k)^\sigma$ if and only if $\Delta^\vee$ is isomorphic to Δ^σ. The result is then a consequence of [Kable 2004, Theorem 7] and of [Anandavardhanan et al. 2004, Corollary 1.6]. $\square$

The general case. First, we notice that the class of σ-induced unitary irreducible representations of G_n is contained in the class of σ-distinguished representations.

Proposition 2.11. *For $n \geq 1$, let π be an irreducible unitary representation of G_n which is σ-induced. Then it is σ-distinguished.*

Proof. Let Δ be a discrete series of G_m with $m \geq 1$, let k be a positive integer, and let α be a real number. Then the representations $\nu_K^\alpha u(\Delta, k) \times ((\nu_K^\alpha u(\Delta, k))^\vee)^\sigma$ and $\pi(u(\Delta, k), \alpha) \times (\pi(u(\Delta, k), \alpha)^\vee)^\sigma$ are σ-distinguished according to the main theorem of [Blanc and Delorme 2008]. But as a product of σ-distinguished representations is σ-distinguished according to [Flicker 1992, Proposition 26], it follows from Remark 1.15 that if π is σ-induced, then it is indeed σ-distinguished. $\square$

It remains to prove the converse to obtain the main result of this paper. First, we make the following obvious but useful observation.

Lemma 2.12. *Let $\pi = u(\Delta_1, k_1) \times \cdots \times u(\Delta_r, k_r) \times \pi(u(\Delta_{r+1}, k_{r+1}), \alpha_{r+1}) \times \cdots \times \pi(u(\Delta_t, k_t), \alpha_t)$ be an irreducible unitary representation of G_n with Δ_i discrete series and real numbers α_i in $(0, 1/2)$. If the integers k_i satisfy $k_i \geq 2$, then π is σ-induced if and only if its highest shifted derivative $\pi^{[-]}$ is σ-induced.*

Proof. With the notations of the statement, according to Lemma 1.5 and Proposition 1.11, the representation $\pi^{[-]}$ is equal to the product

$$u(\Delta_1, k_1 - 1) \times \cdots \times u(\Delta_r, k_r - 1) \\ \times \pi(u(\Delta_{r+1}, k_{r+1} - 1), \alpha_{r+1}) \times \cdots \times \pi(u(\Delta_t, k_t - 1), \alpha_t).$$

Now it is clear that π is σ-self-dual if and only if $\pi^{[-]}$ is and that a representation $u(\Delta, k)$ (with Δ unitary) occurs with odd multiplicity in π if and only if $u(\Delta, k-1)$ occurs with odd multiplicity in $\pi^{[-]}$. The result now follows from the fact that a Speh representation $u(\Delta, k)$ with $k \geq 2$ is σ-distinguished if and only if $u(\Delta, k-1)$ is σ-distinguished, thanks to Corollary 2.9. $\square$

Theorem 2.13. *If π is an irreducible unitary representation of G_n for $n \geq 1$, then π is σ-distinguished if and only it is σ-induced.*

Proof. One direction is Proposition 2.11. Hence, it remains to show that when π is σ-distinguished, it is σ-induced. To do this, we first write π under the form $\pi_1 \times \pi_2$, where π_1 is an irreducible unitary representation of G_{n_1} for some $n_1 \geq 0$ which is a product of the form described in the statement of Lemma 2.12 (i.e., the k_i are ≥ 2) and π_2 is generic unitary of G_{n_2} for $n_2 \geq 0$ (i.e., if you write it as a standard product in Tadić's classification, all the k_i are equal to 1). Notice that π_1 and π_2, and hence n_1 and n_2, are uniquely determined by π. We now prove the statement by induction on n_1.

The case $n_1 = 0$ is true thanks to Theorem 1.16. We thus suppose that n_1 is positive, in which case it is necessarily at least 2 by definition of the representation π_1 (the integers k_i occurring in its definition being at least 2), and we suppose that the statement to prove is true for any irreducible unitary representation $\pi' = \pi'_1 \times \pi'_2$ with $n'_1 < n_1$. By hypothesis, the representation π is σ-distinguished, and hence the representation $\pi^{[-]} = \pi_1^{[-]}$ is σ-distinguished as well thanks to Proposition 2.4. Then, by induction hypothesis, the representation $\pi_1^{[-]}$ must be σ-induced (because if one writes $\pi' = \pi_1^{[-]}$ under the form $\pi'_1 \times \pi'_2$, then we have $n'_1 < n_1$). This implies that the representation π_1 is σ-induced as well according to Lemma 2.12. In particular, it is σ-distinguished by Proposition 2.11. Then, we notice that the representation π_2 is $P^{\sigma}_{n_2}$-distinguished according to Proposition 2.4, as $\pi_2^{[-]}$ is the trivial character of G_0. We can now apply Proposition 2.6 and conclude that π_2 is σ-distinguished, thus σ-induced thanks to Theorem 1.16. This finally implies that π is σ-induced as well. □

Acknowledgments

I thank U.K. Anandavardhanan and A. Minguez for suggesting to study distinction for Speh representations, as well as helpful conversations. I also thank I. Badulescu for answering some questions about these representations. I thank S. Sugiyama for pointing out many typos in the paper. Finally, thanks to the referee's accurate reading and helpful comments, the general presentation of the paper improved significantly, and some proofs were clarified.

References

[Anandavardhanan et al. 2004] U. K. Anandavardhanan, A. C. Kable, and R. Tandon, "Distinguished representations and poles of twisted tensor L-functions", *Proc. Amer. Math. Soc.* **132**:10 (2004), 2875–2883. MR 2005g:11080 Zbl 1122.11033

[Badulescu 2011] A. I. Badulescu, "On p-adic Speh representations", preprint, 2011. arXiv 1110.5080

[Bernstein 1984] J. N. Bernstein, "P-invariant distributions on GL(N) and the classification of unitary representations of GL(N) (non-Archimedean case)", pp. 50–102 in *Lie group representations, II*

(College Park, MD, 1982–1983), edited by R. Herb et al., Lecture Notes in Math. **1041**, Springer, Berlin, 1984. MR 86b:22028 Zbl 0541.22009

[Bernstein and Zelevinsky 1977] J. N. Bernstein and A. V. Zelevinsky, "Induced representations of reductive $\mathfrak{p}$-adic groups, I", *Ann. Sci. École Norm. Sup.* (4) **10**:4 (1977), 441–472. MR 58 #28310 Zbl 0412.22015

[Blanc and Delorme 2008] P. Blanc and P. Delorme, "Vecteurs distributions H-invariants de représentations induites, pour un espace symétrique réductif p-adique G/H", *Ann. Inst. Fourier* (*Grenoble*) **58**:1 (2008), 213–261. MR 2009e:22015 Zbl 1151.22012

[Flicker 1991] Y. Z. Flicker, "On distinguished representations", *J. Reine Angew. Math.* **418** (1991), 139–172. MR 92i:22019 Zbl 0725.11026

[Flicker 1992] Y. Z. Flicker, "Distinguished representations and a Fourier summation formula", *Bull. Soc. Math. France* **120**:4 (1992), 413–465. MR 93j:22033 Zbl 0778.11030

[Kable 2004] A. C. Kable, "Asai L-functions and Jacquet's conjecture", *Amer. J. Math.* **126**:4 (2004), 789–820. MR 2005g:11083 Zbl 1061.11023

[Matringe 2009] N. Matringe, "Conjectures about distinction and local Asai L-functions", *Int. Math. Res. Not.* **2009**:9 (2009), 1699–1741. MR 2011a:22020 Zbl 1225.22014

[Matringe 2011a] N. Matringe, "Derivatives and asymptotics of Whittaker functions", *Represent. Theory* **15** (2011), 646–669. MR 2833471 Zbl 1242.22024

[Matringe 2011b] N. Matringe, "Distinguished generic representations of $\mathrm{GL}(n)$ over p-adic fields", *Int. Math. Res. Not.* **2011**:1 (2011), 74–95. MR 2012f:22032 Zbl 1223.22015

[Offen and Sayag 2008] O. Offen and E. Sayag, "Global mixed periods and local Klyachko models for the general linear group", *Int. Math. Res. Not.* **2008**:1 (2008), Art. ID rnm 136. MR 2009e:22017 Zbl 1158.22021

[Rodier 1982] F. Rodier, "Représentations de $\mathrm{GL}(n, k)$ où k est un corps p-adique", pp. 201–218, Exp. No. 587 in *Bourbaki Seminar, Vol. 1981/1982*, Astérisque **92**, Soc. Math. France, Paris, 1982. MR 84h:22040 Zbl 0506.22019

[Tadić 1986] M. Tadić, "Classification of unitary representations in irreducible representations of general linear group (non-Archimedean case)", *Ann. Sci. École Norm. Sup.* (4) **19**:3 (1986), 335–382. MR 88b:22021 Zbl 0614.22005

[Tadić 1987a] M. Tadić, "Topology of unitary dual of non-Archimedean $\mathrm{GL}(n)$", *Duke Math. J.* **55**:2 (1987), 385–422. MR 89c:22029 Zbl 0668.22006

[Tadić 1987b] M. Tadić, "Unitary representations of $\mathrm{GL}(n)$, derivatives in the non-Archimedean case", pp. 274–282, Ber. No. 281 in *V. Mathematikertreffen Zagreb-Graz* (Mariatrost/Graz, 1986), Ber. Math.-Statist. Sekt. Forschungsgesellsch. Joanneum **274**, Forschungszentrum Graz, Graz, 1987. MR 89c:22028 Zbl 0631.22015

[Zelevinsky 1980] A. V. Zelevinsky, "Induced representations of reductive $\mathfrak{p}$-adic groups, II: On irreducible representations of $\mathrm{GL}(n)$", *Ann. Sci. École Norm. Sup.* (4) **13**:2 (1980), 165–210. MR 83g:22012 Zbl 0441.22014

Received August 6, 2013. Revised November 25, 2013.

NADIR MATRINGE
UNIVERSITÉ DE POITIERS
86000 POITIERS
FRANCE
matringe@math.univ-poitiers.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.461

ON f-BIHARMONIC MAPS AND f-BIHARMONIC SUBMANIFOLDS

YE-LIN OU

We consider f-biharmonic maps, the extrema of the f-bienergy functional. We prove that an f-biharmonic map from a compact Riemannian manifold into a nonpositively curved manifold with constant f-bienergy density is a harmonic map; that any f-biharmonic function on a compact manifold is constant; and that the inversion in the sphere S^{m-1} is a proper f-biharmonic conformal diffeomorphism for $m \geq 3$. We derive equations for f-biharmonic submanifolds (that is, submanifolds whose defining isometric immersions are f-biharmonic maps) and prove that a surface in a manifold (N^n, h) is an f-biharmonic surface if and only if it can be biharmonically conformally immersed into (N^n, h). We also give a complete classification of f-biharmonic curves in three-dimensional Euclidean space. Examples are given of proper f-biharmonic maps and f-biharmonic surfaces and curves.

1. Harmonic, biharmonic, f-harmonic, and f-biharmonic maps

All objects in this paper, including manifolds, tensor fields, and maps, are assumed smooth unless stated otherwise.

We recall the key definitions, focusing on maps on compact Riemannian manifolds M. (For noncompact M, the relevant functionals are integrals over fixed compact domains $K \subset M$, and the criticality conditions must hold for all K.)

Harmonic maps. Harmonic maps are critical points of the energy functional for maps $\phi : (M, g) \to (N, h)$ between Riemannian manifolds:

$$E(\phi) = \frac{1}{2} \int_M |d\phi|^2 v_g.$$

The Euler–Lagrange equation gives the harmonic map equation [Eells and Sampson 1964]

$$\tau(\phi) := \mathrm{Trace}_g \nabla d\phi = 0,$$

Research supported by NSF of Guangxi (P. R. China), 2011GXNSFA018127.
MSC2010: primary 58E20; secondary 53C43.
Keywords: f-biharmonic maps, f-biharmonic submanifolds, f-biharmonic functions, f-biharmonic hypersurfaces, f-biharmonic curves.

where $\tau(\phi) = \mathrm{Trace}_g \nabla d\phi$ is called the tension field of the map ϕ.

Biharmonic maps. Biharmonic maps are critical points of the bienergy functional for maps $\phi : (M, g) \to (N, h)$ between Riemannian manifolds:

$$E_2(\phi) = \frac{1}{2} \int_M |\tau(\phi)|^2 v_g.$$

The Euler–Lagrange equation of this functional gives the biharmonic map equation [Jiang 1986b], namely the vanishing of the *bitension field* $\tau_2(\phi)$ *of* ϕ:

$$\tau_2(\phi) := \mathrm{Trace}_g(\nabla^\phi \nabla^\phi - \nabla^\phi_{\nabla^M})\tau(\phi) - \mathrm{Trace}_g R^N(d\phi, \tau(\phi))d\phi = 0.$$

Here R^N is the curvature operator of (N, h), defined by

$$R^N(X, Y)Z = [\nabla^N_X, \nabla^N_Y]Z - \nabla^N_{[X,Y]}Z.$$

f-harmonic maps. f-harmonic maps are critical points of the f-energy functional for maps $\phi : (M, g) \to (N, h)$ between Riemannian manifolds:

$$E_f(\phi) = \frac{1}{2} \int_M f |d\phi|^2 v_g.$$

Here f is a fixed function $M \to (0, \infty)$. The Euler–Lagrange equation gives the f-harmonic map equation [Course 2004; Ouakkas et al. 2010]

$$\tau_f(\phi) := f\tau(\phi) + d\phi(\mathrm{grad}\, f) = 0.$$

We call $\tau_f(\phi)$ the f-tension field of the map ϕ.

f-biharmonic maps. f-biharmonic maps are critical points of the f-bienergy functional for maps $\phi : (M, g) \to (N, h)$ between Riemannian manifolds:

$$E_{2,f}(\phi) = \frac{1}{2} \int_M f |\tau(\phi)|^2 v_g.$$

The Euler–Lagrange equation gives the f-biharmonic map equation [Lu 2013]

$$\tau_{2,f}(\phi) := f\tau_2(\phi) + (\Delta f)\tau(\phi) + 2\nabla^\phi_{\mathrm{grad}\, f}\tau(\phi) = 0.$$

Bi-f-harmonic maps. Bi-f-harmonic maps are critical points of the bi-f-energy functional for maps $\phi : (M, g) \to (N, h)$ between Riemannian manifolds:

$$E^2_f(\phi) = \frac{1}{2} \int_M |\tau_f(\phi)|^2 v_g. \tag{1}$$

The Euler–Lagrange equation gives the bi-f-harmonic map equation [Ouakkas et al. 2010]

$$\tau^2_f(\phi) := f J^\phi(\tau_f(\phi)) - \nabla^\phi_{\mathrm{grad}\, f}\tau_f(\phi) = 0,$$

where J^ϕ is the Jacobi operator of the map, defined by

$$J^\phi(X) = -(\mathrm{Trace}_g \nabla^\phi \nabla^\phi X - \nabla^\phi_{\nabla^M} X - R^N(d\phi, X)d\phi).$$

Remark. Ouakkas et al. [2010] used the name "f-biharmonic maps" for the critical points of the functional (1). We think that it is more reasonable to call them "bi-f-harmonic maps" as parallel to "biharmonic maps".

We have the following obvious inclusions among the various types of harmonic maps:

$$\{\text{harmonic}\} \subset \{\text{biharmonic}\} \subset \{f\text{-biharmonic}\},$$

$$\{\text{harmonic}\} \subset \{f\text{-harmonic}\} \subset \{\text{bi-}f\text{-harmonic}\}.$$

From now on we will call an f-biharmonic map which is neither harmonic nor biharmonic a *proper f-biharmonic map*.

Harmonic maps as a generalization of important concepts of geodesics, minimal surfaces, and harmonic functions have been studied extensively with tremendous progress in the past 40-plus years. There is voluminous literature about the beautiful theory, important applications, and interesting links of harmonic maps to other areas of mathematics and theoretical physics including nonlinear partial differential equations, holomorphic maps in several complex variables, the theory of stochastic processes, liquid crystals in materials science, and the nonlinear field theory.

The study of biharmonic maps was proposed in [Eells and Lemaire 1983] and Jiang [1986a; 1986b; 1987] made the first serious study on these maps by using the first and second variational formulas of the bienergy functional and specializing on the biharmonic isometric immersions which nowadays are called biharmonic submanifolds. Very interestingly, the concept of biharmonic submanifolds was also introduced in a different way by B. Y. Chen [1991] in his program of understanding the finite-type submanifolds in Euclidean spaces. Since 2000, biharmonic maps have been receiving a growing attention and have become a popular subject of study with great progress. For some recent geometric study of general biharmonic maps see [Baird and Kamissoko 2003; Montaldo and Oniciuc 2006; Ou 2006; 2012b; Balmuş et al. 2007; Ouakkas 2008; Baird et al. 2010; Ou and Lu 2013; Nakauchi et al. 2014; Wang et al. 2014] and the references therein. For some recent study of biharmonic submanifolds see [Jiang 1986a; 1987; Dimitrić 1992; Chen and Ishikawa 1998; Caddeo et al. 2001; 2002; Balmuş et al. 2008; 2013; Ou 2010; Ou and Wang 2011; Ou and Tang 2012; Alías et al. 2013; Chen and Munteanu 2013; Liang and Ou 2013; Nakauchi and Urakawa 2013] and the references therein. For biharmonic conformal immersions and submersions see [Baird et al. 2008; Ou 2009; 2012a; Loubeau and Ou 2010; Wang and Ou 2011] and the references therein.

Lu [2013] introduced f-biharmonic maps and calculated the first variation to obtain the f-biharmonic map equation and the equation for the f-biharmonic conformal maps between the same dimensional manifolds. In this paper, we study some basic properties of f-biharmonic maps and introduce the concept of f-biharmonic submanifolds. We prove that an f-biharmonic map from a compact Riemannian manifold into a nonpositively curved manifold with constant f-bienergy density is a harmonic map (Theorem 2.4); any f-biharmonic function on a compact manifold is constant (Corollary 2.6); and that the inversion in sphere S^{m-1} is a proper f-biharmonic conformal diffeomorphism for $m \geq 3$ (Proposition 2.9). We derive f-biharmonic submanifolds equations (Theorem 3.2 and Corollary 3.4) and prove that a surface in a manifold (N^n, h) is an f-biharmonic surface if and only if it can be biharmonically conformally immersed into (N^n, h) (Corollary 3.6). We also give a complete classification of f-biharmonic curves in three-dimensional Euclidean spaces (Theorem 4.4) according to which proper f-biharmonic curves are some special subclasses of planar curves or general helices in $\mathbb{R}^3$. Many examples of proper f-biharmonic maps and f-biharmonic surfaces and curves are given.

2. Some properties and examples of f-biharmonic maps

As mentioned, f-biharmonic maps are critical points of the f-bienergy functional for maps $\phi : (M, g) \to (N, h)$ between Riemannian manifolds:

$$E_{2,f}(\phi) = \frac{1}{2}\int_M f|\tau(\phi)|^2 v_g.$$

The following theorem was proved in [Lu 2013]. We give a brief outline of the proof for completeness, but note that our notation is different from Lu's.

Theorem 2.1. *A map $\phi : (M, g) \to (N, h)$ between Riemannian manifolds is an f-biharmonic map if and only if*

$$\tau_{2,f}(\phi) := f\tau_2(\phi) + (\Delta f)\tau(\phi) + 2\nabla^{\phi}_{\operatorname{grad} f}\tau(\phi) = 0, \tag{2}$$

where $\tau(\phi)$ and $\tau_2(\phi)$ are the tension and the bitension fields of ϕ respectively.

