

*AG
T*

*Algebraic & Geometric
Topology*

Volume 25 (2025)

Issue 4 (pages 1917–2526)

ALGEBRAIC & GEOMETRIC TOPOLOGY

msp.org/agt

EDITORS

PRINCIPAL ACADEMIC EDITORS

John Etnyre
etnyre@math.gatech.edu
Georgia Institute of Technology

Kathryn Hess
kathryn.hess@epfl.ch
École Polytechnique Fédérale de Lausanne

BOARD OF EDITORS

Julie Bergner	University of Virginia jeb2md@eservices.virginia.edu	Thomas Koberda	University of Virginia thomas.koberda@virginia.edu
Steven Boyer	Université du Québec à Montréal cohf@math.rochester.edu	Markus Land	LMU München markus.land@math.lmu.de
Tara E Brendle	University of Glasgow tara.brendle@glasgow.ac.uk	Christine Lescop	Université Joseph Fourier lescop@ujf-grenoble.fr
Indira Chatterji	CNRS & Univ. Côte d'Azur (Nice) indira.chatterji@math.cnrs.fr	Norihiko Minami	Yamato University minami.norihiko@yamato-u.ac.jp
Octav Cornea	Université de Montreal cornea@dms.umontreal.ca	Andrés Navas	Universidad de Santiago de Chile andres.navas@usach.cl
Alexander Dranishnikov	University of Florida dranish@math.ufl.edu	Robert Oliver	Université Paris 13 bobol@math.univ-paris13.fr
Tobias Ekholm	Uppsala University, Sweden tobias.ekholm@math.uu.se	Jessica S Purcell	Monash University jessica.purcell@monash.edu
Mario Eudave-Muñoz	Univ. Nacional Autónoma de México mario@matem.unam.mx	Birgit Richter	Universität Hamburg birgit.richter@uni-hamburg.de
David Futer	Temple University dfuter@temple.edu	Jérôme Scherer	École Polytech. Féd. de Lausanne jerome.scherer@epfl.ch
John Greenlees	University of Warwick john.greenlees@warwick.ac.uk	Vesna Stojanoska	Univ. of Illinois at Urbana-Champaign vesna@illinois.edu
Matthew Hedden	Michigan State University mhedden@math.msu.edu	Zoltán Szabó	Princeton University szabo@math.princeton.edu
Kristen Hendricks	Rutgers University kristen.hendricks@rutgers.edu	Maggy Tomova	University of Iowa maggy-tomova@uiowa.edu
Hans-Werner Henn	Université Louis Pasteur henn@math.u-strasbg.fr	Daniel T Wise	McGill University, Canada daniel.wise@mcgill.ca
Daniel Isaksen	Wayne State University isaksen@math.wayne.edu	Lior Yanovski	Hebrew University of Jerusalem lior.yanovski@gmail.com

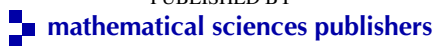
See inside back cover or msp.org/agt for submission instructions.

The subscription price for 2025 is US \$760/year for the electronic version, and \$1110/year (+\$75, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Algebraic & Geometric Topology is indexed by [Mathematical Reviews](#), [Zentralblatt MATH](#), [Current Mathematical Publications](#) and the [Science Citation Index](#).

Algebraic & Geometric Topology (ISSN 1472-2747 printed, 1472-2739 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

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

PUBLISHED BY



mathematical sciences publishers

nonprofit scientific publishing

<https://msp.org/>

© 2025 Mathematical Sciences Publishers

The zero stability for the one-row colored $\mathfrak{sl}_3$ -Jones polynomial

WATARU YUASA

The stability of coefficients of colored ($\mathfrak{sl}_2$ -)Jones polynomials $\{J_{K,n}^{\mathfrak{sl}_2}(q)\}_n$ was discovered by Dasbach and Lin. This stability is now called the zero stability of $J_{K,n}^{\mathfrak{sl}_2}(q)$. Armond showed zero stability for a B -adequate link by using the linear skein theory based on the Kauffman bracket. We prove the zero stability of one-row colored $\mathfrak{sl}_3$ -Jones polynomials $\{J_{K,n}^{\mathfrak{sl}_3}(q)\}_n$ for B -adequate links L with antiparallel twist regions by using the linear skein theory based on Kuperberg's $\mathfrak{sl}_3$ -webs. This implies the existence of many q -series obtained from a quantum invariant associated with $\mathfrak{sl}_3$.