Proof. Since f is fixed, we can use the standard method (see, e.g., [Baird and Kamissoko 2003; Jiang 1986b]) of calculating the first variation of the bienergy functional to obtain

$$\begin{aligned}
\frac{\partial}{\partial t}E_{2,f}(\phi_t)\Big|_{t=0} &= \frac{1}{2}\int_M f\left\{\frac{\partial}{\partial t}\langle\tau(\phi_t), \tau(\phi_t)\rangle\right\}_{t=0} v_g \\
&= -\int_M f\langle\tau(\phi), J^{\phi}(V)\rangle v_g \\
&= \int_M \langle f\tau(\phi), \operatorname{Trace}_g\nabla^{\phi}\nabla^{\phi}V - \nabla^{\phi}_{\nabla^M}V - R^N(d\phi, V)d\phi\rangle v_g.
\end{aligned}$$

Using the symmetry property of the curvature tensor and the divergence theorem we can switch the positions of V and $f\tau(\phi)$ to have

$$\frac{\partial}{\partial t}E_{2,f}(\phi_t)\Big|_{t=0} = -\int_M \langle V, J^{\phi}(f\tau(\phi))\rangle v_g.$$

It follows that ϕ is an f-biharmonic map if and only if the f-bitension field vanishes identically, i.e., $\tau_{2,f}(\phi) = -J^{\phi}(f\tau(\phi)) \equiv 0$. Finally, using [Ou 2006, (7)], we have

$$\begin{aligned}\tau_{2,f}(\phi) = -J^{\phi}(f\tau(\phi)) &= -\{fJ^{\phi}(\tau(\phi)) - (\Delta f)\tau(\phi) - 2\nabla^{\phi}_{\operatorname{grad} f}\tau(\phi)\}\\ &= f\tau_2(\phi) + (\Delta f)\tau(\phi) + 2\nabla^{\phi}_{\operatorname{grad} f}\tau(\phi),\end{aligned}$$

from which the f-biharmonic map equation (2) follows. □

It is well known that for $m \neq 2$, the harmonicity and f-harmonicity of a map $\phi : (M^m, g) \to (N^n, h)$ are related via a conformal change of the domain metric. More precisely:

Proposition 2.2 [Lichnerowicz 1969]. *A map $\phi : (M^m, g) \to (N^n, h)$ with $m \neq 2$ is f-harmonic if and only if $\phi : (M^m, f^{\frac{2}{m-2}} g) \to (N^n, h)$ is a harmonic map.*

In general, this does not generalize to the case of the relationship between biharmonicity and f-biharmonicity, but very interestingly, we have:

Theorem 2.3. *A map $\phi : (M^2, g) \to (N^n, h)$ is an f-biharmonic map if and only if $\phi : (M^2, f^{-1}g) \to (N^n, h)$ is a biharmonic map.*

Proof. On the one hand, we notice that the map $\phi : (M^2, g) \to (N^n, h)$ is an f-biharmonic map if and only if

$$f\tau_2(\phi, g) + (\Delta f)\tau(\phi, g) + 2\nabla^{\phi}_{\operatorname{grad} f}\tau(\phi, g) = 0, \tag{3}$$

which is equivalent to

$$\tau_2(\phi, g) + (\Delta \ln f + |\operatorname{grad} \ln f|^2)\tau(\phi, g) + 2\nabla^{\phi}_{\operatorname{grad} \ln f}\tau(\phi, g) = 0. \tag{4}$$

On the other hand, by [Ou 2009, Corollary 1], the relationship between the bitension field $\tau_2(\phi, g)$ and that of the map $\phi : (M^2, \bar g = F^{-2}g) \to (N^n, h)$ is given by

$$\tau_2(\phi, \bar g) = F^4\big\{\tau^2(\phi, g) + 2(\Delta \ln F + 2\,|\operatorname{grad} \ln F|^2)\tau(\phi, g) + 4\nabla^{\phi}_{\operatorname{grad} \ln F}\,\tau(\phi, g)\big\},$$

which is equivalent to

$$\tau_2(\phi, \bar g) = F^4\big\{\tau^2(\phi, g) + (\Delta \ln F^2 + |\operatorname{grad} \ln F^2|^2)\tau(\phi, g) + 2\nabla^{\phi}_{\operatorname{grad} \ln F^2}\,\tau(\phi, g)\big\}.$$

It follows that the map $\phi : (M^2, \bar g = F^{-2}g) \to (N^n, h)$ is biharmonic if and only if

$$\tau_2(\phi, g) + (\Delta \ln F^2 + |\operatorname{grad} \ln F^2|^2)\tau(\phi, g) + 2\nabla^{\phi}_{\operatorname{grad} \ln F^2}\tau(\phi, g) = 0. \tag{5}$$

Substituting $F^2 = f$ into (5) yields (4). Hence the map $\phi : (M^2, g) \to (N^n, h)$ is f-biharmonic if and only if $\phi : (M^m, f^{-1}g) \to (N^n, h)$ is biharmonic. □

Theorem 2.4. *Any f-biharmonic map $\phi : (M^m, g) \to (N^n, h)$ from a compact Riemannian manifold into a nonpositively curved manifold with constant f-bienergy density* (*i.e.*, $f|\tau(\phi)|^2 = C$) *is a harmonic map.*

Proof. A straightforward computation gives

$$
\begin{aligned}
(6)\quad & \Delta\big(\tfrac{1}{2} f|\tau(\phi)|^2\big)\\
&= \tfrac{1}{2}\Delta\langle f^{\frac{1}{2}}\tau(\phi), f^{\frac{1}{2}}\tau(\phi)\rangle\\
&= (\nabla^\phi_{e_i}\nabla^\phi_{e_i} - \nabla^\phi_{\nabla^M_{e_i}e_i})\langle f^{\frac{1}{2}}\tau(\phi), f^{\frac{1}{2}}\tau(\phi)\rangle\\
&= \langle \nabla^\phi_{e_i} f^{\frac{1}{2}}\tau(\phi), \nabla^\phi_{e_i} f^{\frac{1}{2}}\tau(\phi)\rangle + \langle(\nabla^\phi_{e_i}\nabla^\phi_{e_i} - \nabla^\phi_{\nabla^M_{e_i}e_i}) f^{\frac{1}{2}}\tau(\phi), f^{\frac{1}{2}}\tau(\phi)\rangle\\
&= \langle \nabla^\phi_{e_i} f^{\frac{1}{2}}\tau(\phi), \nabla^\phi_{e_i} f^{\frac{1}{2}}\tau(\phi)\rangle + f\langle(\nabla^\phi_{e_i}\nabla^\phi_{e_i} - \nabla^\phi_{\nabla^M_{e_i}e_i})\tau(\phi), \tau(\phi)\rangle\\
&\qquad + f^{\frac{1}{2}}(\Delta f^{\frac{1}{2}})|\tau(\phi)|^2 + 2f^{\frac{1}{2}}\langle\nabla^\phi_{\mathrm{grad}\, f^{\frac{1}{2}}}\tau(\phi), \tau(\phi)\rangle.
\end{aligned}
$$

Since ϕ is assumed to be f-biharmonic we have

$$
\begin{aligned}
(7)\quad & f\langle(\nabla^\phi_{e_i}\nabla^\phi_{e_i} - \nabla^\phi_{\nabla^M_{e_i}e_i})\tau(\phi), \tau(\phi)\rangle\\
&= \langle f R^N(d\phi(e_i), \tau(\phi))d\phi(e_i) - (\Delta f)\tau(\phi) - 2\nabla^\phi_{\mathrm{grad}\, f}\tau(\phi), \tau(\phi)\rangle\\
&= f\langle R^N(d\phi(e_i), \tau(\phi))d\phi(e_i), \tau(\phi)\rangle - (\Delta f)|\tau(\phi)|^2 - 2\langle\nabla^\phi_{\mathrm{grad}\, f}\tau(\phi), \tau(\phi)\rangle.
\end{aligned}
$$

Substituting (7) into (6) and simplifying the result gives

$$\Delta\big(\tfrac{1}{2} f|\tau(\phi)|^2\big) = f|\nabla^\phi_{e_i}\tau(\phi)|^2 - fR^N(d\phi(e_i), \tau(\phi), d\phi(e_i), \tau(\phi)) - \tfrac{1}{2}(\Delta f)|\tau(\phi)|^2.$$

This, together with the assumptions that $f|\tau(\phi)|^2 = C$, $f > 0$, and

$$(8)\qquad R^N(d\phi(e_i), \tau(\phi), d\phi(e_i), \tau(\phi)) \le 0,$$

allows us to conclude that f is a subharmonic function on the compact manifold (M, g) and hence f is a constant function. It follows that the f-biharmonic map ϕ is actually a biharmonic map from a compact manifold into a nonpositively curved manifold, and thus a harmonic map by a theorem in [Jiang 1986b]. □

Remark. There are many harmonic maps between spheres with constant energy density (called eigenmaps). As our Theorem 2.4 implies that there is no proper f-biharmonic maps from a compact manifold into a nonpositively curved manifold with constant f-bienergy density, it would be interesting to know if there is any proper f-biharmonic map between spheres with constant f-bienergy density.

Proposition 2.5. *A function $u : (M, g) \to \mathbb{R}$ is f-biharmonic if and only if*

$$(9)\qquad f\Delta^2 u + (\Delta f)\Delta u + 2g(\mathrm{grad}\, f, \mathrm{grad}\, \Delta u) = 0, \quad \textit{or, equivalently,}$$

$$(10)\qquad \Delta(f\Delta u) = 0,$$

where $\Delta^2 u = \Delta(\Delta u)$ denotes the bi-Laplacian of u. In other words, a function u is an f-biharmonic function if and only if the product $f\Delta u$ is a harmonic function. In particular, a quasiharmonic function u (i.e., a function $u : (M, g) \to \mathbb{R}$ with $\Delta u = \text{constant} \neq 0$) is an f-biharmonic function if and only if $f : (M, g) \to \mathbb{R}$ is a harmonic function.

Proof. A straightforward computation gives the tension and the bitension fields of $u : (M, g) \to \mathbb{R}$ as

$$\tau(u) = (\Delta u)\frac{\partial}{\partial t} \quad \text{and} \quad \tau_2(u) = (\Delta^2 u)\frac{\partial}{\partial t}. \tag{11}$$

Substituting these into f-biharmonic map equation (2) and performing a further computation we obtain the f-biharmonic function equation (9). The last statement thus follows. □

Corollary 2.6. *Any f-biharmonic function on a compact manifold (M, g) is a constant function.*

Proof. By Proposition 2.5, u is an f-biharmonic function if and only if $f\Delta u$ is a harmonic function. By the well-known fact that any harmonic function on a compact manifold is constant we have $f\Delta u = C$, and hence

$$\Delta u = \frac{C}{f} \tag{12}$$

since $f > 0$ by our assumption. If $C = 0$, then we have $\Delta u = 0$ and hence u is a harmonic function, so u is a constant function in this case. If $C \neq 0$, then (12) implies that u is either a subharmonic or a superharmonic function since f has a fixed sign with $f > 0$. Again, the well-known fact that a subharmonic or superharmonic function on a compact manifold is constant implies that u is constant. This completes the proof of the corollary. □

Example 1. Let $f : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$ be the function $f(x, y, z) = 1/\sqrt{x^2 + y^2 + z^2}$ and let $u : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$ be the function given by $u(x, y, z) = x^2 + y^2 + z^2$. It is easily checked that $\Delta f = 0$, $\Delta u = 6$ and $\Delta^2 u = 0$ and hence f and u satisfy (9). So, $u(x, y, z)$ is an f-biharmonic function on $\mathbb{R}^3 \setminus \{0\}$ for $f(x, y, z)$. Clearly, this f-biharmonic function u is not a harmonic function.

Example 2. Let $f, u : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$ be the functions defined by $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ and $u(x, y, z) = x/(x^2 + y^2 + z^2)$. Then we can check (see also Proposition 2.9) that u is a proper f-biharmonic function which is neither harmonic nor biharmonic.

Corollary 2.7. *Let $f, u : \mathbb{R} \to \mathbb{R}$ be two functions with $f(x) > 0$ for all $x \in \mathbb{R}$. Then u is an f-biharmonic function if and only if*

$$u(x) = \iint \frac{Ax + B}{f}\, dx\, dx + Cx + D, \tag{13}$$

where A, B, C, D are arbitrary constants. In particular:

(I) *For $f(x) = 1 + x^2$, a function $u : \mathbb{R} \to \mathbb{R}$ is f-biharmonic if and only if $u(x) = \frac{1}{2}(Ax - B)\ln(1 + x^2) + (Bx + A)\arctan x + (C - A)x + D$, where A, B, C, D are constants.*

(II) *For $f(x) = e^{-x}$, a function $u : \mathbb{R} \to \mathbb{R}$ is f-biharmonic if and only if $u(x) = (Ax - 2A + B)e^x + Cx + D$, where A, B, C, D are constants.*

Proof. In this case, the f-biharmonic equation (10) reduces to $(fu'')'' = 0$ which has solution (13). Finally, statements (I) and (II) are obtained by elementary integrations. □

Remark. It is easily checked that for $A \neq 0$, $B \neq 0$ the function

$$u(x) = (Ax - 2A + B)e^x + Cx + D$$

is neither a harmonic nor a biharmonic function, so it provides many examples of proper f-biharmonic functions.

Theorem 2.8. *Any f-biharmonic map $\phi : (M^m, g) \to \mathbb{R}^n$ from a compact manifold into a Euclidean space is a constant map.*

Proof. Since the target manifold is a Euclidean space, the curvature is zero. If we write $\phi : (M^m, g) \to \mathbb{R}^n$ as $\phi(p) = (\phi^1(p), \phi^2(p), \ldots, \phi^n(p))$, then we can easily check that

$$\begin{aligned}
\tau(\phi) &= (\Delta\phi^1, \Delta\phi^2, \ldots, \Delta\phi^n),\\
\tau_2(\phi) &= (\Delta^2\phi^1, \Delta^2\phi^2, \ldots, \Delta^2\phi^n),\\
\nabla^{\phi}_{\mathrm{grad}\, f}\tau(\phi) &= (\nabla^{\phi}_{\mathrm{grad}\, f}\Delta\phi^1, \nabla^{\phi}_{\mathrm{grad}\, f}\Delta\phi^2, \ldots, \nabla^{\phi}_{\mathrm{grad}\, f}\Delta\phi^n).
\end{aligned}$$

It follows that the f-biharmonic map equation for ϕ becomes

$$f\Delta^2\phi^\alpha + (\Delta f)\Delta\phi^\alpha + 2g(\mathrm{grad}\, f, \mathrm{grad}\, \Delta\phi^\alpha) = 0, \quad \alpha = 1, 2, \ldots, n.$$

In other words, a map $\phi : (M^m, g) \to \mathbb{R}^n$ from a manifold into a Euclidean space is an f-biharmonic map if and only if each of its component functions is an f-biharmonic function. From this and Corollary 2.6, which states that any f-biharmonic function on a compact manifold is constant, we obtain the theorem. □

Proposition 2.9. *The map $\phi : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}^m \setminus \{0\}$ with $\phi(x) = x/|x|^p$ is an f-biharmonic map for $f(x) = |x|^k$ if and only if (i) $p = 0$, or (ii) $p = m$, or (iii) $k = p + 2$, or (iv) $k = p + 2 - m$. In particular, for $m \geq 3$, the inversion in sphere S^{m-1}, $\phi : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}^m \setminus \{0\}$ with $\phi(x) = x/|x|^2$ is a proper f-biharmonic map for $f(x) = |x|^4$. When $m \neq 4$, this inversion is also a proper f-biharmonic map for $f(x) = |x|^{4-m}$.*

Proof. As we have seen in the proof of Theorem 2.8, a map into a Euclidean space is an f-biharmonic map if and only if each of its component functions is an f-biharmonic function. So, $\phi : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}^m \setminus \{0\}$ with $\phi(x) = x/|x|^p$ is f-biharmonic if and only if the function $u : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}$ with $u(x) = x^i |x|^{-p}$ is an f-biharmonic function for any $i = 1, 2, \ldots, m$. This, by Proposition 2.5, is equivalent to the product $f \Delta u$ being a harmonic function. Using the formula $\triangle^{\mathbb{R}^m}(|x|^\alpha) = \alpha(\alpha - 2 + m)|x|^{\alpha-2}$ and a straightforward computation we have

$$\begin{aligned}\triangle^{\mathbb{R}^m} u &= \triangle^{\mathbb{R}^m}(x^i |x|^{-p}) = x^i \triangle^{\mathbb{R}^m} |x|^{-p} + 2\langle \operatorname{grad} x^i, \operatorname{grad} |x|^{-p}\rangle \\ &= p(p-m)x^i |x|^{-p-2}.\end{aligned}$$

For $f(x) = |x|^k$, we have

$$\begin{aligned}\triangle^{\mathbb{R}^m}(f \triangle^{\mathbb{R}^m} u) &= p(p-m)\triangle^{\mathbb{R}^m}(x^i |x|^{k-p-2}) \\ &= p(p-m)[x^i \triangle^{\mathbb{R}^m} |x|^{k-p-2} + 2\langle \operatorname{grad} x^i, \operatorname{grad} |x|^{k-p-2}\rangle] \\ &= p(p-m)(k-p-2)(k-p+m-2)x^i |x|^{k-p-4}.\end{aligned}$$

It follows that $u(x) = x^i |x|^{-p}$ is an f-biharmonic function with $f = |x|^k$ if and only if $p(p-m)(k-p-2)(k-p+m-2) = 0$. Solving this equation we have (i) $p = 0$, or (ii) $p = m$, or (iii) $k = p+2$, or (iv) $k = p+2-m$, from which the proposition follows. □

Remark. (A) One can check (see also [Balmuş et al. 2007]) that for the cases (i) $p = 0$ and (ii) $p = m$, the maps $\phi = x/|x|^p$ are actually harmonic maps. We know that in these cases these maps are f-biharmonic for any f. For $k = 0$ we have $f(x) = 1$ and hence f-biharmonicity reduces to biharmonicity. In this case, (iii) and (iv) imply that $\phi = x/|x|^p$ is a proper biharmonic map if and only if $p = -2$, or $p = m-2$. Note that the case $p = -2$ was missed in the list of [ibid., Remark 5.8].

(B) For $p \neq 0$, m, and $k \neq 0$, the maps in cases (iii) and (iv) provide infinitely many examples of proper f-biharmonic maps (i.e., which are neither harmonic nor biharmonic maps).

(C) It is well known that the inversion in sphere S^{m-1}, $\phi : \mathbb{R}^m \setminus \{0\} \to \mathbb{R}^m \setminus \{0\}$, $\phi(x) = x/|x|^2$ is a conformal map between the same dimensional Euclidean spaces. Note that the f-biharmonic map equation for conformal maps between the same dimensional spaces was derived in [Lu 2013], however, not a single example of such maps was found. Our Proposition 2.9 shows that there are infinitely many proper f-biharmonic conformal diffeomorphisms and all but one of which are proper f-biharmonic for at least two different choices of f functions. For a study of biharmonic diffeomorphisms see [Baird et al. 2008].

3. f-biharmonic submanifolds

Definition 3.1. A submanifold in a Riemannian manifold is called an *f-biharmonic submanifold* if the isometric immersion defining the submanifold is an f-biharmonic map.

From the definition and the relationships among harmonic, biharmonic and f-biharmonic maps we have the inclusions

$$\{\text{minimal}\} \subset \{\text{biharmonic}\} \subset \{f\text{-biharmonic}\}.$$

From now on we will call an f-biharmonic submanifold a *proper f-biharmonic submanifold* if it is neither a minimal nor a biharmonic submanifold.

Theorem 3.2. *Let $\phi : M^m \to N^{m+1}$ be an isometric immersion of codimension one with mean curvature vector $\eta = H\xi$. Then φ is an f-biharmonic if and only if*

$$\begin{cases} \Delta H - H|A|^2 + H\,\mathrm{Ric}^N(\xi,\xi) + H(\Delta f)/f + 2(\operatorname{grad}\ln f)H = 0, \\ 2A(\operatorname{grad} H) + \frac{m}{2}\operatorname{grad} H^2 - 2H(\mathrm{Ric}^N(\xi))^\top + 2HA(\operatorname{grad}\ln f) = 0, \end{cases} \tag{14}$$

where $\mathrm{Ric}^N : T_qN \to T_qN$ *denotes the Ricci operator of the ambient space defined by* $\langle \mathrm{Ric}^N(Z), W\rangle = \mathrm{Ric}^N(Z, W)$; *$A$ is the shape operator of the hypersurface with respect to the unit normal vector ξ; and* Δ, grad *are the Laplace and the gradient operator of the hypersurface respectively.*

Proof. It is well known that the tension field of the hypersurface is given by

$$\tau(\phi) = mH\xi. \tag{15}$$

From [Ou 2010, Theorem 2.1] we have the bitension field of the hypersurface:

$$\begin{aligned} \tau_2(\phi) = m(\Delta H - H|A|^2 + H\,\mathrm{Ric}^N(\xi,\xi))\xi \\ -m\big(2A(\operatorname{grad} H) + \tfrac{m}{2}(\operatorname{grad} H^2) - 2H(\mathrm{Ric}(\xi))^\top\big). \end{aligned} \tag{16}$$

To compute the term $\nabla^\phi_{\operatorname{grad} f}\tau(\phi)$, we choose a local orthonormal frame $\{e_i\}_{i=1,\dots,m}$ on M so that $\{d\phi(e_1), \dots, d\phi(e_m), \xi\}$ forms an adapted orthonormal frame of the ambient space defined on the hypersurface. Identifying $d\phi(X) = X$, $\nabla^\phi_X W = \nabla^N_X W$ we have

$$\nabla^\phi_{\operatorname{grad} f}\tau(\phi) = m\nabla^N_{\operatorname{grad} f} H\xi = m\{[(\operatorname{grad} f)H]\xi - A(\operatorname{grad} f)\}. \tag{17}$$

Substituting (15), (16) and (17) into the f-biharmonic map equation (2) and simplifying the result we obtain the theorem. $\square$

Corollary 3.3. *A hypersurface $\phi : M^m \to N^{m+1}(C)$ in a space form of constant sectional curvature C is f-biharmonic if and only if its mean curvature function H*

satisfies the equation

$$\begin{cases} \Delta H - H|A|^2 + mCH + H(\Delta f)/f + 2(\operatorname{grad}\ln f)H = 0, \\ 2A(\operatorname{grad} H) + \frac{1}{2}m \operatorname{grad} H^2 + 2HA(\operatorname{grad}\ln f) = 0. \end{cases} \tag{18}$$

Similarly:

Corollary 3.4. *A submanifold $\phi : M^m \to N^n(C)$ in a space form of constant sectional curvature C is f-biharmonic if and only if its mean curvature vector H satisfies the equation*

$$\begin{cases} \Delta^{\perp} H - (\Delta f/f)H - 2\nabla^{\perp}_{\operatorname{grad}\ln f} H + \operatorname{Trace} B(-, A_H -) + CmH = 0, \\ 2\operatorname{Trace} A_{\nabla^{\perp}_{(-)}H}(-) + \frac{1}{2}m \operatorname{grad}(|H|^2) + 2A_H(\operatorname{grad}\ln f) = 0, \end{cases}$$

where $\Delta^{\perp} H = -\operatorname{Trace}(\nabla^{\perp})^2 H$.

Corollary 3.5. *A compact nonzero constant mean curvature f-biharmonic hypersurface $\phi : M^m \to S^{m+1}$ in a sphere with $|A|^2 = \text{constant}$ is biharmonic.*

Proof. Substituting $H = \text{constant} \neq 0$ into the f-biharmonic hypersurface equation (18) we have

$$\begin{cases} \Delta f = (|A|^2 - m) f, \\ A(\operatorname{grad}\ln f) = 0. \end{cases} \tag{19}$$

If $|A|^2$ is constant, we have either $|A|^2 - m = 0$, in which case the first equation of (19) implies that f is a harmonic function, or $|A|^2 - m \neq 0$. In the latter case, the first line of (19) implies that f is either subharmonic or superharmonic since $f > 0$. Since M is compact, the well-known fact that any harmonic (subharmonic or superharmonic) function on a compact manifold is constant implies that f is a constant function. Thus, an f-biharmonic hypersurface is actually biharmonic. □

For classification of biharmonic submanifolds with parallel mean curvature vector and $|A|^2 = \text{constant}$ in sphere see [Balmuş et al. 2013].

In Euclidean space $\mathbb{R}^3$, any biharmonic surface is minimal (see, e.g., [Jiang 1987; Chen and Ishikawa 1998]), so there are no proper biharmonic surfaces. The first question we ask is: Are there proper f-biharmonic surfaces in $\mathbb{R}^3$? We will show that there are infinitely many. We achieve this by using a link between f-biharmonic surfaces and biharmonic conformal immersions of surfaces in a three-manifold. For the study of biharmonic conformal immersions of surfaces in three-manifolds we refer the reader to [Ou 2009; 2012a]. We recall that a surface (i.e., an isometric immersion) $\phi : M^2 \to (N^3, h)$ is said to admit a biharmonic conformal immersion into a three-manifold (N^3, h) if there exists a function $\lambda : M^2 \to (0, \infty)$ such that the conformal immersion $\phi : (M^2, \lambda^{-2}\phi^* h) \to (N^3, h)$ is biharmonic map. In this case, we also say that the surface $\phi : M^2 \to (N^3, h)$ can be biharmonically conformally immersed into the three-manifold (N^3, h) with conformal factor λ.

Corollary 3.6. (i) *A surface $\phi : M^2 \to (N^3, h)$ in a three-manifold is f-biharmonic if and only if the conformal immersion*

$$\phi : (M^2, f^{-1}\phi^* h) \to (N^3, h)$$

is a biharmonic map, i.e., the surface can be biharmonically conformally immersed into (N^3, h) with conformal factor $\lambda = f^{\frac{1}{2}}$.

(ii) *The circular cylinder $\phi : D = \{(\theta, z) \in (0, 2\pi) \times \mathbb{R}\} \to (\mathbb{R}^3, \delta_0)$ with $\phi(\theta, z) = (R\cos\theta, R\sin\theta, z)$ is an f-biharmonic surface for any function f from the family $f = \big(C_2 e^{\pm z/R} - C_1 C_2^{-1} R^2 e^{\mp z/R}\big)/2$, where C_1, C_2 are constants.*

Proof. Statement (i) follows from the definition of an f-biharmonic surface and Theorem 2.3, whilst (ii) is obtained by using (i) and [Ou 2009, Proposition 2]. □

4. f-biharmonic curves

Another special case of f-biharmonic maps is an f-biharmonic curve.

Lemma 4.1. *A curve $\gamma : (a, b) \to (N^m, g)$ parametrized by arclength is an f-biharmonic curve with a function $f : (a, b) \to (0, \infty)$ if and only if*

$$f\big(\nabla^N_{\gamma'}\nabla^N_{\gamma'}\nabla^N_{\gamma'}\gamma' - R^N(\gamma', \nabla^N_{\gamma'}\gamma')\gamma'\big) + 2f'\nabla^N_{\gamma'}\nabla^N_{\gamma'}\gamma' + f''\nabla^N_{\gamma'}\gamma' = 0. \tag{20}$$

Proof. Let $\gamma = \gamma(s)$ be parametrized by arclength. Then $e_1 = \partial/\partial s$ is an orthonormal frame on $((a, b), ds^2)$ and $d\gamma(e_1) = d\gamma(\partial/\partial s) = \gamma'$. Thus, the tension field of the curve is given by $\tau(\gamma) = \nabla^{\gamma}_{e_1} d\gamma(e_1) = \nabla^N_{\gamma'}\gamma'$. It is also easy to see that for a function $f : (a, b) \to (0, \infty)$, $\Delta f = f''$ and $\nabla^{\gamma}_{\operatorname{grad} f}\tau(\gamma) = f'\nabla^N_{\gamma'}\nabla^N_{\gamma'}\gamma'$. Substituting these into the f-biharmonic map equation gives the lemma. □

Theorem 4.2. *A curve $\gamma : (a, b) \to N^n(C)$ parametrized by arclength in an n-dimensional space form is a proper f-biharmonic curve if and only if one of the following cases happens*:

(i) $\kappa_2 = 0$, $f = c_1\kappa_1^{-\frac{3}{2}}$ *and the curvature κ_1 solves the ODE*

$$3\kappa_1'^2 - 2\kappa_1\kappa_1'' = 4\kappa_1^2(\kappa_1^2 - C);$$

(ii) $\kappa_2 \neq 0$, $\kappa_3 = 0$, $\kappa_2/\kappa_1 = c_3$, $f = c_1\kappa_1^{-\frac{3}{2}}$, *and the curvature κ_1 solves the ODE*

$$3\kappa_1'^2 - 2\kappa_1\kappa_1'' = 4\kappa_1^2[(1 + c_3^2)\kappa_1^2 - C].$$

Proof. Let $\gamma : (a, b) \to N^n(C)$ be a curve with arclength parametrization. Let $\{F_i, i = 1, 2, \dots, n\}$ be the Frenet frame along the curve $\gamma(s)$, which is obtained as the orthonormalization of the n-tuple $\big\{\nabla^{(k)}_{\partial/\partial s} d\gamma(\partial/\partial s) \mid k = 1, 2, \dots, n\big\}$. Then we have the following Frenet formula (see, e.g., [Laugwitz 1965]) along the curve:

$$\begin{cases} \nabla^{\gamma}_{\partial/\partial s} F_1 = \kappa_1 F_2, \\ \nabla^{\gamma}_{\partial/\partial s} F_i = -\kappa_{i-1} F_{i-1} + \kappa_i F_{i+1} \qquad \text{for } \; i = 2, 3, \dots, n-1, \\ \nabla^{\gamma}_{\partial/\partial s} F_n = -\kappa_{n-1} F_{n-1}, \end{cases}$$

where $\{\kappa_1, \kappa_2, \dots, \kappa_{n-1}\}$ are the curvatures of the curve γ.

Using this formula and a straightforward computation one finds the tension and the bitension fields of the curve given by

$$\tau(\gamma) = \nabla^N_{\gamma'} \gamma' = \kappa_1 F_2,$$
$$\nabla^N_{\gamma'} \nabla^N_{\gamma'} \gamma' = -\kappa_1^2 F_1 + \kappa_1' F_2 + \kappa_1 \kappa_2 F_3,$$
$$\tau_2(\gamma) = -3\kappa_1 \kappa_1' F_1 + (\kappa_1'' - \kappa_1 \kappa_2^2 - \kappa_1^3 + \kappa_1 C) F_2 + (2\kappa_1' \kappa_2 + \kappa_1 \kappa_2') F_3 + \kappa_1 \kappa_2 \kappa_3 F_4.$$

Substituting these into the f-biharmonic curve equation (20) and comparing the coefficients of both sides we have

$$\begin{cases} -3\kappa_1 \kappa_1' - 2\kappa_1^2 f'/f = 0, \\ \kappa_1'' - \kappa_1 \kappa_2^2 - \kappa_1^3 + \kappa_1 C + \kappa_1 f''/f + 2\kappa_1' f'/f = 0, \\ 2\kappa_1' \kappa_2 + \kappa_1 \kappa_2' + 2\kappa_1 \kappa_2 f'/f = 0, \\ \kappa_1 \kappa_2 \kappa_3 = 0. \end{cases} \tag{21}$$

It is easy to see that if $\kappa_1 = \text{constant} \neq 0$, then the first equation of (21) implies that f is constant and the curve γ is biharmonic. Also, if $\kappa_2 = \text{constant} \neq 0$, then the first and the third equations imply that f is constant and hence the curve γ is biharmonic again.

Now, if $\kappa_2 = 0$, then the f-biharmonic curve equation (21) is equivalent to

$$\begin{cases} 3\kappa_1'/\kappa_1 + 2f'/f = 0, \\ \kappa_1''/\kappa_1 - \kappa_1^2 + C + f''/f + 2(\kappa_1'/\kappa_1)(f'/f) = 0. \end{cases} \tag{22}$$

Integrating the first equation of (22) and substituting the result in to the second we obtain the statements in case (i).

Finally, if $\kappa_1 \neq \text{constant}$ and $\kappa_2 \neq \text{constant}$, then the system (21) is equivalent to

$$\begin{cases} f^2 \kappa_1^3 = c_1^2, \\ (f\kappa_1)'' = f\kappa_1(\kappa_2^2 + \kappa_1^2 - C), \\ f^2 \kappa^2 \kappa_2 = c_2, \\ \kappa_3 = 0. \end{cases} \tag{23}$$

Solving the first equation of (23) we obtain $f = c_1 \kappa_1^{-\frac{3}{2}}$. Substituting the first equation into the third one we obtain $\kappa_2/\kappa_1 = c_3$. Finally, substituting $\kappa_2/\kappa_1 = c_3$ and $f\kappa_1 = c_1 \kappa_1^{-\frac{1}{2}}$ into the second equation we obtain the results stated in case (ii). This completes the proof of the theorem. □

From the proof of Theorem 4.2 we have:

Corollary 4.3. *A curve $\gamma : (a, b) \to N^n(C)$ parametrized by arclength in an n-dimensional space form with constant geodesic curvature is biharmonic.*

It is known [Dimitrić 1992] that any biharmonic curve in a Euclidean space is a geodesic. It would be interesting to know if there is any proper f-biharmonic curve in a Euclidean space. Our next theorem gives a complete classification of proper f-biharmonic curves in $\mathbb{R}^3$ which, together with the fundamental theorem for curves in $\mathbb{R}^3$, can be used to produce many examples of proper f-biharmonic curves in a three-dimensional Euclidean space.

Theorem 4.4. *A curve $\gamma : (a, b) \to \mathbb{R}^3$ parametrized by arclength in a three-dimensional Euclidean space is a proper f-biharmonic curve if and only if*

(i) *γ is a planar curve with $\tau(s)=0$, $\kappa(s)=4c_2/(16+(c_2s+c_3)^2)$, and $f=c_1\kappa^{-\frac{3}{2}}$, where $c_1>0$, $c_2>0$, and c_3 are constants, or*

(ii) *γ is a general helix with $\kappa(s)=4c_2/(16(1+c^2)+(c_2s+c_3)^2)$, $\tau/\kappa=c$, and $f=c_1\kappa^{-\frac{3}{2}}$, where $c\neq 0$, $c_1>0$, $c_2>0$, and c_3 are constants.*

Proof. For the arclength-parametrized curve $\gamma : (a, b) \to \mathbb{R}^3$, we have the curvature $\kappa=\kappa_1$ and the torsion $\tau=\kappa_2$. Applying Theorem 4.2 with $C=0$ we conclude that the curve γ is a proper f-biharmonic curve if and only if

(i) $\tau=0$, $f=c_1\kappa^{-\frac{3}{2}}$ and the curvature κ solves the ODE

$$3\kappa'^2-2\kappa\kappa''=4\kappa^4, \quad \text{or}$$

(ii) $\tau\neq 0$, $\tau/\kappa=c$, $f=c_1\kappa_1^{-\frac{3}{2}}$, and the curvature κ solves the ODE

$$3\kappa'^2-2\kappa\kappa''=4(1+c^2)\kappa^4.$$

Solving the ODEs in each case and noting that $\tau=0$ means the curve is planar and $\tau/\kappa=c$ means the curve is a general helix (Lancret's theorem; see, e.g., [Barros 1997]) we obtain the theorem. □

Remark. (A) Recall that the fundamental theorem for curves in $\mathbb{R}^3$ states that for any given functions $p, q : [s_0, s_1] \to \mathbb{R}$ with $p(s)>0$ for all $s\in[s_0, s_1]$, there exists a unique (up to a rigid motion) curve in $\mathbb{R}^3$ whose curvature and torsion take on the prescribed functions $\kappa(s)=p(s)$, $\tau(s)=q(s)$. This, together with our Theorem 4.4, implies that there are many examples of proper f-biharmonic curves in $\mathbb{R}^3$.

(B) Our classification theorem also implies that proper f-biharmonic curves in $\mathbb{R}^3$ must be special subclasses of planar curves or general helices in $\mathbb{R}^3$. As the following example shows that there are general helices which are not proper f-biharmonic curves.

Example 3. The general helix $\gamma : I \to \mathbb{R}^3$ with $\gamma(s) = \left(\frac{2}{3}(1+\frac{s}{2})^{\frac{3}{2}}, \frac{2}{3}(1-\frac{s}{2})^{\frac{3}{2}}, \frac{s}{\sqrt{2}}\right)$ is never an f-biharmonic curve for any function f.

In fact, one can easily check that $|\gamma'(s)| = 1$ so s is an arclength parameter for the curve. A straightforward computation gives $\kappa(s) = \tau(s) = 1/(2\sqrt{2}\sqrt{4-s^2})$. So, the curve is indeed a general helix with $\tau/\kappa = 1$. Since the curvature is not of the form given in case (ii) of Theorem 4.4 we conclude that the helix is never an f-biharmonic curve for any f.

Finally, we give an example of a proper f-biharmonic planar curve to close this section.

Example 4. The planar curve $\gamma(s) = (4\ln|\sqrt{16+s^2}+s|, \sqrt{16+s^2}\,)$ is a proper f-biharmonic curve.

In fact, we can check that

$$\gamma'(s) = \left(\frac{4}{\sqrt{16+s^2}}, \frac{s}{\sqrt{16+s^2}}\right) \quad \text{and} \quad |\gamma'(s)| = 1.$$

So s is the arclength parameter of the curve. In this case, we have the curvature $\kappa(s) = |\gamma''(s)| = 4/(16+s^2)$ and, by case (i) of Theorem 4.4, the curve γ is a proper f-biharmonic curve with $f = 8c_1(16+s^2)^{\frac{3}{2}}$ for some constant $c_1 > 0$.

References

[Alías et al. 2013] L. J. Alías, S. C. García-Martínez, and M. Rigoli, "Biharmonic hypersurfaces in complete Riemannian manifolds", *Pacific J. Math.* **263**:1 (2013), 1–12. MR 3069073 Zbl 1278.53059

[Baird and Kamissoko 2003] P. Baird and D. Kamissoko, "On constructing biharmonic maps and metrics", *Ann. Global Anal. Geom.* **23**:1 (2003), 65–75. MR 2004c:58033 Zbl 1027.31004

[Baird et al. 2008] P. Baird, A. Fardoun, and S. Ouakkas, "Conformal and semi-conformal biharmonic maps", *Ann. Global Anal. Geom.* **34**:4 (2008), 403–414. MR 2009h:53140 Zbl 1158.53049

[Baird et al. 2010] P. Baird, A. Fardoun, and S. Ouakkas, "Liouville-type theorems for biharmonic maps between Riemannian manifolds", *Adv. Calc. Var.* **3**:1 (2010), 49–68. MR 2011c:58030 Zbl 1185.31004

[Balmuş et al. 2007] A. Balmuş, S. Montaldo, and C. Oniciuc, "Biharmonic maps between warped product manifolds", *J. Geom. Phys.* **57**:2 (2007), 449–466. MR 2007j:53066 Zbl 1108.58011

[Balmuş et al. 2008] A. Balmuş, S. Montaldo, and C. Oniciuc, "Classification results for biharmonic submanifolds in spheres", *Israel J. Math.* **168** (2008), 201–220. MR 2009j:53079 Zbl 1172.58004

[Balmuş et al. 2013] A. Balmuş, S. Montaldo, and C. Oniciuc, "Biharmonic PNMC submanifolds in spheres", *Ark. Mat.* **51**:2 (2013), 197–221. MR 3090194 Zbl 1282.53047

[Barros 1997] M. Barros, "General helices and a theorem of Lancret", *Proc. Amer. Math. Soc.* **125**:5 (1997), 1503–1509. MR 97g:53066 Zbl 0876.53035

[Caddeo et al. 2001] R. Caddeo, S. Montaldo, and C. Oniciuc, "Biharmonic submanifolds of S^3", *Internat. J. Math.* **12**:8 (2001), 867–876. MR 2002k:53123 Zbl 1111.53302

[Caddeo et al. 2002] R. Caddeo, S. Montaldo, and C. Oniciuc, "Biharmonic submanifolds in spheres", *Israel J. Math.* **130** (2002), 109–123. MR 2003c:53090 Zbl 1038.58011

[Chen 1991] B.-Y. Chen, "Some open problems and conjectures on submanifolds of finite type", *Soochow J. Math.* **17**:2 (1991), 169–188. MR 92m:53091 Zbl 0749.53037

[Chen and Ishikawa 1998] B.-Y. Chen and S. Ishikawa, "Biharmonic pseudo–Riemannian submanifolds in pseudo-Euclidean spaces", *Kyushu J. Math.* **52**:1 (1998), 167–185. MR 99b:53078 Zbl 0892.53012

[Chen and Munteanu 2013] B.-Y. Chen and M. I. Munteanu, "Biharmonic ideal hypersurfaces in Euclidean spaces", *Differential Geom. Appl.* **31**:1 (2013), 1–16. MR 3010073 Zbl 1260.53017

[Course 2004] N. Course, *f-harmonic maps*, Ph.D. thesis, University of Warwick, 2004, Available at http://neilcourse.99k.org/thesis/Neil_Course_f-harmonic_maps.pdf.

[Dimitrić 1992] I. Dimitrić, "Submanifolds of E^m with harmonic mean curvature vector", *Bull. Inst. Math. Acad. Sinica* **20**:1 (1992), 53–65. MR 93g:53087 Zbl 0778.53046

[Eells and Lemaire 1983] J. Eells and L. Lemaire, *Selected topics in harmonic maps*, CBMS Regional Conference Series in Mathematics **50**, Amer. Math. Soc., 1983. MR 85g:58030 Zbl 0515.58011

[Eells and Sampson 1964] J. Eells, Jr. and J. H. Sampson, "Harmonic mappings of Riemannian manifolds", *Amer. J. Math.* **86** (1964), 109–160. MR 29 #1603 Zbl 0122.40102

[Jiang 1986a] G. Y. Jiang, "2-harmonic isometric immersions between Riemannian manifolds", *Chinese Ann. Math. Ser. A* **7**:2 (1986), 130–144. In Chinese. MR 87k:53140 Zbl 0596.53046

[Jiang 1986b] G. Y. Jiang, "2-harmonic maps and their first and second variational formulas", *Chinese Ann. Math. Ser. A* **7**:4 (1986), 389–402. In Chinese; translated in *Note Mat.* **1**:1 (2008), 209–232. MR 88i:58039 Zbl 0628.58008

[Jiang 1987] G. Y. Jiang, "Some nonexistence theorems on 2-harmonic and isometric immersions in Euclidean space", *Chinese Ann. Math. Ser. A* **8**:3 (1987), 377–383. In Chinese. MR 89a:53071 Zbl 0637.53071

[Laugwitz 1965] D. Laugwitz, *Differential and Riemannian geometry*, Academic Press, New York, 1965. MR 30 #2406 Zbl 0139.38903

[Liang and Ou 2013] T. Liang and Y.-L. Ou, "Biharmonic hypersurfaces in a conformally flat space", *Results Math.* **64**:1-2 (2013), 91–104. MR 3095129 Zbl 1275.58013

[Lichnerowicz 1969] A. Lichnerowicz, "Applications harmoniques et variétés kähleriennes", pp. 341–402 in *Symposia Mathematica, III* (INDAM, Rome, 1968/69), Academic Press, London, 1969. MR 41 #7598 Zbl 0193.50101

[Loubeau and Ou 2010] E. Loubeau and Y.-L. Ou, "Biharmonic maps and morphisms from conformal mappings", *Tohoku Math. J.* (2) **62**:1 (2010), 55–73. MR 2011c:53153 Zbl 1202.53061

[Lu 2013] W.-J. Lu, "On f-biharmonic maps between Riemannian manifolds", preprint, 2013. arXiv 1305.5478

[Montaldo and Oniciuc 2006] S. Montaldo and C. Oniciuc, "A short survey on biharmonic maps between Riemannian manifolds", *Rev. Un. Mat. Argentina* **47**:2 (2006), 1–22. MR 2008a:53063 Zbl 1140.58004

[Nakauchi and Urakawa 2013] N. Nakauchi and H. Urakawa, "Biharmonic submanifolds in a Riemannian manifold with non-positive curvature", *Results Math.* **63**:1-2 (2013), 467–474. MR 3009698 Zbl 1261.58011

[Nakauchi et al. 2014] N. Nakauchi, H. Urakawa, and S. Gudmundsson, "Biharmonic maps into a Riemannian manifold of non-positive curvature", *Geom. Dedicata* **169** (2014), 263–272. MR 3175248

[Ou 2006] Y.-L. Ou, "p-harmonic morphisms, biharmonic morphisms, and nonharmonic biharmonic maps", *J. Geom. Phys.* **56**:3 (2006), 358–374. MR 2006e:53117 Zbl 1083.58015

[Ou 2009] Y.-L. Ou, "On conformal biharmonic immersions", *Ann. Global Anal. Geom.* **36**:2 (2009), 133–142. MR 2010g:53118 Zbl 1178.53054

[Ou 2010] Y.-L. Ou, "Biharmonic hypersurfaces in Riemannian manifolds", *Pacific J. Math.* **248**:1 (2010), 217–232. MR 2011i:53097 Zbl 1205.53066

[Ou 2012a] Y.-L. Ou, "Biharmonic conformal immersions into 3-dimensional manifolds", preprint, 2012. arXiv 1209.2104

[Ou 2012b] Y.-L. Ou, "Some constructions of biharmonic maps and Chen's conjecture on biharmonic hypersurfaces", *J. Geom. Phys.* **62**:4 (2012), 751–762. MR 2888980 Zbl 1239.58010

[Ou and Lu 2013] Y.-L. Ou and S. Lu, "Biharmonic maps in two dimensions", *Ann. Mat. Pura Appl.* (4) **192**:1 (2013), 127–144. MR 3011327 Zbl 1266.58004

[Ou and Tang 2012] Y.-L. Ou and L. Tang, "On the generalized Chen's conjecture on biharmonic submanifolds", *Michigan Math. J.* **61**:3 (2012), 531–542. MR 2975260 Zbl 1268.58015

[Ou and Wang 2011] Y.-L. Ou and Z.-P. Wang, "Constant mean curvature and totally umbilical biharmonic surfaces in 3-dimensional geometries", *J. Geom. Phys.* **61**:10 (2011), 1845–1853. MR 2012g:53120 Zbl 1227.58004

[Ouakkas 2008] S. Ouakkas, "Biharmonic maps, conformal deformations and the Hopf maps", *Differential Geom. Appl.* **26**:5 (2008), 495–502. MR 2009h:53142 Zbl 1159.58009

[Ouakkas et al. 2010] S. Ouakkas, R. Nasri, and M. Djaa, "On the f-harmonic and f-biharmonic maps", *JP J. Geom. Topol.* **10**:1 (2010), 11–27. MR 2011j:53114 Zbl 1209.58014

[Wang and Ou 2011] Z.-P. Wang and Y.-L. Ou, "Biharmonic Riemannian submersions from 3-manifolds", *Math. Z.* **269**:3-4 (2011), 917–925. MR 2860270 Zbl 1235.53065

[Wang et al. 2014] Z.-P. Wang, Y.-L. Ou, and H.-C. Yang, "Biharmonic maps from a 2-sphere", *J. Geom. Phys.* **77** (2014), 86–96. MR 3157904 Zbl 1284.58007

Received June 22, 2013.