57K10, 57K14, 57K16

1 Introduction

The colored $\mathfrak{g}$ -Jones polynomial of a knot K is a quantum invariant obtained from an irreducible representation of a simple Lie algebra $\mathfrak{g}$. We will discuss some stability of coefficients of the one-row colored $\mathfrak{sl}_3$ -Jones polynomials $\{J_{K,n}^{\mathfrak{sl}_3}(q) \in \mathbb{Z}[q^{\pm 1/2}] \mid n \in \mathbb{N}\}$, which is a quantum invariant of K associated with irreducible representations of $\mathfrak{sl}_3$ corresponding to the one-row Yang diagram (n) . This kind of stability for the colored ($\mathfrak{sl}_2$ -)Jones polynomials was discovered by Dasbach and Lin [5; 6]. They showed that some leading coefficients, concerning the degree of q , of $\{J_{K,n}^{\mathfrak{sl}_2}(q)\}_n$ are independent of the colorings n (where $n+1$ is the dimension of an irreducible representation) if a knot K is alternating. They also conjectured that the first n coefficients of $J_{K,N}^{\mathfrak{sl}_2}(q)$ are constant for all N greater than n if K is alternating. Armond [1] proved this conjecture for a larger class of links, called adequate links, which contain alternating links. Independently, Garoufalidis and Lê [9] proved more general stability, called k -stability, for alternating links, where k is a nonnegative integer. In the sense of Garoufalidis and Lê, the stability proved in [1] corresponds to the zero stability. The k -stability also ensures the existence of a q -series called the k -limit, which is closely related to quantum modular forms. The 0-limit is also known as the *tail* of K .

Definition 1.1 Let $\hat{J}_{K,n}^{\mathfrak{sl}_2}(q) := \pm q^d J_{K,n}^{\mathfrak{sl}_2}(q) = a_0 + \sum_{i=1}^{\infty} a_i q^i$ be a normalization of the colored Jones polynomial $J_{K,n}^{\mathfrak{sl}_2}(q)$ of a knot K , where the sign is

chosen so that a_0 is positive. The *tail* of K is a q -series $\Phi_K(q) \in \mathbb{Z}[[q]]$ satisfying

$$\Phi_K(q) - \hat{J}_{K,n}^{\mathfrak{sl}_2}(q) \in q^{n+1} \mathbb{Z}[[q]]$$

for any positive integer n .

Note that the integrality theorem for the colored Jones polynomial proved by Lê [20] claims that the coefficients of $\hat{J}_{K,n}(q)$ become integral, and therefore its tail $\Phi_K(q)$ belongs to $\mathbb{Z}[[q]]$. Armond and Dasbach [3] showed that the tail of an adequate knot is determined by its reduced B -graph.¹ A similar result was obtained by Garoufalidis, Norin and Vuong [11]. They stated that the first three coefficients of $\Phi_K(q)$ of an alternating link K are described in terms of its reduced Tait graph. From these results, we can see that the tail is not useful in distinguishing links. However, tails of knots and links give us interesting q -series related to quantum modular forms. For example, Garoufalidis and Lê [9] showed that tails of alternating links are described as a generalization of Nahm sums. In particular, the tail of a $(2, m)$ -torus link is the (false) theta series. In work of Armond and Dasbach [2], Hajij [13] and Yuasa [25], an Andrews–Gordon-type identity for the (false) theta series was derived from two explicit formulas for the tail of a $(2, m)$ -torus link. Explicit formulas for tails of other knots and links have been studied by Garoufalidis and Lê [9], Elhamdadi and Hajij [7], Keilthy and Osburn [15] and Beirne and Osburn [4], and for quantum spin networks by Hajij [13].

Our goal is to develop a study of the stability and tails for $J_{K,n}^{\mathfrak{sl}_2}(q)$ to quantum invariants $J_{K,\lambda}^{\mathfrak{g}}(q)$ associated with a higher-rank simple Lie algebra $\mathfrak{g}$. Many problems arise when we consider higher-rank cases. For example, we have to choose a sequence of irreducible representations to consider the stability because the colored $\mathfrak{g}$ -Jones polynomial of a knot is parametrized by dominant weights. Moreover, the explicit computation of the colored $\mathfrak{g}$ -Jones polynomials of a given knot is much more difficult than in the $\mathfrak{sl}_2$ case.

The aim of this paper is to show zero stability of the one-row colored $\mathfrak{sl}_3$ -Jones polynomial $\{J_{K,n}^{\mathfrak{sl}_3}(q)\}_n$ of a B -adequate link K with antiparallel twist regions. The one-row coloring n for K means that all components of K are colored by the irreducible representation of the highest weight $n\varpi_1$ (or we write it as $(n, 0)$), where $\{\varpi_i\}_{i=1,2}$ correspond to the fundamental weights of $\mathfrak{sl}_3$. There are some studies on the explicit computation of the colored $\mathfrak{sl}_3$ Jones polynomial: for the trefoil knot by Lawrence [19], for the $(2, 2m+1)$ - and $(4, 5)$ -torus knots with general coloring by Garoufalidis, Morton and Vuong [10; 12], for 2-bridge links with one-row coloring by Yuasa [24], and for pretzel links with one-row coloring by Kawasoe [14]. These explicit formulas give tails of the colored $\mathfrak{sl}_3$ -Jones polynomial of some links; see Garoufalidis and Vuong [12] and Yuasa [25; 26]. For the λ -colored $\mathfrak{g}$ -Jones polynomial of the (a, b) -torus knot when $\mathfrak{g}$ is a simple Lie algebra of rank 2, Garoufalidis and Vuong [12] proved the k -stability for any k . They used the formula of Rosso and Jones in [22] to prove it. Here we prove the zero stability of the one-row colored $\mathfrak{sl}_3$ -Jones polynomial for any antiparallel B -adequate link using the linear skein theory for $\mathfrak{sl}_3$ developed by Kuperberg [18]. Our proof is inspired by work of Armond [1] using the Kauffman bracket.

Theorem 1 (zero stability for the one-row colored $\mathfrak{sl}_3$ -Jones polynomial; see Theorem 4.16) *For any antiparallel B -adequate link L , there exists $\Phi_L^{\mathfrak{sl}_3}(q)$ in $\mathbb{Z}[[q]]$ such that*

$$\Phi_L^{\mathfrak{sl}_3}(q) - \hat{J}_{L,n}^{\mathfrak{sl}_3}(q) \in q^{n+1}\mathbb{Z}[[q]].$$

¹They consider A -graphs. However, this corresponds to B -graphs in our convention. That is, our q is q^{-1} in [3].

An antiparallel B -adequate link is an oriented link whose representative is a B -adequate link diagram with only antiparallel twist regions; see [Definition 4.1](#).

This result and its proof are an extension of the work on the zero stability of colored $\mathfrak{sl}_2$ -Jones polynomials in [\[1\]](#) to $\mathfrak{sl}_3$. We will discuss the zero stability for general B -adequate links in a forthcoming paper.

This paper is organized as follows. In [Section 2](#), we introduce the $\mathfrak{sl}_3$ version of the linear skein theory and review properties of $\mathfrak{sl}_3$ -webs and $\mathfrak{sl}_3$ -clasps. In [Section 3](#), we discuss a lower bound on the minimum degree of a clasped $\mathfrak{sl}_3$ -web. In [Section 4](#), we prove the zero stability of the one-row colored $\mathfrak{sl}_3$ -Jones polynomials by calculating clasped $\mathfrak{sl}_3$ -webs. In the [appendix](#), we prove some new formulas for the clasped $\mathfrak{sl}_3$ -webs used in this paper.