YE-LIN OU
DEPARTMENT OF MATHEMATICS
TEXAS A&M UNIVERSITY – COMMERCE
PO BOX 3011
COMMERCE, TX 75429-3011
UNITED STATES

yelin.ou@tamuc.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 271, No. 2, 2014
dx.doi.org/10.2140/pjm.2014.271.479

UNITARY PRINCIPAL SERIES OF SPLIT ORTHOGONAL GROUPS

ALESSANDRA PANTANO, ANNEGRET PAUL AND SUSANA SALAMANCA RIBA

We prove the nonunitarity of a large set of parameters for Langlands quotients of minimal principal series of the orthogonal group SO(n+1, n), by showing that the set of unitary principal series parameters of SO(n+1, n) embeds into a (known) union of spherical unitary parameters for certain split orthogonal groups. In an earlier paper, we proved the nonunitarity of the genuine principal series of the metaplectic group Mp($2n$) attached to the same set of parameters. We conjecture that the set of parameters is complete in both cases and prove the conjecture for small rank groups and in the case of unipotent parameters.

1. Introduction

For $G = \mathrm{SO}(n{+}1, n)$ or the real metaplectic group $\mathrm{Mp}(2n)$, let MA be the Levi factor of a minimal parabolic subgroup of G. For every irreducible representation δ of M and every real character ν of A, we choose a minimal parabolic subgroup $P = MAN$ of G making ν weakly dominant, and we denote by

$$I_G(\delta, \nu) := \mathrm{Ind}_P^G(\delta \otimes \nu) \tag{1-1}$$

the (minimal) *principal series representation* of G induced from the representation $\delta \otimes \nu \otimes \mathrm{triv}$ of P. In the case of the metaplectic group, we assume that the representation is genuine, that is, does not factor to the symplectic group. Let $J_G(\delta, \nu)$ be the Langlands quotient of $I_G(\delta, \nu)$, that is, the distinguished irreducible composition factor containing the minimal K-type. We are interested in determining for which pairs (δ, ν) the irreducible representation $J_G(\delta, \nu)$ is unitarizable. We call this set the complementary series $\mathrm{CS}(G)$ of G. The *spherical* complementary series of $\mathrm{SO}(n{+}1, n)_0$ (with δ trivial) is denoted $\mathrm{CS}(\mathrm{SO}(n{+}1, n)_0, \delta_0)$. Our work is motivated by the following conjecture.

This material is based on work supported by NSF Grants DMS-0554278, DMS-0967583, and DMS-0967168.
MSC2010: 22E45.
Keywords: orthogonal groups, intertwining operators, petite K-types, complementary series, theta correspondence, unipotent representations, spherical unitary dual, Weyl group representations.

Conjecture 1. *There is a natural, well-defined bijection*

$$\mathrm{CS}(G) \longleftrightarrow \bigcup_{p+q=n} \mathrm{CS}(\mathrm{SO}(p+1,p)_0,\delta_0)\times \mathrm{CS}(\mathrm{SO}(q+1,q)_0,\delta_0).$$

Because the spherical unitary dual of every split orthogonal group is known, by [Barbasch 2010], this would give a complete description of the unitary principal series for these two families of groups G.

For the two groups under consideration, A is isomorphic to $\mathbb{R}^n$, and M is isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$, in the case of $\mathrm{SO}(n+1,n)$, and to $\mathbb{Z}/4\mathbb{Z}\times(\mathbb{Z}/2\mathbb{Z})^{n-1}$ for $\mathrm{Mp}(2n)$; moreover there is a natural one-to-one correspondence between M-types (i.e., irreducible representations of M) of $\mathrm{SO}(n+1,n)$ and *genuine* M-types of $\mathrm{Mp}(2n)$. Every M-type of $\mathrm{SO}(n+1,n)$ and every genuine M-type of $\mathrm{Mp}(2n)$ is contained in a unique *fine* K-type for G. Fix an M-type δ. We call the set of real parameters ν for which $J_G(\delta,\nu)$ is unitary the δ-*complementary series* of G:

$$\mathrm{CS}(G,\delta) := \{\nu\in\mathfrak{a}^*_{\mathbb{R}} \mid J_G(\delta,\nu) \text{ is unitary}\}. \tag{1-2}$$

For all $w\in W$, $J_G(\delta,\nu)\simeq J_G(w\delta,w\nu)$; hence

$$\mathrm{CS}(G,w\cdot\delta) = w^{-1}\cdot \mathrm{CS}(G,\delta). \tag{1-3}$$

It follows that $\mathrm{CS}(G,\delta)$ is invariant under the action of the stabilizer W^δ of δ in W and depends only on the orbit of the M-type δ under the action of W. Here W is the Weyl group of A in G, which may be identified with the Weyl group of the root system of G.

The Weyl groups of $\mathrm{SO}(n+1,n)$ and of $\mathrm{Mp}(2n)$ are isomorphic. Moreover if an M-type of $\mathrm{SO}(n+1,n)$ and a genuine M-type of $\mathrm{Mp}(2n)$ correspond to each other in the above mentioned bijection, then their stabilizers are also isomorphic. The W-orbits of M-types of $\mathrm{SO}(n+1,n)$ and of genuine M-types of $\mathrm{Mp}(2n)$ are parametrized by pairs of nonnegative integers (p,q) with $p+q=n$; we choose a representative $\delta^{p,q}$ in each orbit (see (2-15)). Then the $W^{\delta^{p,q}}$-action leads to a natural splitting of each real parameter ν into a pair $(\nu^p,\nu^q)\in\mathbb{R}^p\times\mathbb{R}^q$ (see Section 3).

In [Pantano et al. 2010], we proved that the $\delta^{p,q}$-complementary series of $\mathrm{Mp}(2n)$ embeds into the product of the spherical complementary series of $\mathrm{SO}(p+1,p)_0$ with that of $\mathrm{SO}(q+1,q)_0$. In this paper, we show the analogous result for $\mathrm{SO}(n+1,n)$. In particular, the following theorem and the corresponding result for $\mathrm{Mp}(2n)$ make Conjecture 1 more precise.

Theorem 2. *Let* $G=\mathrm{SO}(n+1,n)$, *and let* $\nu=(\nu_1,\dots,\nu_n)$ *be a real character of* A. *For each pair of nonnegative integers* (p,q) *such that* $p+q=n$, *write* $\nu=(\nu^p|\nu^q)$ *with*

(1-4) $$\nu^p := (\nu_1, \dots, \nu_p) \quad \textit{and} \quad \nu^q := (\nu_{p+1}, \dots, \nu_n).$$

The map

(1-5) $$\mathrm{CS}(G, \delta^{p,q}) \to \mathrm{CS}(\mathrm{SO}(p+1, p)_0, \delta_0) \times \mathrm{CS}(\mathrm{SO}(q+1, q)_0, \delta_0)$$

taking ν *to* (ν^p, ν^q) *is a well-defined injection. Here* δ_0 *denotes the trivial* M*-type.*

The theorem asserts that if the (Hermitian) Langlands quotient $J(\delta^{p,q}, (\nu^p|\nu^q))$ of G is unitary, then the spherical (Hermitian) Langlands quotients $J(\delta_0, \nu^p)$ of $\mathrm{SO}(p+1, p)_0$ and $J(\delta_0, \nu^q)$ of $\mathrm{SO}(q+1, q)_0$ must both be unitary. Due to [Barbasch 2010], these conditions can be explicitly checked. We give a description of the spherical unitary parameters for $\mathrm{SO}(p+1, p)_0$ in Section 8.

Corollary 3. *Let* $G = \mathrm{SO}(n+1, n)$ *and let* $\nu = (\nu^p|\nu^q)$ *be a real character of* A, *as in (1-4). If the spherical Langlands quotient* $J(\delta_0, \nu^p)$ *of* $\mathrm{SO}(p+1, p)_0$ *or the spherical Langlands quotient* $J(\delta_0, \nu^q)$ *of* $\mathrm{SO}(q+1, q)_0$ *is not unitary, then the Langlands quotient* $J(\delta^{p,q}, (\nu^p|\nu^q))$ *of* $\mathrm{SO}(n+1, n)$ *is also not unitary.*

The theorem gives *nonunitarity certificates* for $\mathrm{SO}(n+1, n)$. In general, proving the unitarity of a representation is much harder than showing that it is not unitary. We conjecture that the spherical complementary series of $\mathrm{SO}(p+1, p)_0 \times \mathrm{SO}(q+1, q)_0$ gives an exhaustive parametrization of unitary parameters for the Langlands quotients of the $\delta^{p,q}$-principal series of both $G = \mathrm{Mp}(2n)$ and $\mathrm{SO}(n+1, n)$. To prove this we must show that, for each pair of parameters

$$\nu^p \in \mathrm{CS}(\mathrm{SO}(p+1, p)_0, \delta_0) \quad \text{and} \quad \nu^q \in \mathrm{CS}(\mathrm{SO}(q+1, q)_0, \delta_0),$$

the Langlands quotient $J_G(\delta^{p,q}, (\nu^p|\nu^q))$ of G is unitary for both $G = \mathrm{Mp}(2n)$ and $\mathrm{SO}(n+1, n)$. In this paper, we show that it is sufficient to prove this for one of the two families of groups.

Theorem 4. *Suppose the Langlands quotient* $J_G(\delta^{p,q}, (\nu^p|\nu^q))$ *is unitary for all* $\nu^p \in \mathrm{CS}(\mathrm{SO}(p+1, p)_0, \delta_0)$ *and* $\nu^q \in \mathrm{CS}(\mathrm{SO}(q+1, q)_0, \delta_0)$, *and all* $p + q = n$, *for* $G = \mathrm{Mp}(2n)$. *Then the same is true for* $G = \mathrm{SO}(n+1, n)$; *and vice versa.*

In [Pantano et al. 2010], we proved the unitarity of the principal series of $\mathrm{Mp}(2n)$ attached to our list of parameters for some small rank cases; for the general case, we exhibited two large families of spherical unitary parameters for the product $\mathrm{SO}(p+1, p)_0 \times \mathrm{SO}(q+1, q)_0$, which give rise to $\delta^{p,q}$-complementary series of $\mathrm{Mp}(2n)$. In this paper, we show some of the $\mathrm{SO}(n+1, n)$ analogues of these results. In particular, we obtain:

Theorem 5. *Let* $n = p + q \leq 4$ *and take elements* $\nu^p \in \mathrm{CS}(\mathrm{SO}(p+1, p)_0, \delta_0)$ *and* $\nu^q \in \mathrm{CS}(\mathrm{SO}(q+1, q)_0, \delta_0)$. *Then* $J_G(\delta^{p,q}, (\nu^p|\nu^q))$ *of* G *is unitary for both* $G = \mathrm{Mp}(2n)$ *and* $\mathrm{SO}(n+1, n)$.

Proving the unitarity of the δ-principal series for our collection of parameters might use the ideas for Barbasch's proof [2010] in the spherical case. First using normalized parabolic induction and deformation of parameters, Barbasch reduces the question of unitarity to *unipotent* parameters. These are parameters that correspond to special unipotent representations, that is, those that are attached to unipotent Arthur parameters [1989]. Then he proves the unitarity of these special unipotent representations.

Arthur packets and "unipotent" parameters are defined for linear groups only; for genuine (p, q)-principal series representations of the metaplectic group, we call a parameter $\nu = (\nu^p|\nu^q)$ *unipotent* if it is a unipotent parameter for $\mathrm{SO}(n+1, n)$.

In his recent book [Arthur 2013], the author proves that for split classical groups, all special unipotent representations are unitary. We check that $J_{\mathrm{SO}(n+1,n)}(\delta^{p,q}, \nu)$ is special unipotent if and only if ν^p and ν^q are spherical unipotent parameters for $\mathrm{SO}(p+1, p)_0$ and $\mathrm{SO}(q+1, q)_0$, respectively. Using the theta correspondence for dual pairs of the form $(O(r, s), \mathrm{Sp}(2m, \mathbb{R}))$, we prove that the unitarity of the special unipotent principal series of $\mathrm{SO}(n+1, n)$ implies the unitarity of the corresponding representation of $\mathrm{Mp}(2n)$. We obtain the following result.

Theorem 6. *If* ν^p, ν^q *are spherical unipotent parameters for* $\mathrm{SO}(p+1, p)_0$ *and* $\mathrm{SO}(q+1, q)_0$, *respectively*, *then* $\nu = (\nu^p|\nu^q) \in \mathrm{CS}(G, \delta^{p,q})$ *for* $G = \mathrm{Mp}(2n)$ *and* $\mathrm{SO}(n+1, n)$.

Sections 2 through 4 of this paper contain the proof of Theorem 2. In Section 2, we collect some structural facts about our groups and the K-types. Section 3 contains a careful outline of the main argument. Here we reduce the proof of our main theorem to an explicit matching of W-types (Theorem 9). The heart of the calculations is in Section 4, with the proof of Theorem 9. In Sections 5 through 7, we address what we know about the unitarity of the δ-principal series. Section 5 is devoted to proving Theorem 5. In Section 6, we use the theta correspondence to relate the complementary series of $\mathrm{SO}(n+1, n)$ and $\mathrm{Mp}(2n)$ to each other. Unipotent parameters and their unitarity are discussed in Section 7, and in Section 8, we give a description of the spherical unitary parameters for $\mathrm{SO}(n+1, n)$.

2. The structure of SO(n+1, n)

Let $G = \mathrm{SO}(n+1, n)$ be defined by

$$G = \mathrm{SO}(n+1, n) := \{g \in \mathrm{SL}(2n+1, \mathbb{R}) \colon (g^t)Jg = J\}, \tag{2-1}$$

with $J = \mathrm{diag}(-I_{n+1}, I_n)$. Here I_s denotes the identity matrix of size $s \times s$. We denote by $\mathfrak{g}_0$ the Lie algebra of G, and by $\mathfrak{g}$ its complexification. Let $\mathfrak{k}_0$ be the maximal compact Cartan subalgebra of $\mathfrak{g}_0$ corresponding to the Cartan involution $\theta(X) = -X^t$, and let $K = S(O(n+1) \times O(n))$ be the corresponding compact subgroup of G. Write $\mathfrak{g}_0 = \mathfrak{k}_0 \oplus \mathfrak{p}_0$ for the Cartan decomposition of $\mathfrak{g}_0$, and let $\mathfrak{a}_0$ be the maximal abelian subspace

$$\mathfrak{a}_0 = \left\{ X(B) := \begin{pmatrix} 0 & B \\ B^t & 0 \end{pmatrix} : B = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & s_n \\ 0 & 0 & \dots & s_{n-1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & s_2 & \dots & 0 & 0 \\ s_1 & 0 & \dots & 0 & 0 \end{pmatrix}, s_1, \dots, s_n \in \mathbb{R} \right\} \tag{2-2}$$

of $\mathfrak{p}_0$. For all $i = 1, \dots, n$, let $\epsilon_i \in \mathfrak{a}_0^*$ be defined by $\epsilon_i(X(B)) = s_i$. Then the restricted roots are

$$\Delta(\mathfrak{g}_0, \mathfrak{a}_0) = \{\pm\epsilon_i \pm \epsilon_j\}_{1 \le i < j \le n} \cup \{\pm\epsilon_k\}_{k=1,\dots,n}. \tag{2-3}$$

They form a root system Δ of type B_n. The corresponding Weyl group $W = W(\Delta)$ is isomorphic to $S_n \ltimes (\mathbb{Z}/2\mathbb{Z})^n$, and consists of all permutations and sign changes on n coordinates. Note that W can be realized as $N_K(A)/Z_K(A)$, where A is the vector group $\exp(\mathfrak{a}_0)$.

For each root $\alpha \in \Delta$ we choose a Lie algebra homomorphism

$$\phi_\alpha : \mathfrak{sl}(2, \mathbb{R}) \to \mathfrak{g}_0 = \mathfrak{so}(n+1, n), \tag{2-4}$$

as in [Vogan 1981, (4.3.6)], and we let G_α be the corresponding connected subgroup of $\mathrm{SO}(n+1, n)$. Moreover we define

$$Z_\alpha := \phi_\alpha \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2-5}$$

$$\sigma_\alpha := \exp\Big(\frac{\pi}{2} Z_\alpha\Big), \tag{2-6}$$

$$m_\alpha := \exp(\pi Z_\alpha) = \sigma_\alpha^2. \tag{2-7}$$

We make the following choices:

$$\begin{aligned} Z_{\epsilon_i+\epsilon_j} &= (E_{n+2-j,n+2-i} - E_{n+2-i,n+2-j}) + (E_{n+1+i,n+1+j} - E_{n+1+j,n+1+i}), \\ Z_{\epsilon_i-\epsilon_j} &= (E_{n+2-j,n+2-i} - E_{n+2-i,n+2-j}) - (E_{n+1+i,n+1+j} - E_{n+1+j,n+1+i}) \\ Z_{\epsilon_k} &= 2E_{1,n+2-k} - 2E_{n+2-k,1}. \end{aligned}$$

Then

$$\sigma_{\epsilon_k} = I - 2(E_{1,1} + E_{n+2-k,n+2-k}), \quad m_{\epsilon_k} = I,$$

and for $\alpha = \epsilon_i \pm \epsilon_j$,

$$\begin{aligned} \sigma_\alpha &= I - (E_{n+2-i,n+2-i} + E_{n+2-j,n+2-j} + E_{n+1+i,n+1+i} + E_{n+1+j,n+1+j}) + Z_\alpha, \\ m_\alpha &= I - 2(E_{n+2-i,n+2-i} + E_{n+2-j,n+2-j} + E_{n+1+i,n+1+i} + E_{n+1+j,n+1+j}). \end{aligned}$$

As usual, the symbol $E_{i,j}$ denotes the (i,j)-elementary matrix.

The centralizer of A in K is denoted by M, and consists of all elements

$$T(t_1,\dots,t_n)=\operatorname{diag}(1;t_n,t_{n-1},\dots,t_1;t_1,t_2,\dots,t_n), \tag{2-8}$$

with $t_1,\dots,t_n=\pm 1$. It is an abelian group of order 2^n, isomorphic to $(\mathbb{Z}/2\mathbb{Z})^n$. The dual group $\widehat{M}$ is generated by the M-types $\{\delta_i : 1\le i\le n\}$, where

$$\delta_i(T(t_1,\dots,t_n)) := t_i. \tag{2-9}$$

For every subset S of $\{1,2,\dots,n\}$, we denote by δ_S the irreducible representation of M satisfying

$$\delta_S(T(t_1,\dots,t_n))=\prod_{i\in S} t_i. \tag{2-10}$$

The Weyl group W acts on the set of irreducible representations of M by

$$(s_\alpha\cdot\delta)(m) := \delta(\sigma_\alpha^{-1} m\sigma_\alpha) \quad \text{for all } m\in M \text{ and } \alpha\in\Delta. \tag{2-11}$$

The stabilizer of δ in W is the subgroup

$$W^\delta := \{w\in W : w\cdot\delta\simeq\delta\} \tag{2-12}$$

of W. It is easy to check that

$$s_{\epsilon_k}\cdot\delta_S=\delta_S \quad \text{and} \quad s_{\epsilon_i\pm\epsilon_j}\cdot\delta_S=\begin{cases}\delta_S & \text{if } i,j\in S \text{ or } i,j\in S^C,\\ \delta_{S\triangle\{i,j\}} & \text{otherwise,}\end{cases} \tag{2-13}$$

for all $1\le i<j\le n$ and all $k=1,\dots,n$. Here S^C denotes the complement of S in $\{1,2,\dots,n\}$, and the symbol $(S\triangle T)$ denotes the symmetric difference of the two subsets. If $q=\#S$ (the cardinality of S) and $p=\#(S^C)$, this is the Weyl group of a root system of type $B_p\times B_q$.

Equation (2-13) also shows that

$$W\cdot\delta_S=\{\delta_T : \#S=\#T\}. \tag{2-14}$$

This implies that the Weyl group orbits of $\widehat{M}$ can be parametrized by pairs of nonnegative integers (p,q) with $p+q=n$; for each such pair, we choose a representative

$$\delta^{p,q} := \delta_{\{p+1,p+2,\dots,n\}}=\delta_{p+1}\cdot\delta_{p+2}\cdots\delta_n. \tag{2-15}$$

With this notation, $W^{\delta^{p,q}}=W(B_p)\times W(B_q)$, and the trivial M-type is $\delta^{n,0}$.

Let δ be an irreducible representation of M. A root $\alpha\in\Delta$ is called "good" for δ if $\delta(m_\alpha)\ne -1$. Otherwise we say that α is a "bad root" for δ. The set of good roots for $\delta=\delta_S$ is

$$\Delta_{\delta_S}=\{\pm\epsilon_i\pm\epsilon_j : i,j\in S \text{ or } i,j\in S^C\}\cup\{\pm\epsilon_k : 1\le k\le n\}. \tag{2-16}$$

If $q = \#S$ and $p = \#(S^C)$, this is a root system of type $B_p \times B_q$.

Remark 7. For every M-type δ, the Weyl group of the root system Δ_δ coincides with the stabilizer W^δ of δ in W.

Recall the definition of *fine* K-types given in [Vogan 1981, Section 4] (see also [Adams et al. 2007, Definition 4.9]). The fine K-types of $\mathrm{SO}(n+1, n)$ are given by

$$\mu_{p,q} = \mathrm{triv} \otimes \Lambda^q(\mathbb{C}^n) \tag{2-17}$$

for each value of $0 \le q \le n$. The restriction of $\mu_{p,q}$ to M is

$$W \cdot \delta^{p,q}$$

(see (2-15)), and its highest weight is

$$\begin{cases} (0, \dots, 0; \underbrace{1, \dots, 1}_{q}, 0, \dots, 0; +) & \text{if } q \le n/2, \\ (0, \dots, 0; \underbrace{1, \dots, 1}_{n-q}, 0, \dots, 0; -) & \text{otherwise.} \end{cases}$$

Recall that since $K = S(O(n+1) \times O(n))$ is disconnected, the highest weight does not necessarily determine the K-type uniquely, so we use a sign to distinguish two representations with the same highest weight.

Remark 8. (a) The restriction of a fine K-type to M consists of the W-orbit of a single M-type.

(b) Every M-type δ is contained in the restriction to M of a unique fine K-type μ_δ.