Acknowledgments

The author gratefully thanks the referees for their careful reading, helpful comments and great efforts to improve this paper. This work was supported by Grant-in-Aid for Early-Career Scientists grants JP19K14528, JP23K12972.

2 $\mathfrak{sl}_3$ -webs and $\mathfrak{sl}_3$ -clasps

We mainly work with a space of $\mathfrak{sl}_3$ -webs, which is a linear combination of oriented planar trivalent graphs with coefficients in $\mathcal{R} = \mathbb{Z}((q^{1/6}))$. Let us introduce some useful symbols for elements in $\mathcal{R}$. We define

- a *quantum integer* by $[n] := (q^{n/2} - q^{-n/2}) / (q^{1/2} - q^{-1/2})$ for any nonnegative integer $n \in \mathbb{Z}_{\geq 0}$, and
- a *quantum binomial coefficient* by $\begin{bmatrix} n \\ k \end{bmatrix} := [n]! / ([k]![n-k]!)$ for $0 \leq k \leq n$, and $\begin{bmatrix} n \\ k \end{bmatrix} = 0$ for $k > n$, where $[n]! := [n][n-1] \cdots [1]$.

Let us define $\mathfrak{sl}_3$ -web spaces based on [\[18\]](#). We consider a surface Σ equipped with signed marked points (P, s) , where $P \subset \partial\Sigma$ is a finite set and $s : P \rightarrow \{+, -\}$ a map.

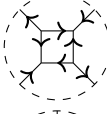
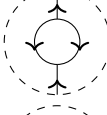
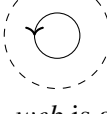
A *tangled trivalent graph* on Σ is an immersion of a directed graph G into Σ satisfying:

- (1) The valency of a vertex of G is 1 or 3.
- (2) All crossing points are transversal double points of two edges with under/over-passing information.
- (3) The set of univalent vertices of G coincides with P .
- (4) A neighborhood of a vertex in Σ is one of



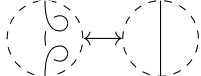
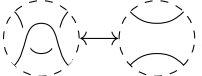
A tangled trivalent graph is *flat* if it has no crossings. An *elliptic face* of a flat trivalent graph G is a 0-gon (ie a disk), 2-gon or 4-gon in the set of connected components of $\Sigma \setminus G$ which does not touch the boundary of Σ .

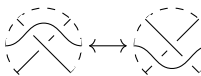
Definition 2.1 ($\mathfrak{sl}_3$ -web spaces [18]) Let $\mathcal{G}(s; \Sigma)$ be the set of the boundary-fixing isotopy classes of tangled trivalent graphs on Σ . The $\mathfrak{sl}_3$ -web space $\mathcal{W}(s; \Sigma)$ is the quotient of the $\mathcal{R}$ -module spanned by $\mathcal{G}(s; \Sigma)$ modulo the following $\mathfrak{sl}_3$ -skein relations:

-  $= q^{1/3} \langle \text{left side of crossing} \rangle - q^{-1/6} \langle \text{right side of crossing} \rangle$.
-  $= q^{-1/3} \langle \text{left side of crossing} \rangle - q^{1/6} \langle \text{right side of crossing} \rangle$.
-  $= \langle \text{left side of 4-gon} \rangle + \langle \text{right side of 4-gon} \rangle$ (the 4-gon relation).
-  $= [2] \langle \text{left side of 2-gon} \rangle$ (the 2-gon relation).
-  $= [3] \langle \text{empty disk} \rangle = \langle \text{loop} \rangle$ (the trivial loop relation).

An $\mathfrak{sl}_3$ -web is an element in $\mathcal{W}(s; \Sigma)$ and a *basis web* is an $\mathfrak{sl}_3$ -web represented by a graph in $\mathcal{G}(s; \Sigma)$ with no elliptic faces.

The $\mathfrak{sl}_3$ -skein relations realize the *Reidemeister moves* (R1') and (R2)–(R4):

(R1') , (R2) ,

(R3) , (R4) .