3. Nonunitarity certificates for complementary series of SO(n+1, n)

For $G = \mathrm{SO}(n+1, n)$, recall the definition of the δ-complementary series $\mathrm{CS}(G, \delta)$ in (1-2). This is a closed set because unitarity is a closed condition. As seen in Section 2, it suffices to consider the complementary series attached to a single M-type in each W-orbit for the action of the Weyl group on $\widehat{M}$. Such orbits are parametrized by pairs of nonnegative integers (p, q) with $p + q = n$. In each orbit, we choose the representative $\delta^{p,q}$ introduced in (2-15); the corresponding fine K-type is $\mu_{\delta^{p,q}}$ (see (2-17)). The stabilizer $W^{\delta^{p,q}} \cong W(B_p) \times W(B_q)$ acts on a (real) continuous parameter $\nu \in \mathfrak{a}^*_{\mathbb{R}}$ by sign changes and separate permutations of the first p and the last q coordinates. This leads to a natural splitting of each parameter ν into a pair $(\nu^p|\nu^q)$ as in Theorem 2.

Theorem 2 gives a comparison between the set of (real) unitary parameters for principal series representations of different groups. On the one hand, we have a $\delta^{p,q}$-principal series of the group $G = \mathrm{SO}(n+1, n)$; on the other hand, we have a spherical principal series for the group

$$G^{\delta^{p,q}} = \mathrm{SO}(p+1,\, p)_0 \times \mathrm{SO}(q+1,\, q)_0.$$

This group is intrinsically related to the M-type $\delta^{p,q}$ of G, and only depends on the system of good roots $\Delta_{\delta^{p,q}}$ (see (2-16)) associated to this M-type. Precisely, $G^{\delta^{p,q}}$ is the connected real split group corresponding to the root system

$$\Delta(G^{\delta^{p,q}}) := \Delta_{\delta^{p,q}}.$$

Consider a parameter $\nu = (\nu^p | \nu^q)$ as in Theorem 2. Write $I_{G^{\delta^{p,q}}}(\delta_0, \nu)$ for the spherical principal series of $G^{\delta^{p,q}}$ with parameter ν, and consider the possibly nonspherical representation $I(\delta^{p,q}, \nu)$ of $\mathrm{SO}(n+1, n)$ with the same parameter. The long Weyl group element of both G and $G^{\delta^{p,q}}$ is equal to $-\mathrm{Id}$, hence

$$w_0 \cdot \nu = -\nu, \quad w_0 \cdot \delta_0 = \delta_0, \quad w_0 \cdot \delta^{p,q} = \delta^{p,q}. \tag{3-1}$$

These are exactly the conditions that ν, δ_0 and $\delta^{p,q}$ must satisfy so that the above mentioned principal series representations admit an invariant Hermitian form. If ν is (weakly) dominant, then the Langlands quotients $J_{G^{\delta^{p,q}}}(\delta_0, \nu)$ and $J(\delta^{p,q}, \nu)$ are the quotients of the appropriate principal series by the radical of the Hermitian form. Hence they inherit a nondegenerate Hermitian form, and they are unitary if and only if the original form on the principal series is (positive) semidefinite.

To study the unitarity of a Langlands quotient, one needs to look at the signature of the Hermitian intertwining operators on the principal series which induce the form. Luckily these intertwining operators are very well understood. A thorough description can be found in [Barbasch et al. 2008] for split linear groups (such as $\mathrm{SO}(n+1, n)$ and $G^{\delta^{p,q}}$). We will not review the theory here but only recall the main results. The interested reader may consult the reference above for details.

First consider the spherical Langlands quotient $J_{G^{\delta^{p,q}}}(\delta_0, \nu)$ of the group $G^{\delta^{p,q}}$. Hecke algebra considerations reduce the study of the unitarity of $J_{G^{\delta^{p,q}}}(\delta_0, \nu)$ to the analysis of the signature of certain (relatively simple) "algebraic" operators. Precisely there is one operator $A(w_0, \psi, \nu)$ for every representation ψ of the Weyl group of $G^{\delta^{p,q}}$; the representation $J_{G^{\delta^{p,q}}}(\delta_0, \nu)$ is unitary if and only if

$$A(w_0, \psi, \nu) \text{ is positive semidefinite, for all } \psi \in \widehat{W(G^{\delta^{p,q}})}.$$

Barbasch [2010] has identified a small set of $W(G^{\delta^{p,q}})$-types (called "relevant") that detect unitarity, in the sense that $J_{G^{\delta^{p,q}}}(\delta_0, \nu)$ is unitary if and only if

$$A(w_0, \psi, \nu) \text{ is positive semidefinite, for all relevant } \psi \in \widehat{W(G^{\delta^{p,q}})}. \tag{3-2}$$

Next we consider the Langlands quotient $J(\delta^{p,q}, \nu)$ of $\mathrm{SO}(n+1, n)$. For every K-type μ, there is a family of (much harder) "analytic" intertwining operators

$$T(w_0, \mu, \delta^{p,q}, \nu), \tag{3-3}$$

with the property that the Langlands quotient $J(\delta^{p,q}, \nu)$ is unitary if and only if

$$T(w_0, \mu, \delta^{p,q}, \nu) \text{ is positive semidefinite, for all } \mu \in \widehat{K}.$$

The operator $T(w_0, \mu, \delta^{p,q}, \nu)$ is defined on the space $\mathrm{Hom}_K(\mu, I_G(\delta^{p,q}, \nu))$, which is isomorphic to

$$V_\mu[\delta^{p,q}] := \mathrm{Hom}_M(\mu, \delta^{p,q})$$

by Frobenius reciprocity. The stabilizer $W^{\delta^{p,q}}$ of the M-type $\delta^{p,q}$ acts naturally on this space. Note that the group $W^{\delta^{p,q}}$ coincides with the Weyl group of the system $\Delta_{\delta^{p,q}}$ of good roots for $\delta^{p,q}$:

$$W(\Delta_{\delta^{p,q}}) \simeq W(B_p) \times W(B_q) \simeq W(G^{\delta^{p,q}}).$$

Hence, for every K-type μ, we obtain a representation ψ_μ of the Weyl group of $G^{\delta^{p,q}}$ on the domain $V_\mu[\delta^{p,q}]$ of the intertwining operator $T(w_0, \mu, \delta^{p,q}, \nu)$.

The operator $T(w_0, \mu, \delta^{p,q}, \nu)$ is in general hard to compute, but if the K-type μ is sufficiently small (more precisely, "petite"; see [Barbasch et al. 2008, Sections 4.5 and 4.6] for a precise definition), then $T(w_0, \mu, \delta^{p,q}, \nu)$ depends only on the $W(\Delta_{\delta^{p,q}})$-structure of $V_\mu[\delta^{p,q}]$. One measure of the size of a K-type is its level (see Definition 11).

The following facts are crucial:

(1) Every K-type of level at most 2 is automatically petite.

(2) If μ is petite, then the "analytic" operator on μ coincides with the "algebraic" operator on the $W(\Delta_{\delta^{p,q}})$-type $V_\mu[\delta^{p,q}]$:

$$T(w_0, \mu, \delta^{p,q}, \nu) = A(w_0, V_\mu[\delta^{p,q}], \delta_0, \nu).$$

(3) For $G = \mathrm{SO}(n+1, n)$, every relevant $W(\Delta_{\delta^{p,q}})$-type ψ occurs in the representation of $W(\Delta_{\delta^{p,q}})$ on the space $V_\mu[\delta^{p,q}]$ for some petite K-type μ (of level 2).

The first two claims are well known; the proof already appears in [Barbasch et al. 2008] (see also [Pantano et al. 2010] for the corresponding results for double covers of split groups such as $\mathrm{Mp}(2n)$). The third claim is Theorem 9 below.

Theorem 9. *For every relevant $W^{\delta^{p,q}}$-type ψ, there exists a K-type μ of level at most 2 such that*

$$\psi_\mu \cong \psi. \tag{3-4}$$

The proof of Theorem 9 is given in Section 4. This concludes the proof of Theorem 2.

Because the spherical unitary dual of split groups of type B is known, Theorem 2 and Corollary 3 provide a set of nonunitarity certificates for (Langlands quotients of) minimal principal series of $\mathrm{SO}(n+1, n)$. We give an example.

By [Adams et al. 2007, Lemma 14.6], the spherical Langlands quotient of $O(k+1, k)$, with parameter, $(a_1 \geq a_2 \geq \cdots \geq a_k)$ is not unitary if the last coordinate a_k is strictly greater than $\frac{1}{2}$, or if there is a jump strictly greater than 1 between two consecutive coordinates. Because $O(k+1, k)$ and $\mathrm{SO}(k+1, k)_0$ have the same spherical complementary series, Corollary 3 implies that an analogous result must be true for $\mathrm{SO}(n+1, n)$.

Proposition 10. *Let $G = \mathrm{SO}(n+1, n)$ and let $\nu = (\nu_1, \ldots, \nu_n)$ be a real character of A. We may assume that*

$$\nu_1 \geq \cdots \geq \nu_p \geq 0 \quad \textit{and} \quad \nu_{p+1} \geq \cdots \geq \nu_n \geq 0.$$

Suppose that any of the following conditions holds:

(1) $\nu_p > \frac{1}{2}$ *or* $\nu_n > \frac{1}{2}$, *or*

(2) $\nu_i - \nu_{i+1} > 1$, *for some i with $1 \leq i \leq p-1$, or $p+1 \leq i \leq n-1$.*

Then the Langlands quotient $J(\delta^{p,q}, \nu)$ of $\mathrm{SO}(n+1, n)$ is not unitary.

4. A matching of petite K-types with relevant W^δ-types

Given $\delta^{p,q}$ as in the previous section, recall the stabilizer of $\delta^{p,q}$ in W

$$W^{\delta^{p,q}} \simeq W(B_p) \times W(B_q) \subseteq W(B_n).$$

We let $W(B_p)$ act on the first p coordinates, and $W(B_q)$ on the last q coordinates.

The M-type $\delta^{p,q}$ is contained in the (unique) fine K-type

$$\mu_{\delta^{p,q}} = \mathrm{triv} \otimes \Lambda^q(\mathbb{C}^n).$$

Here $\mathbb{C}^n$ represents the standard representation of $O(n)$ with (standard) basis $\{v_1, v_2, \ldots, v_n\}$. For every K-type μ whose restriction to M contains $\delta^{p,q}$, we denote by E_μ the vector space carrying the K-type μ and by $E_\mu(\delta^{p,q})$ the isotypic component of the M-type $\delta^{p,q}$ inside E_μ. In particular, when μ is the fine K-type $\mu_{\delta^{p,q}}$, we set

$$E_{\mu_{\delta^{p,q}}} = \Lambda^q(\mathbb{C}^n),$$

and we let $E_{\mu_{\delta^{p,q}}}(\delta^{p,q})$ be the one-dimensional space of $\Lambda^q(\mathbb{C}^n)$ spanned by the vector

$$u = v_{p+1} \wedge v_{p+2} \wedge \cdots \wedge v_n. \tag{4-1}$$

Given the M-type $\delta^{p,q}$ and a K-type μ containing $\delta^{p,q}$, one can look at the representation ψ_μ of $W^{\delta^{p,q}}$ on $\mathrm{Hom}_M(\mu, \delta^{p,q})$, or, equivalently, on

$$V_\mu[\delta^{p,q}] := \mathrm{Hom}_{\mathbb{C}}(E_\mu(\delta^{p,q}), E_{\mu_{\delta^{p,q}}}(\delta^{p,q})).$$

For each $[\sigma] \in W^{\delta^{p,q}}$ we choose a representative σ in K. Then, for $T \in V_\mu[\delta^{p,q}]$, we define $\psi_\mu([\sigma])(T)$ to be the map from $E_\mu(\delta^{p,q})$ to $E_{\mu_{\delta^{p,q}}}(\delta^{p,q})$ given by

$$\big(\psi_\mu([\sigma])(T)\big)(w) = \mu_{\delta^{p,q}}(\sigma)\big(T(\mu(\sigma^{-1})w)\big) \quad \text{for all } w \in E_\mu(\delta^{p,q}). \tag{4-2}$$

We are interested in computing the set of representations ψ_μ associated to petite K-types μ.

The notion of "petite" K-type for real split groups is carefully explained in [Barbasch et al. 2008, Sections 4.5 and 4.6]. For the purpose of this paper, it is sufficient to consider K-types of level at most 2, which are necessarily petite. We recall the definition of the "level" of a K-type. Recall the elements Z_α from Section 2.

Definition 11 [Adams et al. 2007, Section 4]. An irreducible representation μ of K is said to be level k if $|\gamma| \le k$ for every root α and every eigenvalue γ of $d\mu(iZ_\alpha)$.

Recall that $W^{\delta^{p,q}}$ is isomorphic to the Weyl group of the group

$$G^{\delta^{p,q}} = \mathrm{SO}(p+1, p)_0 \times \mathrm{SO}(q+1, q)_0.$$

The relevant $W^{\delta^{p,q}}$-types are a minimal set of irreducible representations of $W^{\delta^{p,q}}$ that detect nonunitarity for spherical Langlands quotients of $G^{\delta^{p,q}}$ (see (3-2)).

Theorem 12 [Barbasch 2004]. *The following is a set of relevant $W(B_k)$-types for the group* $\mathrm{SO}(k+1, k)_0$:

$$\{(k-m, m) \times (0) : 0 \le m \le [k/2]\} \cup \{(k-m) \times (m) : 0 \le m \le k\}. \tag{4-3}$$

Relevant $W(B_p) \times W(B_q)$-types of $\mathrm{SO}(p+1, p)_0 \times \mathrm{SO}(q+1, q)_0$ *are of the form*

$$\psi \otimes \mathrm{triv} \quad \text{or} \quad \mathrm{triv} \otimes \tau, \tag{4-4}$$

with ψ and τ a relevant W-type for $\mathrm{SO}(p+1, p)_0$ *and* $\mathrm{SO}(q+1, q)_0$, *respectively.*

The parametrization of $W(B_k)$-types in terms of pairs of partitions can be found, for example, in [Pantano et al. 2010, Section 9]. We give here just a short description of the $W(B_k)$-types we need. Recall that $W(B_k)$ is a semidirect product of the symmetric group S_k by the abelian normal subgroup $(\mathbb{Z}/2\mathbb{Z})^k$. The irreducible representations of S_k are parametrized by partitions of k, with the trivial partition (k) corresponding to the trivial representation of S_k. If $c + d = k$, with $d \le c$, then the symbol $(c, d) \times (0)$ (or simply (c, d)) denotes the pullback to $W(B_k)$ of the irreducible representation of S_k corresponding to the partition (c, d). This representation is a summand of dimension $\binom{k}{c}(k-2c+1)/(k-c+1)$ of the permutation module $M^{(c,d)} = \mathrm{Ind}_{S_c\times S_d}^{S_k}(\mathrm{triv})$, with $(\mathbb{Z}/2\mathbb{Z})^k$ acting trivially.

If a and b are nonnegative integers with $a+b=k$ then the symbol $(a)\times(b)$ denotes the irreducible representation of $W(B_k)$ of dimension $\binom{k}{a}$ induced from the one-dimensional representation of $(S_a\times S_b)\ltimes(\mathbb{Z}/2\mathbb{Z})^k$ in which S_a, S_b, and $(\mathbb{Z}/2\mathbb{Z})^a$ act trivially, and $(\mathbb{Z}/2\mathbb{Z})^b$ acts by sign. Here $(S_a\times S_b)$ is the stabilizer in S_k of the character $(\mathrm{triv})^a\otimes(\mathrm{sign})^b$ of $(\mathbb{Z}/2\mathbb{Z})^k$.

Theorem 9 asserts that, for every M-type $\delta^{p,q}$ of $\mathrm{SO}(n+1, n)$ and every relevant $W^{\delta^{p,q}}$-type ψ, there exists a K-type μ of level at most 2 such that

$$\psi_\mu \cong \psi.$$

We now describe the matching explicitly. We may restrict our attention to the case $p \geq q$. Indeed, for all choices of the parameters, we have

$$J(\delta^{q,p}, (\nu^q|\nu^p)) \simeq J(\delta^{p,q}, (\nu^p|\nu^q)) \otimes \chi, \tag{4-5}$$

with χ the nontrivial unitary character of $\mathrm{SO}(n+1, n)$.

The matching is presented in Table 1 (and will be proved in the next few sections).

Here we use partitions to parametrize some of the representations of $O(n)$ (and later of $O(n,\mathbb{C})$ and $\mathrm{GL}(n,\mathbb{C})$). See, for example, [Fulton and Harris 1991, Lectures 6 and 19]. For example, if the partition $\lambda = (\lambda_1, \ldots \lambda_k)$ has at most $n/2$ parts, it parametrizes a representation of $O(n)$ with highest weight $(\lambda_1, \ldots, \lambda_k, 0, \ldots, 0)$.

Remark 13. All K-types recorded in this table are level at most 2.

Proof. By definition, fine K-types have level at most 1; hence every irreducible constituent of the tensor product of two fine K-types is of level at most two. This

The relevant $W^{\delta^{p,q}}$-type ψ	A petite K-type μ such that $\psi_\mu = \psi$
$((p-k)\times(k))\otimes\mathrm{triv}$	$\Lambda^k(\mathbb{C}^{n+1})\otimes\Lambda^{q+k}(\mathbb{C}^n)$
$(p-k, k)\otimes\mathrm{triv}$	$\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\Big[\mathrm{triv}\otimes\Big(V^{O(n)}_{(2^k,1^q)}\Big)\Big]$
$\mathrm{triv}\otimes((q-k)\times(k))$	$\Lambda^k(\mathbb{C}^{n+1})\otimes\Lambda^{q-k}(\mathbb{C}^n)$
$\mathrm{triv}\otimes(q-k, k)$	$\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\Big[\mathrm{triv}\otimes\Big(V^{O(n)}_{(2^k,1^{q-2k})}\Big)\Big]$

Table 1. Matching.

implies, for example, that the K-types

$$\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\Big[\mathrm{triv}\otimes\Big(V^{O(n)}_{(2^k,1^{q-2k})}\Big)\Big]\subseteq\big[\mathrm{triv}\otimes\Lambda^k(\mathbb{C}^n)\big]\otimes\big[\mathrm{triv}\otimes\Lambda^{q-k}(\mathbb{C}^n)\big]$$

and

$$\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\Big[\mathrm{triv}\otimes\Big(V^{O(n)}_{(2^k,1^{q})}\Big)\Big]\subseteq\big[\mathrm{triv}\otimes\Lambda^k(\mathbb{C}^n)\big]\otimes\big[\mathrm{triv}\otimes\Lambda^{q+k}(\mathbb{C}^n)\big]$$

are level at most 2. Every K-type of the form

$$\Lambda^a(\mathbb{C}^{n+1})\otimes\Lambda^b(\mathbb{C}^n)$$

is also level at most 2, since Z_{ϵ_k} acts trivially on $\Lambda^b(\mathbb{C}^n)$ and it acts on $\Lambda^a(\mathbb{C}^{n+1})$ with eigenvalues $0, \pm 2i$ and $Z_{\epsilon_i\pm\epsilon_j}$ acts on both $\Lambda^b(\mathbb{C}^n)$ and $\Lambda^a(\mathbb{C}^{n+1})$ with eigenvalues $0, \pm i$. $\square$

4.1. ***The $W(B_q)$-type $(q-k)\times(k)$.*** Consider the K-type

$$\mu := \Lambda^k(\mathbb{C}^{n+1})\otimes\Lambda^{q-k}(\mathbb{C}^n), \tag{4-6}$$

where $\mathbb{C}^{n+1}$ and $\mathbb{C}^n$ are the standard representations of $O(n+1)$ and $O(n)$, respectively, with bases $\{e_1, e_2, \dots, e_{n+1}\}$ and $\{v_1, v_2, \dots, v_n\}$.

The restriction of μ to M contains the M-type $\delta^{p,q}$ with multiplicity $\binom{q}{k}$. The $\delta^{p,q}$-isotypic component inside μ is spanned by the vectors:

$$w_J := e_{q+2-i_k}\wedge e_{q+2-i_{k-1}}\wedge\cdots\wedge e_{q+2-i_1}\otimes v_{p+j_1}\wedge v_{p+j_2}\wedge\cdots\wedge v_{p+j_{q-k}}, \tag{4-7}$$

where $J=\{1\le j_1<j_2<\cdots<j_{q-k}\le q\}$ is a subset of $\{1, 2, \dots, q\}$ of cardinality $q-k$, and $I=\{1\le i_1<i_2<\cdots<i_k\le q\}$ is the complement of J in the same set. Indeed one can check that

$$T(t_1, t_2, \dots, t_n)\cdot w_J = \delta^{p,q}(T(t_1, t_2, \dots, t_n))\cdot w_J \tag{4-8}$$

for all $T(t_1, t_2, \dots, t_n)$ in M, and that this is the entire $\delta^{p,q}$-isotypic subspace of μ.

We study the representation ψ_μ of $W^{\delta^{p,q}}$ on the $\binom{q}{k}$-dimensional space

$$V_\mu[\delta^{p,q}] = \mathrm{Hom}_M(\mu, \delta^{p,q}).$$

For each $L=\{l_1<l_2<\cdots<l_{q-k}\}\subset\{1, 2, \dots, q\}$, set

$$T_L(w_J) = \begin{cases} u & \text{if } J = L, \\ 0 & \text{otherwise.}\end{cases} \tag{4-9}$$

Recall that the vector $u = v_{p+1}\wedge v_{p+2}\wedge\cdots\wedge v_n$ is a basis for the $\delta^{p,q}$-isotypic inside $\mu_{\delta^{p,q}}$. The maps $\{T_L\}$ form a basis of $V_\mu[\delta^{p,q}]$. Note that

$$\sigma_{\epsilon_l}\cdot v_s = +v_s \quad\text{and}\quad \sigma_{\epsilon_l}\cdot e_t = \begin{cases} -e_t & \text{if } t=1 \text{ or } t=n+2-l, \\ +e_t & \text{otherwise,}\end{cases} \tag{4-10}$$

for all $l, s = 1, \dots, n$ and $t = 1, \dots, n+1$. Hence:

- $\mu_{\delta^{p,q}}(\sigma_{\epsilon_l})u = u$.
- $\mu(\sigma_{\epsilon_l})w_J = \mu(\sigma_{\epsilon_l}^{-1})w_J = \begin{cases} -w_J & \text{if } l \in p + \{1, 2, \dots, q\} \setminus J, \\ +w_J & \text{otherwise,} \end{cases}$

and consequently

$$\psi_\mu(s_{\epsilon_l})T_J = \begin{cases} -T_J & \text{if } l \in p + \{1, 2, \dots, q\} \setminus J, \\ +T_J & \text{otherwise.} \end{cases} \tag{4-11}$$

Similarly we observe that

$$\sigma_{\epsilon_l - \epsilon_{l+1}} \cdot v_s = \begin{cases} v_{s+1} & \text{if } s = l, \\ -v_{s-1} & \text{if } s = l+1, \\ v_s & \text{if } s \neq l, l+1, \end{cases} \tag{4-12}$$

and

$$\sigma_{\epsilon_l - \epsilon_{l+1}} \cdot e_t = \begin{cases} +e_{t-1} & \text{if } t = n+2-l, \\ -e_{t+1} & \text{if } t = n+2-(l+1), \\ +e_t & \text{if } t \neq n+2-l, n+2-(l+1) \end{cases} \tag{4-13}$$

for $s = 1, \dots, n$, $t = 1, \dots, n+1$ and $l = 1, \dots, p-1$ or $l = p+1, \dots, n-1$. Hence:

- $\mu_{\delta^{p,q}}(\sigma_{\epsilon_l - \epsilon_{l+1}})u = u$.
- $\mu(\sigma_{\epsilon_l - \epsilon_{l+1}})w_J = \begin{cases} +w_J & \text{if } l < p, \\ +w_J & \text{if } l > p \text{ and either } l, l+1 \in p+J \\ & \quad \text{or } l, l+1 \in p+\{1, 2, \dots, q\} \setminus J, \\ -w_{J \triangle \{l-p, l+1-p\}} & \text{otherwise,} \end{cases}$

and since $\mu(\sigma_{\epsilon_l - \epsilon_{l+1}})w_J = \mu(\sigma_{\epsilon_l - \epsilon_{l+1}}^{-1})w_J$,

$$\psi_\mu(\sigma_{\epsilon_l - \epsilon_{l+1}})T_J = \begin{cases} +T_J & \text{if } l < p, \\ +T_J & \text{if } l > p \text{ and either } l, l+1 \in p+J \\ & \quad \text{or } l, l+1 \in p+\{1, 2, \dots, q\} \setminus J, \\ -T_{J \triangle \{l-p, l+1-p\}} & \text{otherwise.} \end{cases} \tag{4-14}$$

This information is enough to characterize the representation ψ_μ of $W^{\delta^{p,q}}$ on $V_\mu[\delta^{p,q}]$. Recall that $W(B_p)$ and $W(B_q)$ are realized as the (appropriate) subgroups of $W(B_n)$ acting on the coordinates $\{1, \dots, p\}$ and $\{p+1, \dots, n\}$, respectively. Let S_{q-k} and $(\mathbb{Z}/2\mathbb{Z})^{q-k}$ be the subgroups of $W(B_q)$ acting on the first $q-k$ coordinates $\{p+1, p+2, \dots, n-k\}$, and let S_k, $(\mathbb{Z}/2\mathbb{Z})^k$ be the ones acting on the last k coordinates $\{n-k+1, n-k+2, \dots, n\}$. Here the $\mathbb{Z}/2\mathbb{Z}$ factors are generated by the σ_{ϵ_l}, and the symmetric groups by the $\sigma_{\epsilon_l - \epsilon_{l+1}}$. Note that:

- The restriction of ψ_μ to $W(B_p)$ is trivial, hence ψ_μ is of the form $\mathrm{triv}\otimes\psi'$ for some representation ψ' of $W(B_q)$.
- The groups S_{q-k}, S_k and $(\mathbb{Z}/2\mathbb{Z})^{q-k}$ act trivially on the vector $T_{\{1,\dots,q-k\}}$, and $(\mathbb{Z}/2\mathbb{Z})^k$ acts on it by sign. Hence the restriction of ψ' to the group $(S_{q-k}\times S_k)\ltimes(\mathbb{Z}/2\mathbb{Z})^q$ contains the one-dimensional representation

$$[(q-k)\otimes(k)]\cdot[(\mathrm{triv})^{q-k}\otimes(\mathrm{sign})^k]. \tag{4-15}$$

By Frobenius reciprocity, ψ' contains the irreducible representation

$$(q-k)\times(k) := \mathrm{Ind}_{(S_{q-k}\times S_k)\ltimes(\mathbb{Z}/2\mathbb{Z})^q}^{W(B_q)}[(q-k)\otimes(k)]\cdot[(\mathrm{triv})^{q-k}\otimes(\mathrm{sign})^k]. \tag{4-16}$$

Actually $\psi' = (q-k)\times(k)$ for dimensional reasons.

We conclude that $\psi_\mu = \mathrm{triv}\otimes((q-k)\times(k))$, as claimed.

4.2. *The $W(B_q)$-type $(q-k,k)$.* Consider the (possibly reducible) representation

$$\mu := \mathrm{triv}\otimes[\Lambda^k(\mathbb{C}^n)\otimes\Lambda^{q-k}(\mathbb{C}^n)] \tag{4-17}$$

of K, where triv denotes the trivial representation of $O(n+1)$ and $\mathbb{C}^n$ denotes the standard representation of $O(n)$ (with basis $\{v_1, v_2, \dots, v_n\}$).

Note that μ contains the M-type $\delta^{p,q}$ with multiplicity $\binom{q}{k}$. For every subset $J=\{1\le j_1<j_2<\cdots<j_{q-k}\le q\}$, let $I=\{i_1<i_2<\cdots<i_k\}$ be its complement in the set $\{1,2,\dots,q\}$, and let

$$w_J := v_{p+i_1}\wedge v_{p+i_2}\wedge\cdots\wedge v_{p+i_k}\otimes v_{p+j_1}\wedge v_{p+j_2}\wedge\cdots\wedge v_{p+j_{q-k}}. \tag{4-18}$$

Then the vectors $\{w_J\}$ span the $\delta^{p,q}$-isotypic component inside μ.

We study the representation ψ_μ of $W^{\delta^{p,q}}$ on the $\binom{q}{k}$-dimensional space $V_\mu[\delta^{p,q}]$. For all $L=\{1\le l_1<l_2<\cdots<l_{q-k}\le q\}$, set

$$T_L(w_J) = \begin{cases} u & \text{if } J=L,\\ 0 & \text{otherwise.}\end{cases} \tag{4-19}$$

Then the maps $\{T_L\}$ form a basis of $V_\mu[\delta^{p,q}]$. Note that

$$\sigma_{\epsilon_l}\cdot v_s = +v_s \quad \text{for all } l,s=1,\dots,n; \tag{4-20}$$

hence

$$\mu_{\delta^{p,q}}(\sigma_{\epsilon_l})u = u \quad\text{and}\quad \mu(\sigma_{\epsilon_l})w_J = w_J \quad \text{for all } J. \tag{4-21}$$

This implies that

$$\psi_\mu(s_{\epsilon_l})T_J = T_J \quad \text{for all } J, \tag{4-22}$$

so ψ_μ is really a representation of $S_p\times S_q$. Next we show that S_p acts trivially.

Recall from (4-12) that for all $s = 1, \dots, n$ and all $l = 1, \dots, p-1$ or $l = p+1, \dots, n-1$, we have

$$\sigma_{\epsilon_l - \epsilon_{l+1}} \cdot v_s = \begin{cases} v_{s+1} & \text{if } s = l, \\ -v_{s-1} & \text{if } s = l+1, \\ v_s & \text{if } s \neq l, l+1. \end{cases} \tag{4-23}$$

Hence:

- $\mu_{\delta^{p,q}}(\sigma_{\epsilon_l - \epsilon_{l+1}})u = u$.
- $\mu(\sigma_{\epsilon_l - \epsilon_{l+1}})w_J = \begin{cases} +w_J & \text{if } l < p, \\ +w_J & \text{if } l > p \text{ and either } l, l+1 \in p+J \\ & \quad \text{or } l, l+1 \in p + \{1, 2, \dots, q\} \setminus J, \\ -w_{J \triangle \{l-p, l+1-p\}} & \text{otherwise,} \end{cases}$

and since $\mu(\sigma_{\epsilon_l - \epsilon_{l+1}})w_J = \mu(\sigma^{-1}_{\epsilon_l - \epsilon_{l+1}})w_J$,

$$\psi_\mu(\sigma_{\epsilon_l - \epsilon_{l+1}})T_J = \begin{cases} +T_J & \text{if } l < p, \\ +T_J & \text{if } l > p \text{ and either } l, l+1 \in p+J \\ & \quad \text{or } l, l+1 \in p + \{1, 2, \dots, q\} \setminus J, \\ -T_{J \triangle \{l-p, l+1-p\}} & \text{otherwise.} \end{cases} \tag{4-24}$$

Therefore S_p acts indeed trivially, so ψ_μ is of the form $\operatorname{triv} \otimes \psi'$ for some representation ψ' of S_q. Finally we prove that ψ' equals the permutation module $\operatorname{Ind}_{S_{q-k} \times S_k}^{S_q} \operatorname{triv} = M^{(q-k,k)}$. Write

$$\operatorname{Hom}_M(\mu, \delta^{p,q}) = \bigoplus_{\substack{L \subseteq \{1,\dots,q\} \\ |L| = q-k}} U_L, \tag{4-25}$$

with $U_L := \mathbb{C}T_L$. The symmetric group S_q permutes the subspaces U_L transitively. Set $L_0 := \{1, 2, \dots, q-k\}$ and $U_0 := U_{L_0}$, and let H be the stabilizer of U_0 in S_q (i.e., the set of all η in S_q such that $\eta U_0 = U_0$). Note that

- $H \simeq S_{q-k} \times S_k$. (We identify S_{q-k} and S_k with the subgroups of S_q acting on the first $q-k$ coordinates and the last k coordinates, respectively.)
- U_0 is stable under H and carries the trivial representation of H.
- The S_q-module $V_\mu[\delta^{p,q}]$ is induced from the H-module U_0 (see [Serre 1977, Proposition 19]).

Therefore

$$\psi' = \operatorname{Ind}_{S_{q-k} \times S_k}^{S_q} \operatorname{triv} = M^{(q-k,k)}, \tag{4-26}$$

and $\psi_\mu = \operatorname{triv} \otimes M^{(q-k,k)}$.

The representation $\mathrm{triv} \otimes M^{(q-k,k)}$ is clearly reducible because the permutation module $M^{(q-k,k)}$ of S_q decomposes as a direct sum of Specht modules

$$M^{(q-k,k)} = \bigoplus_{\lambda \trianglerighteq (q-k,k)} \mathcal{S}^{\lambda} = \bigoplus_{a=0}^{k} \mathcal{S}^{(q-a,a)}. \tag{4-27}$$

Here $\trianglerighteq$ denotes the dominance (partial) ordering on partitions: If $\lambda = (\lambda_1, \lambda_2, \dots)$ and $\mu = (\mu_1, \mu_2, \dots)$ are partitions of N then $\lambda \trianglerighteq \mu$ if $\sum_{i=1}^{r} \lambda_i \geq \sum_{i=1}^{r} \mu_i$ for all r. With abuse of notation, we let $(q-0,0)$ denote the trivial partition. Then

$$\mathrm{triv} \otimes M^{(q-k,k)} = \bigoplus_{a=0}^{k} \mathrm{triv} \otimes (q-a,a). \tag{4-28}$$

The module $\psi_\mu = V_\mu[\delta^{p,q}]$ is also reducible because the representation μ of K decomposes as a direct sum of K-types. We need to identify the irreducible component of μ containing $(q-k,k)$.

The first task is to compute the decomposition of the tensor product

$$\Lambda^k(\mathbb{C}^n) \otimes \Lambda^{q-k}(\mathbb{C}^n)$$

into $O(n)$-types. We do this in three steps: First we decompose $\Lambda^k(\mathbb{C}^n) \otimes \Lambda^{q-k}(\mathbb{C}^n)$ into irreducible representations of $\mathrm{GL}(n, \mathbb{C})$. Then we decompose each $\mathrm{GL}(n, \mathbb{C})$-type occurring in such a decomposition as a direct sum of $O(n, \mathbb{C})$-types. Finally we restrict to $O(n)$.

We will use Weyl's construction of irreducible representations of $\mathrm{GL}(n, \mathbb{C})$ and $O(n, \mathbb{C})$ (see, for example, [Fulton and Harris 1991, Lectures 6 and 19]). By Pieri's formula,

$$\Lambda^k(\mathbb{C}^n) \otimes \Lambda^{q-k}(\mathbb{C}^n) = V^{\mathrm{GL}(n)}_{(1^k)} \otimes V^{\mathrm{GL}(n)}_{(1^{q-k})} = \bigoplus_{a=0}^{k} V^{\mathrm{GL}(n)}_{(2^a, 1^{q-2a})}. \tag{4-29}$$

Note that the partitions $(2^a, 1^{q-2a})$ have at most

$$a + (q-2a) = q - a \leq q \leq n/2 \tag{4-30}$$

parts, so we can apply Littlewood's restriction formula [1944]:

$$\begin{aligned} \mathrm{Res}^{\mathrm{GL}(n,\mathbb{C})}_{O(n,\mathbb{C})} V^{\mathrm{GL}(n)}_{(2^a,1^{q-2a})} &= \bigoplus_{\nu : \#\mathrm{parts} \leq n/2} \Bigg(\sum_{\xi : \text{ even parts}} N_{\nu,\xi,(2^a,1^{q-2a})} \Bigg) V^{O(n)}_{\nu} \\ &= \bigoplus_{b=0}^{a} \underbrace{N_{(2^b,1^{q-2a}),(2^{a-b}),(2^a,1^{q-2a})}}_{=1} V^{O(n)}_{(2^b,1^{q-2a})} = \bigoplus_{b=0}^{a} V^{O(n)}_{(2^b,1^{q-2a})}. \end{aligned} \tag{4-31}$$

Here the $N_{\nu,\xi,\lambda}$ are the Littlewood–Richardson numbers. We deduce that

$$\Lambda^k(\mathbb{C}^n)\otimes\Lambda^{q-k}(\mathbb{C}^n)=\bigoplus_{a=0}^{k}\Bigl(\bigoplus_{b=0}^{a}V^{O(n)}_{(2^b,1^{q-2a})}\Bigr) \tag{4-32}$$

(as a representation of $O(n,\mathbb{C})$), and

$$\begin{aligned}\mu:=&\ \mathrm{triv}\otimes\bigl[\Lambda^k(\mathbb{C}^n)\otimes\Lambda^{q-k}(\mathbb{C}^n)\bigr]\\ =&\bigoplus_{a=0}^{k}\Bigl(\bigoplus_{b=0}^{a}\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\mathrm{triv}\otimes V^{O(n)}_{(2^b,1^{q-2a})}\Bigr)\end{aligned} \tag{4-33}$$

(as a representation of $K=S(O(n+1)\times O(n))$). Note that the K-representation

$$\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\mathrm{triv}\otimes V^{O(n)}_{(2^b,1^{q-2a})} \tag{4-34}$$

embeds in the tensor product

$$\bigl[\mathrm{triv}\otimes\Lambda^b(\mathbb{C}^n)\bigr]\otimes\bigl[\mathrm{triv}\otimes\Lambda^{q-2a+b}(\mathbb{C}^n)\bigr]. \tag{4-35}$$

Every M-type $\delta^{r,s}$ appearing in this tensor product satisfies

$$s\le q-2a+2b\le q \tag{4-36}$$

hence $\delta^{r,s}\ne\delta^{p,q}$ for all $b\ne a$, and

$$\mathrm{Hom}_M\bigl(\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\mathrm{triv}\otimes V^{O(n)}_{(2^b,1^{q-2a})},\delta^{p,q}\bigr)=\{0\}\quad\text{for all }\ b\ne a. \tag{4-37}$$

It follows that, as a representation of $W^{\delta^{p,q}}$,

$$\begin{aligned}\psi_\mu=&\ \mathrm{Hom}_M(\mu,\delta^{p,q})\\ =&\bigoplus_{a=0}^{k}\Bigl(\bigoplus_{b=0}^{a}\mathrm{Hom}_M\bigl(\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\mathrm{triv}\otimes V^{O(n)}_{(2^b,1^{q-2a})},\delta^{p,q}\bigr)\Bigr)\\ =&\bigoplus_{a=0}^{k}\mathrm{Hom}_M\bigl(\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\mathrm{triv}\otimes V^{O(n)}_{(2^a,1^{q-2a})},\delta^{p,q}\bigr).\end{aligned} \tag{4-38}$$

We also know that

$$\psi_\mu=\bigoplus_{a=0}^{k}\mathrm{triv}\otimes(q-a,a). \tag{4-39}$$

Equations (4-38) and (4-39) hold for all $k=0,\dots,[q/2]$. A simple induction argument shows that

$$\mathrm{Hom}_M\bigl(\mathrm{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))}\mathrm{triv}\otimes V^{O(n)}_{(2^a,1^{q-2a})},\delta^{p,q}\bigr)=\mathrm{triv}\otimes(q-a,a), \tag{4-40}$$

for all $a = 1 \dots k$. In particular,

$$\operatorname{Hom}_M\big(\operatorname{Res}^{O(n+1,\mathbb{C})\times O(n,\mathbb{C})}_{S(O(n+1)\times O(n))} \operatorname{triv}\otimes V^{O(n)}_{(2^k,1^{q-2k})}, \delta^{p,q}\big) = \operatorname{triv}\otimes(q-k,k). \tag{4-41}$$

4.3. *The $W(B_p)$-types $(p-k, k)$ and $(p-k)\times(k)$.* The calculations are similar to the ones done in Sections 4.2 and 4.1, respectively. We leave the details to the diligent reader.

5. Some small rank examples

In [Pantano et al. 2010], we proved Theorem 5 for metaplectic groups, except for the case $(p, q) = (2, 2)$ and $\nu^p = \nu^q = (\frac{3}{2}, \frac{1}{2})$. In this case, $J_{\mathrm{Mp}(8)}(\delta^{2,2}, \nu)$ coincides with the lowest K-type constituent of a weakly fair $A_{\mathfrak{q}}(\lambda)$ module with $\mathfrak{q} = \mathfrak{l}+\mathfrak{u}$, where $L \cong \widetilde{U}(2, 2)$ and $\lambda = -\rho(\mathfrak{u})$. As a constituent of an $A_{\mathfrak{q}}(\lambda)$ module in the weakly fair range, it is unitary by a result of Vogan [1993] (see [Knapp and Vogan 1995, Theorem 0.54]). The unitarity of this representation will also follow from Theorem 6 since ν is a unipotent parameter (see Definition 27). This completes the proof of Theorem 5 for metaplectic groups.

In this section, we prove Theorem 5 for orthogonal groups as well. The proof is analogous to the one for the metaplectic groups.

For nonnegative integers (p, q) such that $p+q \le 4$, and parameters $\nu = (\nu^p|\nu^q)$ with $\nu^p \in \mathrm{CS}(\mathrm{SO}(p+1, p)_0, \delta_0)$ and $\nu^q \in \mathrm{CS}(\mathrm{SO}(q+1, q)_0, \delta_0)$, we need to show that $J(\delta^{p,q}, \nu)$ is unitary.

The main tool for the proof is unitary induction with deformation of parameters. In all cases except one, we think of $J(\delta^{p,q}, \nu)$ as an irreducible subquotient of an induced representation

$$I(\nu^q) = \operatorname{Ind}^G_{P_I}\big(J(\delta^{p,0}, \nu^p)\otimes\delta^{0,q}\otimes\nu^q\big), \tag{5-1}$$

where $P_I = M_I A_I N_I$ is a parabolic subgroup of G with Levi factor

$$M_I A_I \cong \mathrm{SO}(p+1, p)\times \mathrm{GL}(1, \mathbb{R})^q.$$

Here $\delta^{0,q}$ is the product of q sign characters on the $\mathbb{Z}/2\mathbb{Z}$ parts of $\mathrm{GL}(1, \mathbb{R})$. The idea is to irreducibly deform the parameter ν^q to 0. Since $J(\delta^{p,0}, \nu^p)$ is unitary, and the Hermitian form on the induced representation can change signature only at reducibility points, this will prove that $J(\delta^{p,q}, \nu)$ is unitary. For this, we need to know when an induced representation stays irreducible under deformation.

Let $P = MAN$ be a minimal parabolic subgroup of $\mathrm{SO}(n+1, n)$, and let $P_I = M_I A_I N_I$ be a parabolic subgroup containing P. Then

$$P\cap M_I = MA_M N_M \tag{5-2}$$

is a minimal parabolic subgroup of M_I, and $A = A_M A_I$. For each pair of (real) characters $\nu_M \in \mathfrak{a}^*_{M,\mathbb{R}}$ and $\nu_I \in \mathfrak{a}^*_{I,\mathbb{R}}$, write $\nu = (\nu_M|\nu_I)$ for the corresponding character of A. Let δ be a character of M, and let $J_{M_I}(\delta, \nu_M)$ be the Langlands subquotient of the principal series

$$\mathrm{Ind}_{M A_M N_M}^{M_I}(\delta \otimes \nu_M). \tag{5-3}$$

Note that $J_{M_I}(\delta, \nu_M)$ is always irreducible.

Recall the notion of good and bad roots from page 484 (see (2-16)).

Proposition 14. *Consider the induced representation*

$$I(\nu_I) := \mathrm{Ind}_{M_I A_I N_I}^{\mathrm{SO}(n+1,n)}\big(J_{M_I}(\delta, \nu_M) \otimes \nu_I\big) \tag{5-4}$$

of $\mathrm{SO}(n+1, n)$. *Set* $\nu = (\nu_M|\nu_I)$, *and assume that* ν_I *satisfies*

$$\begin{cases} \langle \nu, \beta^\vee \rangle \notin 2\mathbb{Z}+1 & \textit{for all } \beta \in \Delta(\mathfrak{n}_I) \textit{ that are good for } \delta, \\ \langle \nu, \beta^\vee \rangle \notin 2\mathbb{Z}-\{0\} & \textit{for all } \beta \in \Delta(\mathfrak{n}_I) \textit{ that are bad for } \delta. \end{cases} \tag{5-5}$$

Then $I(\nu_I)$ *is irreducible.*

Proof. We claim that under the conditions (5-5), the operator $T(w_0, \mu, \delta^{p,q}, \nu)$ (see (3-3)) has no zero eigenvalues for any K-type μ containing $\delta^{p,q}$ and hence is invertible. This easily follows from [Barbasch et al. 2008, Theorem 2.10]. Then, as for the corresponding result for the metaplectic group [Pantano et al. 2010, Proposition 8.9], the proposition follows from (the proof of) [Knapp and Zuckerman 1977, Theorem 8] and [Speh and Vogan 1980, Corollary 3.9]. □

As for $\mathrm{Mp}(2n)$ in [Pantano et al. 2010], we have the following consequence.