The above means that $\mathfrak{sl}_3$ -webs on the left and right sides represent the same element in an $\mathfrak{sl}_3$ -web space for any choice of orientation of the edges.

It is easy to see that any tangled trivalent graph decomposes into a sum of basis webs by using the $\mathfrak{sl}_3$ -skein relations. In fact, the set of basis webs consists of a basis of the $\mathfrak{sl}_3$ -web space:

Theorem 2.2 [18; 23] *The set of basis webs on a surface Σ with signed marked points $s : P \rightarrow \{+, -\}$ is a basis of $\mathcal{W}(s; \Sigma)$ as a $\mathbb{Z}[q^{\pm 1/6}]$ -module.*

In some cases, one can give the set of basis webs via an argument concerning the Euler characteristic.

Example 2.3 Let D be a disk with a basepoint $* \in \partial D$. We identify signed marked points on $\partial D \setminus \{*\}$ with a sequence of signs. Then the following isomorphisms hold for any $\epsilon \in \{+, -\}$:

- (1) $\mathcal{W}(\emptyset; D)$ of a disk D with no marked points is isomorphic to a free $\mathcal{R}$ -module spanned by the empty diagram $\emptyset$.
- (2) $\mathcal{W}(\epsilon; D) = 0$.
- (3) $\mathcal{W}(\epsilon\epsilon; D) = 0$.
- (4) $\mathcal{W}(\epsilon\bar{\epsilon}; D)$ is a free $\mathcal{R}$ -module spanned by an oriented simple arc.
- (5) $\mathcal{W}(\epsilon\epsilon\bar{\epsilon}; D) = 0$.
- (6) $\mathcal{W}(\epsilon\epsilon\epsilon; D)$ is a free $\mathcal{R}$ -module spanned by a tripod with a sink or source vertex.

In the above, $\bar{\epsilon}$ means the opposite sign of ϵ .

We review a diagrammatic definition of an $\mathfrak{sl}_3$ -clasp introduced in [18; 21; 17; 26] and note some useful properties. The $\mathfrak{sl}_3$ -clasp plays a similar role to the Jones–Wenzl projector in the Kauffman bracket skein theory.

In what follows, we will mainly consider tangled trivalent graphs or $\mathfrak{sl}_3$ -webs in a rectangle $D = [0, 1] \times [0, 1]$. We assume that the set of marked points lies in the top edge $I_1 = [0, 1] \times \{1\}$ and the bottom edge $I_0 = [0, 1] \times \{0\}$, and a basepoint $*$ at $(0, 0)$. In this situation, the set of marked points is divided into $P^{(0)}$ and $P^{(1)}$, where $P^{(j)} := P \cap I_j$, and we denote the assignment of signs by $s^{(j)}: P^{(j)} \rightarrow \{+, -\}$ for $j = 0, 1$. One can identify $s^{(j)}$ with a sequence of signs on $[0, 1] \times \{j\}$ arranged from 0 to 1. We abbreviate $\mathcal{G}(s^{(0)} \sqcup s^{(1)}; D)$ and $\mathcal{W}(s^{(0)} \sqcup s^{(1)}; D)$ by $\mathcal{G}(s^{(0)}, \bar{s}^{(1)})$ and $\mathcal{TL}(s^{(0)}, \bar{s}^{(1)})$, respectively, where $\bar{s}^{(1)}$ is a sequence consisting of the opposite signs of $s^{(1)}$.² When we describe diagrams representing $\mathfrak{sl}_3$ -webs in D , we omit from the notation the signs, the basepoint, and the boundary of D . The composition $\mathcal{TL}(s_1, s_2) \otimes \mathcal{TL}(s_0, s_1) \rightarrow \mathcal{TL}(s_0, s_2)$ is defined by gluing the top side of an $\mathfrak{sl}_3$ -web in $\mathcal{TL}(s_0, s_1)$ and the bottom side of an $\mathfrak{sl}_3$ -web in $\mathcal{TL}(s_1, s_2)$.