Corollary 15. *In the setting of Proposition 14, let* $J(\nu_I)$ *be the (irreducible) Langlands subquotient of* $I(\nu_I)$. *Let* $R \subset \mathfrak{a}^*_{I,\mathbb{R}}$ *be any connected region in the complement of the hyperplane arrangement defined in (5-5). If* $J(\nu_I)$ *is unitary for some value of* ν_I *in* R, *then* $J(\nu)$ *is unitary throughout the closure of* R. *In particular, for* ν *in the unit cube (i.e., if* $0 \le |\nu_i| \le \frac{1}{2}$ *for* $i = 1 \ldots n$), $J(\nu)$ *is unitary.*

Theorem 16. *Let* $G = \mathrm{SO}(n+1, n)$. *Choose* (p, q) *with* $p + q = n$, *and let* $\nu = (\nu^p|\nu^q)$, *with* $\nu^p = (a_1, \ldots, a_p)$ *and* $\nu^q = (a_{p+1}, \ldots, a_n)$, *be a character of* A *such that*

$$\nu^p \in \mathrm{CS}(\mathrm{SO}(p+1, p)_0, \delta_0) \quad \textit{and} \quad \nu^q \in \mathrm{CS}(\mathrm{SO}(q+1, q)_0, \delta_0). \tag{5-6}$$

Suppose that

- *For all* $j = 1, \ldots, p$, *either* $0 \le |a_j| \le 3/2$ *or* $a_j \in \mathbb{Z} + \frac{1}{2}$.
- *For all* $j = p+1, \ldots, n$, $0 \le |a_j| \le \frac{1}{2}$ *(i.e.,* ν^q *belongs to the unit cube).*

Then the Langlands subquotient $J(\delta^{p,q}, \nu)$ *of* $\mathrm{SO}(n+1, n)$ *is unitary.*

Proof. The proof is completely analogous to the proof of [Pantano et al. 2010, Theorem 8.11]. If $P_I = M_I A_I N_I$ and $I(\nu^q)$ are as given in (5-1), then up to sign, the roots of N_I are

$$\begin{array}{ll} \epsilon_i & \text{for } p+1 \leq i \leq n, \\ \epsilon_i \pm \epsilon_j & \text{for } p+1 \leq i < j \leq n, \\ \epsilon_i \pm \epsilon_j & \text{for } 1 \leq i \leq p \text{ and } p+1 \leq j \leq n. \end{array}$$

The roots in the first two rows are good for $\delta^{p,q}$, the remaining ones are bad. Under our assumptions on ν^p, if ν^q is in the interior of the unit cube, then the conditions (5-5) are satisfied. By Proposition 14, $I(\nu^q)$ is irreducible. Moreover $J(0) = I(0)$ is unitary. By Corollary 15, $J(\delta^{p,q}, \nu)$ is unitary for all ν^q in the (closed) unit cube. □

The spherical complementary series parameters for $\mathrm{SO}(p+1, p)_0$ for $p = 1, 2, 3$ are given at the end of Section 8. It is easily checked that they all satisfy the hypotheses of the theorem, and for $q = 1, 2$, the spherical complementary series parameters all belong to the unit cube, except for the isolated point $(\frac{3}{2}, \frac{1}{2})$. Consequently Theorem 16 now implies the unitarity of all representations under consideration (for $n \leq 4$) except $J(\delta^{2,2}, \nu)$, with $\nu^p = \nu^q = (\frac{3}{2}, \frac{1}{2})$. Just as for the corresponding representation of $\mathrm{Mp}(8)$, we can realize it as a constituent of an $A_{\mathfrak{q}}(\lambda)$-module at the edge of the weakly fair range, which proves that it is unitary. See [Knapp and Vogan 1995, Example 3, Chapter VIII, §5] for a detailed discussion of this (reducible) module of $\mathrm{SO}(5, 4)$. Note that its unitarity also follows from Theorem 29.

Collecting all these results, we have now proved Theorem 5.

6. The theta correspondence

For the case $p = n$, $q = 0$, the authors of [Adams et al. 2007] use the theta correspondence to prove that if ν is a spherical unitary parameter for $\mathrm{SO}(n+1, n)$ then $\nu \in \mathrm{CS}(\mathrm{Mp}(2n), \delta^{n,0})$. We will generalize the argument to relate the complementary series of the two families of groups to each other in more generality.

First we collect the facts about the theta correspondence that we need (some of them were already recalled in [Pantano et al. 2010]).

Let (G, G') be a reductive dual pair in $\mathrm{Sp}(2N, \mathbb{R})$, that is, G and G' are reductive subgroups of $\mathrm{Sp}(2N, \mathbb{R})$ which are mutual centralizers. Write $\widetilde{G}$ and $\widetilde{G}'$ for the preimages of G and G' in $\mathrm{Mp}(2N)$ under the covering map. Howe [1989] defines a correspondence between irreducible representations of $\widetilde{G}$ and those of $\widetilde{G}'$, and shows that this correspondence is a bijection between subsets of the genuine admissible duals of the two groups. Moreover subjugated to the correspondence is a bijection between K- and K'-types in the space of joint harmonics $\mathscr{H}$. Here K and K' are maximal compact subgroups of $\widetilde{G}$ and $\widetilde{G}'$, respectively. The K- and K'-types are assigned a degree and the correspondence satisfies the following property: If π corresponds to π' in the correspondence for the dual pair (G, G'), and μ is a K-type of minimal degree occurring in π, then μ occurs in $\mathscr{H}$, and corresponds to a K'-type μ' which occurs in and is of minimal degree in π'.

In general, the theta correspondence does not preserve unitarity. However we have preservation of unitarity in the stable range.

Definition 17. The dual pair $(O(r, s), \mathrm{Sp}(2n, \mathbb{R}))$ is said to be in the stable range with $O(r, s)$ the smaller member if $n \geq r+s$. It is in the stable range with $\mathrm{Sp}(2n, \mathbb{R})$ the smaller member if $\min\{r, s\} \geq 2n$.

Theorem 18 [Li 1989c]. *Suppose (G, G') is a dual pair in the stable range with G the smaller member. If π is an irreducible genuine unitary representation of $\widetilde{G}$, then π occurs in the correspondence for the dual pair and corresponds to a unitary representation of $\widetilde{G}'$.*

Theorem 19 [Li 1989a]. *Let G' be a reductive group which is a member of some Type I reductive dual pairs. Let π' be a unitary irreducible genuine representation of $\widetilde{G}'$, of low rank* (in the sense of [Howe 1982]). *Then there exist a unitary character ξ of $\widetilde{G}'$, a reductive group G such that (G, G') is a dual pair in the stable range with G the smaller member, and a unitary genuine representation π of $\widetilde{G}$ such that π corresponds to $\pi' \otimes \xi$.*

As explained in [Pantano et al. 2010, Section 8.1], the correspondence for dual pairs of the form $(\mathrm{Sp}(2n, \mathbb{R}), O(r, s))$ with $r+s$ odd can be regarded as a correspondence between genuine irreducible representations of $\mathrm{Mp}(2n)$ and irreducible representations of $O(r, s)$. This depends on some choices, which we make the same way as we did in [Pantano et al. 2010].

For principal series representations of $O(m+1, m)$, we use the following notation. For each pair of nonnegative integers (p, q) such that $p+q=m$, we write $I_{O(m+1,m)}(\delta^{p,q}, \nu)$ for the principal series representation of $O(m+1, m)$ with lowest $(O(m+1) \times O(m))$-type

$$(0, \dots, 0; +) \otimes (\underbrace{1, \dots, 1}_{j}, 0, \dots, 0; \epsilon),$$

with $(j, \epsilon) = (q, +)$ if $p \geq q$, and $(p, -)$ if $p < q$. The corresponding Langlands subquotient will be denoted $J_{O(m+1,m)}(\delta^{p,q}, \nu)$. We say that a parameter ν belongs to the complementary series $\mathrm{CS}(O(m+1, m), \delta^{p,q})$ if $J_{O(m+1,m)}(\delta^{p,q}, \nu)$ is unitarizable (as a representation of $O(m+1, m)$ or $\mathrm{SO}(m+1, m)$).

For each positive integer m, write ρ_m for the infinitesimal character of the trivial representation of $\mathrm{SO}(2m+1)$. If $\nu \in \mathbb{C}^n$, then $(\rho_m|\nu)$ denotes the $(m+n)$-tuple obtained by tacking the coordinates of ν onto ρ_m:

$$(\rho_m|\nu) = \left(m-\tfrac{1}{2}, m-\tfrac{3}{2}, \dots, \tfrac{3}{2}, \tfrac{1}{2}, \nu_1, \dots, \nu_n\right).$$

For dual pairs of the form under consideration, the theta correspondence gives rise to the following correspondence of infinitesimal characters.

Theorem 20 [Przebinda 1996]. *Let (G, G') be a reductive dual pair with G an odd orthogonal group, and G' a symplectic group, or vice versa. Let π and π' be representations of $\widetilde{G}$ and $\widetilde{G}'$, with infinitesimal characters γ and γ', respectively, which correspond to each other. Assume that the rank r of G is greater than or equal to the rank r' of G'. Then $\gamma = (\rho_{r-r'}|\gamma')$.*

We need two more results about the correspondence for symplectic-orthogonal pairs.

Theorem 21 [Adams and Barbasch 1998]. *Let $n \le m$, and write $m = n+k$. Let $p+q = n$, and let $\nu = (\nu^p|\nu^q) \in \mathbb{R}^n$. Let $(G(m), G'(n)) = (\mathrm{Sp}(2m,\mathbb{R}), O(n+1,n))$ or $(O(m+1,m), \mathrm{Sp}(2n,\mathbb{R}))$. In the correspondence for the dual pair $(G(m), G'(n))$,*

$$J_{G(m)}(\delta^{p+k,q}, (\rho_k|\nu)) \longleftrightarrow J_{G'(n)}(\delta^{p,q}, \nu).$$

Proposition 22 [Adams and Barbasch 1998]. *Suppose that π and π' map to each other in the theta correspondence for the dual pair $(G, G') = (\mathrm{Sp}(2n,\mathbb{R}), O(r,s))$ with $r+s$ odd. If $2n > r+s$, then the lowest $\widetilde{U}(n)$-types of π are of minimal degree in π. Similarly, if $2n < r+s$, then the lowest $(O(r)\times O(s))$-types of π' are of minimal degree in π'.*

Instead of recalling Howe's definition of the rank of a representation (see [Howe 1982]), we give a theorem which leads to an alternative definition of "low rank".

Theorem 23 [Li 1989b; 1997]. *Let G be $\mathrm{Mp}(2n)$ or $O(n+1,n)$, and let π be an irreducible unitary representation of G. Set*

$$r_G = \begin{cases} n & \text{if } G = \mathrm{Mp}(2n) \text{ or if } G = O(n+1,n) \text{ with } n \text{ even},\\ n-1 & \text{if } G = O(n+1,n) \text{ with } n \text{ odd}.\end{cases} \tag{6-1}$$

Let $WF(\pi) \subseteq \mathfrak{g}_0^* \simeq \mathfrak{g}_0$ denote the wave front set of π, and write $\mathrm{rank}(WF(\pi))$ for the maximal rank of the elements of $WF(\pi)$ as matrices. Then π is of "low rank" in the sense of Howe if and only if

$$\mathrm{rank}(WF(\pi)) < r_G.$$

Proposition 24. *Let (p, q) be a pair of nonnegative integers such that $p+q = n$, and let k be any integer satisfying $k \ge n+2$.*

(1) *For all parameters $\nu = (\nu^p|\nu^q)$, if*

$$\big((\rho_k|\nu^p)|\nu^q\big) \in \mathrm{CS}(\mathrm{SO}(n+k+1, n+k), \delta^{p+k,q}),$$

then $J_{O(n+k+1,n+k)}(\delta^{p+k,q}, ((\rho_k|\nu^p)|\nu^q))$ is of low rank.

(2) *For all parameters* $\nu = (\nu^p|\nu^q)$, *if*

$$\big((\rho_k|\nu^p)|\nu^q\big) \in \mathrm{CS}(\mathrm{Mp}(2(n+k)), \delta^{p+k,q}),$$

then $J_{\mathrm{Mp}(2(n+k))}(\delta^{p+k,q}, ((\rho_k|\nu^p)|\nu^q))$ *is of low rank.*

Proof. For part (1), note that for $q = 0$ this result already appears in [Adams et al. 2007, Fact 2, Section 14]; the same argument applied there goes over word for word if $q > 0$. The first step is to realize $J_{O(n+k+1,n+k)}(\delta^{p+k,q}, ((\rho_k|\nu^p)|\nu^q))$ as a composition factor of an induced representation

$$\mathrm{Ind}_L^{O(n+k+1,n+k)}(\mathrm{triv} \otimes \xi),$$

where $L = O(k+1, k) \times \mathrm{GL}(1, \mathbb{R})^n$, and ξ is a one-dimensional character of $\mathrm{GL}(1, \mathbb{R})^n$. The wave front set is then contained in the closure of the Richardson orbit for L, hence its rank is bounded above by the rank of that orbit. The same calculation done in [Adams et al. 2007] shows that this rank is strictly less than $n+k-1$. By Theorem 23, the representation $J_{O(n+k+1,n+k)}(\delta^{p+k,q}, ((\rho_k|\nu^p)|\nu^q))$ of $O(n+k+1, n+k)$ is of low rank.

For part (2), Theorem 21 implies that the Langlands quotient $J_{\mathrm{Mp}(2(n+k))}(\delta^{p+k,q}, ((\rho_k|\nu^p)|\nu^q))$ corresponds to $J_{O(n+1,n)}(\delta^{p,q}, \nu)$ in the correspondence for the pair $(\mathrm{Sp}(2(n+k), \mathbb{R}), O(n+1, n))$. By [Li 1989b, Proposition 1], if π is any representation of $\mathrm{Mp}(2(n+k))$ coming from the duality correspondence with some representation of $O(n+1, n)$, then the rank of the wave front set of π is at most $2n+1$. In particular,

$$\mathrm{rank}\big(WF(J_{\mathrm{Mp}(2(n+k))}(\delta^{p+k,q}, ((\rho_k|\nu^p)|\nu^q)))\big) \le 2n+1.$$

If $k > n+1$ (as in our assumptions), then this is strictly less than the split rank of $\mathrm{Mp}(2(n+k))$, hence the representation $J_{\mathrm{Mp}(2(n+k))}(\delta^{p+k,q}, (\rho_k|\nu))$ of $\mathrm{Mp}(2(n+k))$ is of low rank by Theorem 23. □

One consequence of Conjecture 1 would be that the (p, q)-complementary series of $\mathrm{SO}(n+1, n)$ and the genuine (p, q)-complementary series of $\mathrm{Mp}(2n)$ coincide. We would like to show that if ν is a unitary parameter for one of the groups (for a given choice of p and q), then it is unitary for the other group as well. The theta correspondence provides such an argument if we know that a closely related parameter ν' for a larger group of the same type is also unitary.

Theorem 25. (1) *Let* $\nu = (\nu^p|\nu^q) \in \mathrm{CS}(\mathrm{SO}(n+1, n), \delta^{p,q})$. *If* $(\rho_{n+2}|\nu)$ *is in* $\mathrm{CS}(\mathrm{SO}(2n+3, 2n+2), \delta^{p+n+2,q})$ *then* $\nu \in \mathrm{CS}(\mathrm{Mp}(2n), \delta^{p,q})$.

(2) *Let* $\nu = (\nu^p|\nu^q) \in \mathrm{CS}(\mathrm{Mp}(2n), \delta^{p,q})$. *If*

$$(\rho_{n+2}|\nu) \in \mathrm{CS}(\mathrm{Mp}(4n+4), \delta^{p+n+2,q})$$

then $\nu \in \mathrm{CS}(\mathrm{SO}(n+1, n), \delta^{p,q})$.

Proof. For part (1), assume $\nu = (\nu^p|\nu^q) \in \mathrm{CS}(\mathrm{SO}(n+1, n), \delta^{p,q})$, and $(\rho_{n+2}|\nu)$ is in $\mathrm{CS}(\mathrm{SO}(2n+3, 2n+2), \delta^{p+n+2,q})$. We want to show that $J_{\mathrm{Mp}(2n)}(\delta^{p,q}, \nu)$ is unitary. By our assumption, $\pi_s = J_{O(2n+3,2n+2)}(\delta^{p+n+2,q}, (\rho_{n+2}|\nu))$ is unitary. By Proposition 24 (1), π_s is of low rank. Theorem 19 implies that there are a character ξ of $O(2n+3, 2n+2)$, a group $\mathrm{Mp}(2m)$ with $m \le n+1$, and a unitary representation π of $\mathrm{Mp}(2m)$ such that $\pi \leftrightarrow \pi_s \otimes \xi$ in the theta correspondence for the dual pair $(\mathrm{Sp}(2m, \mathbb{R}), O(2n+3, 2n+2))$. By Proposition 22, the lowest K-type of $\pi_s \otimes \xi$ is of minimal degree for the dual pair, and must therefore occur in $\mathcal{H}$. By the explicit correspondence in $\mathcal{H}$ (see [Adams and Barbasch 1998, Proposition 2.1]), ξ must be trivial. If we can show that $m = n$ then we are done, since by Theorem 21, π_s corresponds to $J_{\mathrm{Mp}(2n)}(\delta^{p,q}, \nu)$ in the correspondence for the dual pair $(\mathrm{Sp}(2n, \mathbb{R}), O(2n+3, 2n+2))$, so that $J_{\mathrm{Mp}(2n)}(\delta^{p,q}, \nu)$ must be unitary.

So suppose $m > n$. Then $m = n+1$. Theorem 21 tells us that π_s corresponds to $J_{\mathrm{Mp}(2n+2)}(\delta^{p+1,q}, \nu')$ for this dual pair, where $\nu' = (n+\frac{3}{2}, \nu_1, \dots, \nu_n)$ (since it is obtained from $(\rho_{n+2}|\nu)$ by removing the coordinates of ρ_{n+1}). By [Pantano et al. 2010, Proposition 7.7] (the analog, for the metaplectic group, of Proposition 10), this is not a unitary parameter for any principal series of $\mathrm{Mp}(2n+2)$. So $m \le n$.

Now suppose that $m < n$, say $n = m+k$. By Theorem 20, the infinitesimal character of π is obtained from $(\rho_{n+2}|\nu)$ by removing the coordinates of $\rho_{2n+2-m} = \rho_{n+2+k}$. If $k > 0$ then this means that ν contains a coordinate $n+\frac{5}{2}$. This implies that one of the conditions of Proposition 10 holds, hence ν is not a unitary parameter. This contradicts our assumption, so we must have $m = n$, and so $\pi = J_{\mathrm{Mp}(2n)}(\delta^{p,q}, \nu)$ is unitary.

For part (2), let $\nu = (\nu^p|\nu^q) \in \mathrm{CS}(\mathrm{Mp}(2n), \delta^{p,q})$, and assume $(\rho_{n+2}|\nu)$ is in $\mathrm{CS}(\mathrm{Mp}(4n+4), \delta^{p+n+2,q})$. We want to show that $J_{O(n+1,n)}(\delta^{p,q}, \nu)$, and hence $J_{\mathrm{SO}(n+1,n)}(\delta^{p,q}, \nu)$, is unitary. By our assumption,

$$\pi_s = J_{\mathrm{Mp}(4n+4)}(\delta^{n+p+2,q}, (\rho_{n+2}|\nu))$$

is unitary. By Proposition 24 (2), π_s is of low rank. Theorem 19 implies that there are a character ξ of $\mathrm{Mp}(4n+4)$, a group $O(r, s)$ with $r+s \le 2n+2$ and $r+s$ odd, and a unitary representation π of $O(r, s)$ such that $\pi_s \otimes \xi \leftrightarrow \pi$ in the correspondence for the pair $(\mathrm{Sp}(4n+4, \mathbb{R}), O(r, s))$. The group $\mathrm{Mp}(4n+4)$ has no nontrivial characters, so ξ must be trivial. Also, by Proposition 22, the fine K-type of π_s occurs in $\mathcal{H}$. By the correspondence in $\mathcal{H}$, this can only happen when $r = s+1$. We want to show that $s = n$; then $\pi = J_{O(n+1,n)}(\delta^{p,q}, \nu)$ is unitary.

The condition $r+s \le 2n+2$ implies $2s+1 \le 2n+2$. By integrality, $s \le n$. Assume $s < n$. By Theorem 20 , the infinitesimal character of π is obtained from $(\rho_{n+2}|\nu)$ by removing the coordinates of ρ_{2n+2-s}. If $s < n$ then this contains a coordinate $n+\frac{5}{2}$, which must come from ν. Then ν satisfies one of the conditions of [Pantano et al. 2010, Proposition 7.7], hence is not a unitary parameter. Consequently $s = n$, and the theorem is proved. □

The following is a direct consequence of the description of the spherical unitary parameters for the split groups of type B (see Section 8).

Proposition 26 [Barbasch 2010]. *If* $\nu \in \mathrm{CS}(\mathrm{SO}(n+1, n), \delta_0)$, *then*

$$(\rho_m|\nu) \in \mathrm{CS}(\mathrm{SO}(n+m+1, n+m), \delta_0) \quad \textit{for all } m > 0.$$

Using this observation, Theorem 25 now immediately implies Theorem 4.

7. Unipotent representations

In this section, we identify the principal series parameters which are attached to special unipotent representations of $\mathrm{SO}(n+1, n)$, and discuss the unitarity of these modules.

Barbasch [2010] attaches parameters of spherical principal series representations of $\mathrm{SO}(n+1, n)_0$ to nilpotent orbits in $\mathfrak{sp}(2n, \mathbb{C})$. Fix a Cartan subalgebra $\mathfrak{h}^\vee$ of $\mathfrak{sp}(2n, \mathbb{C})$. This algebra is naturally isomorphic to $\mathfrak{a}^*$. Given a nilpotent orbit $\mathbb{O}^\vee$, let $\{e^\vee, h^\vee, f^\vee\}$ be an $\mathfrak{sl}_2$ triple with $f^\vee \in \mathbb{O}^\vee$ and $h^\vee \in \mathfrak{h}^\vee$. Then the corresponding spherical parameters are of the form

$$\nu = \frac{h^\vee}{2} + \gamma,$$

where $\gamma \in \mathfrak{z}(\{e^\vee, h^\vee, f^\vee\})$, the centralizer of the triple. A spherical parameter ν is called unipotent if it is of the form $\nu = h^\vee/2$ (see Definition 34). Since our (p, q)-principal series parameters are given by pairs of spherical parameters (for $\mathrm{SO}(p+1, p)_0$ and $\mathrm{SO}(q+1, q)_0$), we can attach them to pairs of nilpotent orbits analogously.

Definition 27. Fix nonnegative integers p and q such that $p+q = n$. A parameter $\nu = (\nu^p|\nu^q)$ for a (p, q)-principal series of $\mathrm{SO}(n+1, n)$ (or $\mathrm{Mp}(2n)$) is called unipotent if both ν^p and ν^q are spherical unipotent parameters for $\mathrm{SO}(p+1, p)_0$ and $\mathrm{SO}(q+1, q)_0$, respectively.

Proposition 28. *Given* p *and* q *such that* $p + q = n$, *a parameter* $\nu = (\nu^p|\nu^q)$ *is unipotent if and only if* $J(\delta^{p,q}, \nu)$ *is a special unipotent representation of* $\mathrm{SO}(n + 1, n)$.

Proof. Recall first that a Langlands parameter [Langlands 1970] for $G = \mathrm{SO}(n+1, n)$ is a conjugacy class by $\mathrm{Sp}(2n, \mathbb{C})$ of homomorphisms

$$\varphi : W_{\mathbb{R}} \longrightarrow \mathrm{Sp}(2n, \mathbb{C}) \times \Gamma,$$

satisfying certain conditions. Here $\Gamma = \{\gamma, 1\}$ is the Galois group of $\mathbb{C}$ over $\mathbb{R}$, and $W_{\mathbb{R}}$ is the Weil group of $\mathbb{R}$, the group generated by $\mathbb{C}^\times$ and an element τ, subject to the relations $\tau z \tau^{-1} = \bar{z}$ for $z \in \mathbb{C}^\times$, $\tau^2 = -1$. To each such parameter φ is associated an "L-packet", which is a finite set Π_φ of irreducible admissible representations of G. The assignment is explained in detail in [Borel 1979].

For principal series representations of $\mathrm{SO}(n+1, n)$, the L-packets are in fact singletons. Fix a maximal torus T of $\mathrm{Sp}(2n, \mathbb{C})$, and $\pi = J(\delta^{p,q}, \nu)$. Then the Langlands parameter φ of π can be chosen to be

$$\varphi(\tau) = \big((\underbrace{1, \dots, 1}_{p}, \underbrace{-1, \dots, -1}_{q}), \gamma\big) \in T \times \Gamma; \tag{7-1a}$$

$$\varphi(re^{i\theta}) = r^{2\nu} = \big((r^{2\nu_1}, \dots, r^{2\nu_n}), 1\big). \tag{7-1b}$$

Now recall from [Arthur 1989, §4] (see also [Adams et al. 1992, Chapters 22 and 26]) that an *Arthur parameter* for $G = \mathrm{SO}(n+1, n)$ is a conjugacy class by $\mathrm{Sp}(2n, \mathbb{C})$ of certain maps

$$\psi : W_{\mathbb{R}} \times \mathrm{SL}(2, \mathbb{C}) \longrightarrow \mathrm{Sp}(2n, \mathbb{C}) \times \Gamma,$$

whose restriction to $\mathrm{SL}(2, \mathbb{C})$ must be a holomorphic homomorphism into $\mathrm{Sp}(2n, \mathbb{C})$. To each Arthur parameter is attached a finite set of irreducible admissible representations, called an *Arthur packet*. If ψ is an Arthur parameter, then

$$\varphi_\psi : W_{\mathbb{R}} \longrightarrow \mathrm{Sp}(2n, \mathbb{C}) \times \Gamma, \tag{7-2}$$

defined by

$$\varphi_\psi(w) = \psi\left(w, \begin{pmatrix} |w|^{\frac{1}{2}} & 0 \\ 0 & |w|^{-\frac{1}{2}} \end{pmatrix}\right) \tag{7-3}$$

for all $w \in W_{\mathbb{R}}$, is a Langlands parameter. Here, if $w = z\tau$ then $|w| = |z|$.

An Arthur parameter ψ is called *unipotent* if

$$\psi(\mathbb{C}^\times) = \{(1, 1)\}. \tag{7-4}$$

We call representations which are contained in an L-packet Π_{φ_ψ} for ψ a unipotent parameter *special unipotent*. If the principal series representation $\pi = J(\delta^{p,q}, \nu)$ is special unipotent, then the corresponding Arthur parameter ψ must satisfy (7-4) and (7-1a), and the image of $\mathrm{SL}(2, \mathbb{C})$ must lie in the centralizer C of $\psi(\tau)$. This centralizer is isomorphic to $\mathrm{Sp}(2p, \mathbb{C}) \times \mathrm{Sp}(2q, \mathbb{C})$. It follows that our parameters are in 1-1 correspondence with C-orbits of holomorphic maps of $\mathrm{SL}(2, \mathbb{C})$ into C. The nontrivial orbits are given by C-conjugacy classes of embeddings of $\mathfrak{sl}(2, \mathbb{C})$ into $\mathfrak{sp}(2p, \mathbb{C}) \oplus \mathfrak{sp}(2q, \mathbb{C})$, and such classes of embeddings are in turn in 1-1 correspondence with nonzero nilpotent C-orbits on $\mathfrak{sp}(2p, \mathbb{C}) \oplus \mathfrak{sp}(2q, \mathbb{C})$, so that the set of Arthur parameters under consideration is indeed in 1-1 correspondence with pairs of nilpotent orbits as claimed. Unwinding the definitions, we see that given such a pair of orbits $(\mathbb{O}_p^\vee, \mathbb{O}_q^\vee)$, the corresponding Arthur parameter ψ determines a Langlands parameter φ_ψ which is the parameter of a (p, q)-principal series. Moreover

$$d\psi\left(\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}\right), \tag{7-5}$$

which, in the case of principal series, is the continuous parameter of the representations attached to φ_ψ, is then $(h_p^\vee/2, h_q^\vee/2)$, where $h_p^\vee$ and $h_q^\vee$ are the middle elements of $\mathfrak{sl}_2$ triples for $\mathbb{O}_p^\vee$ and $\mathbb{O}_q^\vee$, respectively. This completes the proof of our proposition. □

In his recent book, Arthur [2013] reformulates and proves several of the conjectures of [Arthur 1989]. In particular, for certain quasisplit classical groups including $\mathrm{SO}(n+1, n)$, he proves that all representations in certain Arthur packets are local components of automorphic representations, and therefore unitary (see Theorem 1.5). These include the Arthur packets attached to unipotent Arthur parameters. Moreover he proves (see Proposition 7.4.1) that these Arthur packets contain the L-packets that are attached to them. Consequently, we obtain the following result.

Theorem 29 [Arthur 2013]. *Special unipotent representations of* $\mathrm{SO}(n+1, n)$ *are unitary.*

Theorem 29, Proposition 28, part (1) of Theorem 25, and the observation that if ν is a unipotent parameter, then so is $(\rho_{n+2}|\nu)$, now easily imply Theorem 6.

8. The spherical unitary dual of $\mathrm{SO}(n+1, n)_0$.

In this section, we give an explicit description of the spherical unitary dual of split groups of type B. All the results are known, and due to D. Barbasch [2010; 2008]. See also [Pantano et al. 2010, Section 11] for a more detailed account.

Let $G = \mathrm{SO}(n+1, n)_0$, and let $\check{\mathfrak{g}} = \mathfrak{sp}(2n, \mathbb{C})$ be the complex dual Lie algebra, with Cartan subalgebra $\check{\mathfrak{h}}$. The spherical unitary dual of G is a disjoint union of sets, parametrized by nilpotent orbits in $\check{\mathfrak{g}}$. Recall that nilpotent orbits in $\mathfrak{sp}(2n, \mathbb{C})$ are parametrized by partitions of $2n$ in which every odd part occurs with even multiplicity.

Definition 30. Let ν be a real parameter in $\check{\mathfrak{h}}$, and let $\mathbb{O}$ be a nilpotent orbit in $\check{\mathfrak{g}}$. Let $\check{h} \in \check{\mathfrak{h}}$ be the middle element of an $\mathfrak{sl}(2)$ triple associated to $\mathbb{O}$. We say that ν is attached to $\mathbb{O}$ if

(1) $\nu = \check{h}/2 + \kappa$, for some semisimple element κ in the centralizer $\mathfrak{z}_{\check{\mathfrak{g}}}(\mathbb{O})$, and

(2) whenever $\mathbb{O}'$ is another nilpotent orbit in $\check{\mathfrak{g}}$ such that $\nu = \check{h}'/2 + \kappa'$, for some $\kappa' \in \mathfrak{z}_{\check{\mathfrak{g}}}(\check{\mathbb{O}}')$ semisimple, then $\mathbb{O}' \subset \bar{\mathbb{O}}$.

If ν is a real parameter in $\check{\mathfrak{h}}$, we can identify ν with an element of $\mathfrak{a}^*_{\mathbb{R}}$, and consider the irreducible spherical representation $J(\delta_0, \nu)$ of G.

Definition 31. A parameter $\nu \in \mathfrak{a}^*_{\mathbb{R}}$ is in the $\mathbb{O}$-complementary series if

(1) ν is attached to $\mathbb{O}$, and

(2) $J(\delta_0, \nu)$ is unitary.

The zero-complementary series, that is, the complementary series attached to the trivial nilpotent orbit plays a special role.

Theorem 32 [Barbasch 2010]. *For every nilpotent orbit $\mathbb{O}$ in $\check{\mathfrak{g}}$, let $G^0(\mathbb{O})$ be the connected real split group whose complex dual Lie algebra is $\mathfrak{z}_{\check{\mathfrak{g}}}(\mathbb{O})$. Let $\nu = \check{h}/2+\kappa$ be a parameter attached to the nilpotent orbit $\mathbb{O}$. Then ν is in the $\mathbb{O}$-complementary series for the group G if and only if κ is in the zero-complementary series for the group $G^0(\mathbb{O})$.*

The zero-complementary series of all real split groups is known, thanks to D. Barbasch. We recall the result for the groups we need.

Theorem 33 [Barbasch 2010]. *The zero-complementary series for split groups of type B_k, C_k and D_k consists of the following dominant parameters*:

B_k. *The set of all $\nu = (\kappa_1, \ldots, \kappa_k)$ such that $0 \le \kappa_1 \le \kappa_2 \le \cdots \le \kappa_k < \frac{1}{2}$.*

C_k. *The set of all $\nu = (\kappa_1, \ldots, \kappa_k)$ such that there exists an index $i = 2, \ldots, k$ with the property that*

$$0 \le \kappa_1 \le \cdots \le \kappa_i < 1 - \kappa_{i-1} < \kappa_{i+1} < \cdots < \kappa_k < 1,$$

and, for every $i \le j < k$, there is an odd number of κ_l with $1 \le l < i$ such that $\kappa_j < 1 - \kappa_l < \kappa_{j+1}$.

D_k. *Similar to type C_k. If k is even, replace every occurrence of κ_1 by $|\kappa_1|$. If k is odd, replace every occurrence of κ_1 by* 0.

Note that the choice of dominant parameters is not the standard one.

To compute $\mathfrak{z}_{\check{\mathfrak{g}}}(\mathbb{O})$, let λ be the partition corresponding to $\mathbb{O}$; denote the parts of λ by a_l, and their multiplicity by r_l (r_l is even if a_l is odd):

$$\lambda = (\underbrace{a_1, \ldots, a_1}_{r_1}, \ldots, \underbrace{a_m, \ldots, a_m}_{r_m}).$$

Then $\mathfrak{z}_{\check{\mathfrak{g}}}(\mathbb{O})$ is a product of symplectic and orthogonal Lie algebras. There is a factor $\mathfrak{sp}(r_l)$ for each odd part, and a factor $\mathfrak{so}(r_l)$ for each even part.

We describe the contribution of an odd part a of λ to ν; we refer the reader to the appendix of [Pantano et al. 2010] for the other cases. If $r_a = 2n_a$, the partition λ

contains n_a pairs of the form (a, a). The j-th pair (a, a) contributes a string

$$-\left(\frac{a-1}{2}\right), -\left(\frac{a-3}{2}\right), \dots, -1, 0, +1, \dots, +\left(\frac{a-3}{2}\right), +\left(\frac{a-1}{2}\right)$$

(of length a) to $\check{h}/2$, and a string $\left(\kappa_j^{(a)}, \kappa_j^{(a)}, \dots, \kappa_j^{(a)}\right)$ (also of length a) to κ. Moreover the part a contributes a factor $\mathfrak{sp}(2n_a)$ to the stabilizer of the orbit, and a factor $SO(n_a+1, n_a)_0$ to the group $G^0(\check{\mathbb{O}})$. For ν to be a unitary parameter, we impose the condition that the parameter $\left(\kappa_1^{(a)}, \dots, \kappa_{n_a}^{(a)}\right)$ belongs to the zero-complementary series for $SO(n_a+1, n_a)_0$ (see Theorem 33).

Definition 34 (Barbasch). A parameter $\nu = \check{h}/2 + \kappa$ is called *spherical unipotent* if $\kappa = 0$.

Finally we give an explicit list of the spherical unitary parameters (in the fundamental Weyl chamber, FWC) for $SO(n+1, n)_0$ with $n \leq 3$.

(1) For $n = 1$, the closed interval $[0, \frac{1}{2}]$.

(2) For $n = 2$,
   - The intersection of the unit cube with the FWC: $\{0 \leq \nu_2 \leq \nu_1 \leq \frac{1}{2}\}$,
   - The isolated point $(\frac{3}{2}, \frac{1}{2})$.

(3) For $n = 3$,
   - The intersection of the unit cube with the FWC: $\{0 \leq \nu_3 \leq \nu_2 \leq \nu_1 \leq \frac{1}{2}\}$.
   - The segment from $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ to $(1, \frac{1}{2}, 0)$:
   $$\left\{(\tfrac{1}{2}+t, \tfrac{1}{2}, \tfrac{1}{2}-t), \text{ for } 0 \leq t \leq \tfrac{1}{2}\right\}.$$
   - The segment from $(1, \frac{1}{2}, 0)$ to $(\frac{3}{2}, \frac{1}{2}, \frac{1}{2})$:
   $$\left\{(1+t, \tfrac{1}{2}, t), \text{ for } 0 \leq t \leq \tfrac{1}{2}\right\}.$$
   - The segment from $(1, 1, 0)$ to $(\frac{3}{2}, \frac{1}{2}, \frac{1}{2})$:
   $$\left\{(1+t, 1-t, t), \text{ for } 0 \leq t \leq \tfrac{1}{2}\right\}.$$
   - The segment from $(\frac{3}{2}, \frac{1}{2}, 0)$ to $(\frac{3}{2}, \frac{1}{2}, \frac{1}{2})$:
   $$\left\{(\tfrac{3}{2}, \tfrac{1}{2}, t), \text{ for } 0 \leq t \leq \tfrac{1}{2}\right\}.$$
   - The isolated point $(\frac{5}{2}, \frac{3}{2}, \frac{1}{2})$.

Acknowledgements

We thank Jeffrey Adams for explaining Arthur's results on unipotent representations and their relevance to our work. Moreover we are grateful to the referee for several valuable suggestions.

References

[Adams and Barbasch 1998] J. Adams and D. Barbasch, "Genuine representations of the metaplectic group", *Compositio Math.* **113**:1 (1998), 23–66. MR 99h:22013 Zbl 0913.11022

[Adams et al. 1992] J. Adams, D. Barbasch, and D. A. Vogan, Jr., *The Langlands classification and irreducible characters for real reductive groups*, Progress in Mathematics **104**, Birkhäuser, Boston, 1992. MR 93j:22001 Zbl 0756.22004

[Adams et al. 2007] J. Adams, D. Barbasch, A. Paul, P. E. Trapa, and D. A. Vogan, Jr., "Unitary Shimura correspondences for split real groups", *J. Amer. Math. Soc.* **20**:3 (2007), 701–751. MR 2008i:22008 Zbl 1114.22009

[Arthur 1989] J. G. Arthur, "Unipotent automorphic representations: conjectures", pp. 13–71 in *Orbites unipotentes et représentations, II: Groupes p-adiques et réels*, edited by M. Andler, Astérisque **171-172**, Société Mathématique de France, Paris, 1989. MR 91f:22030 Zbl 0728.22014

[Arthur 2013] J. G. Arthur, *The endoscopic classification of representations: orthogonal and symplectic groups*, American Mathematical Society Colloquium Publications **61**, American Mathematical Society, Providence, RI, 2013. MR 3135650 Zbl 06231010

[Barbasch 2004] D. Barbasch, "Relevant and petite K-types for split groups", pp. 35–71 in *Functional analysis VIII* (Dubrovnik, 2003), edited by D. Bakić et al., Various Publ. Ser. (Aarhus) **47**, Aarhus University, Aarhus, 2004. MR 2006d:22020 Zbl 1070.22006

[Barbasch 2010] D. Barbasch, "The unitary spherical spectrum for split classical groups", *J. Inst. Math. Jussieu* **9**:2 (2010), 265–356. MR 2011d:22014 Zbl 1188.22010

[Barbasch et al. 2008] D. Barbasch, D. Ciubotaru, and A. Pantano, "Unitarizable minimal principal series of reductive groups", pp. 63–136 in *Representation theory of real reductive Lie groups*, edited by J. Arthur et al., Contemp. Math. **472**, American Mathematical Society, Providence, RI, 2008. MR 2011e:22020 Zbl 1178.22017

[Borel 1979] A. Borel, "Automorphic L-functions", pp. 27–61 in *Automorphic forms, representations and L-functions, Part 2* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, American Mathematical Society, Providence, RI, 1979. MR 81m:10056 Zbl 0412.10017

[Fulton and Harris 1991] W. Fulton and J. Harris, *Representation theory: a first course*, Graduate Texts in Mathematics **129**, Springer, New York, 1991. MR 93a:20069 Zbl 0744.22001

[Howe 1982] R. Howe, "On a notion of rank for unitary representations of the classical groups", pp. 224–331 in *Harmonic analysis and group representation* (Cortona, 1980), edited by A. Figà Talamanca, CIME Summer Schools **82**, Liguori, Naples, 1982. Reprinted by Springer, Berlin, 2011. MR 86j:22016

[Howe 1989] R. Howe, "Transcending classical invariant theory", *J. Amer. Math. Soc.* **2**:3 (1989), 535–552. MR 90k:22016 Zbl 0716.22006

[Knapp and Vogan 1995] A. W. Knapp and D. A. Vogan, Jr., *Cohomological induction and unitary representations*, Princeton Mathematical Series **45**, Princeton University Press, 1995. MR 96c:22023 Zbl 0863.22011

[Knapp and Zuckerman 1977] A. W. Knapp and G. Zuckerman, "Classification theorems for representations of semisimple Lie groups", pp. 138–159 in *Non-commutative harmonic analysis* (Marseille-Luminy, 1976), edited by J. Carmona and M. Vergne, Lecture Notes in Math. **587**, Springer, Berlin, 1977. MR 57 #16474 Zbl 0353.22011

[Langlands 1970] R. P. Langlands, "Problems in the theory of automorphic forms", pp. 18–61 in *Lectures in modern analysis and applications, III*, edited by C. T. Taam, Lecture Notes in Math. **170**, Springer, Berlin, 1970. MR 46 #1758 Zbl 0225.14022

[Li 1989a] J.-S. Li, "On the classification of irreducible low rank unitary representations of classical groups", *Compositio Math.* **71**:1 (1989), 29–48. MR 90k:22027 Zbl 0694.22012

[Li 1989b] J.-S. Li, "On the singular rank of a representation", *Proc. Amer. Math. Soc.* **106**:2 (1989), 567–571. MR 89k:22029 Zbl 0682.22009

[Li 1989c] J.-S. Li, "Singular unitary representations of classical groups", *Invent. Math.* **97**:2 (1989), 237–255. MR 90h:22021 Zbl 0694.22011

[Li 1997] J.-S. Li, "Unipotent representations attached to small nilpotent orbits", 1997, Available at http://www.math.umd.edu/~jda/seattle_proceedings/li.ps. Lecture notes from a conference on *Representation theory of real and p-adic reductive groups*, (Seattle, WA, 1997).

[Littlewood 1944] D. E. Littlewood, "On invariant theory under restricted groups", *Philos. Trans. Roy. Soc. London. Ser. A.* **239** (1944), 387–417. MR 7,6e Zbl 0060.04403

[Pantano et al. 2010] A. Pantano, A. Paul, and S. A. Salamanca-Riba, "Unitary genuine principal series of the metaplectic group", *Represent. Theory* **14** (2010), 201–248. MR 2011a:22013 Zbl 1200.22006

[Przebinda 1996] T. Przebinda, "The duality correspondence of infinitesimal characters", *Colloq. Math.* **70**:1 (1996), 93–102. MR 96m:22034 Zbl 0854.22017

[Serre 1977] J.-P. Serre, *Linear representations of finite groups*, Graduate Texts in Mathematics **42**, Springer, New York, 1977. MR 56 #8675 Zbl 0355.20006

[Speh and Vogan 1980] B. Speh and D. A. Vogan, Jr., "Reducibility of generalized principal series representations", *Acta Math.* **145**:3-4 (1980), 227–299. MR 82c:22018 Zbl 0457.22011

[Vogan 1981] D. A. Vogan, Jr., *Representations of real reductive Lie groups*, Progress in Mathematics **15**, Birkhäuser, Boston, 1981. MR 83c:22022 Zbl 0469.22012

[Vogan 1993] D. A. Vogan, Jr., "Unipotent representations and cohomological induction", pp. 47–70 in *The Penrose transform and analytic cohomology in representation theory* (South Hadley, MA, 1992), edited by M. Eastwood et al., Contemp. Math. **154**, American Mathematical Society, Providence, RI, 1993. MR 94j:22017 Zbl 0822.22009

Received July 7, 2013. Revised June 24, 2014.

ALESSANDRA PANTANO
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, IRVINE
IRVINE, CA 92697
UNITED STATES
apantano@uci.edu

ANNEGRET PAUL
DEPARTMENT OF MATHEMATICS
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MI 49008
UNITED STATES
annegret.paul@wmich.edu

SUSANA SALAMANCA RIBA
DEPARTMENT OF MATHEMATICS
NEW MEXICO STATE UNIVERSITY
LAS CRUCES, NM 88003
UNITED STATES
ssalaman@nmsu.edu

CONTENTS

Volume 271, no. 1 and no. 2

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

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.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 271 No. 2 October 2014