We first define the $\mathfrak{sl}_3$ -clasp in $\mathcal{TL}(-^m, -^m)$.

Definition 2.4 (one-row colored $\mathfrak{sl}_3$ -clasps) The $\mathfrak{sl}_3$ -clasp JW_{-m} described by a white box with $m \in \mathbb{Z}_{>0}$ is defined as follows:

$$\begin{aligned}
 (1) \quad \text{JW}_{-} &= \begin{array}{c} \uparrow \\ \boxed{} \\ \downarrow \end{array} := \begin{array}{c} \uparrow \\ \downarrow \end{array} \\
 (2) \quad \text{JW}_{-m+1} &= \begin{array}{c} \uparrow^{m+1} \\ \boxed{} \\ \downarrow \end{array} := \begin{array}{c} \uparrow^m \\ \boxed{} \\ \downarrow \end{array} - \frac{[m]}{[m+1]} \begin{array}{c} \uparrow^m \\ \boxed{} \\ \downarrow^m \end{array} \begin{array}{c} \uparrow \\ \downarrow \end{array}.
 \end{aligned}$$

In the above, an edge labeled by m represents the m -parallelization of the edge. JW_{+^m} is defined by the same diagram with oppositely directed edges.

²We take the opposite sign $\bar{s}^{(1)}$ to be consistent with the composition.

We next introduce the $\mathfrak{sl}_3$ -clasp in $\mathcal{TL}(-m+n, -m+n)$.

Definition 2.5 (two-row colored $\mathfrak{sl}_3$ -clasps) We define

$$\text{JW}_{-m+n} = \begin{array}{c} m \quad n \\ \downarrow \quad \downarrow \\ \boxed{} \\ \uparrow \quad \uparrow \\ m \quad n \end{array} = \sum_{i=0}^{\min\{m,n\}} (-1)^i \frac{[m]_i [n]_i}{[m+n+1]_i} \begin{array}{c} m \quad i \quad n \\ \downarrow \quad \downarrow \quad \downarrow \\ \boxed{} \quad \boxed{} \\ \uparrow \quad \uparrow \quad \uparrow \\ m-i \quad i \quad n-i \end{array}.$$

One can define JW_{+m-n} in the same way.

For convenience in the computation of $\mathfrak{sl}_3$ -webs, we will introduce “stair-step” and “triangle” webs in Definitions 2.6 and 2.7, which were also used in [16; 17; 24; 8]. In these definitions, the orientation of edges of $\mathfrak{sl}_3$ -webs is not explicitly given because it is uniquely determined according to your choice of the orientation of an edge around the box.

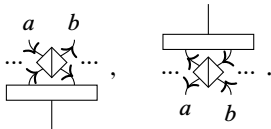
Definition 2.6 For positive integers n and m , a *stair-step web* is defined recursively by

$$\begin{array}{c} 1 \\ \diagdown \quad \diagup \\ n \quad n \\ \diagup \quad \diagdown \\ 1 \end{array} = n \left\{ \begin{array}{c} \text{---} \quad \text{---} \\ | \quad | \\ \text{---} \quad \text{---} \\ | \quad | \\ \text{---} \quad \text{---} \end{array} \right\} \quad \text{and} \quad \begin{array}{c} m \\ \diagdown \quad \diagup \\ n \quad n \\ \diagup \quad \diagdown \\ m \end{array} = n \begin{array}{c} m-1 \quad 1 \\ \diagdown \quad \diagup \\ n \quad n \\ \diagup \quad \diagdown \\ m-1 \quad 1 \end{array} \quad \text{for } m > 1.$$

Definition 2.7 For positive integers n , define the *triangle webs* by

$$\begin{array}{c} 1 \quad 1 \\ \diagdown \quad \diagup \\ 1 \\ \diagup \quad \diagdown \\ 1 \end{array} = \begin{array}{c} \diagdown \quad \diagup \\ | \\ \diagup \quad \diagdown \end{array} \quad \text{and} \quad \begin{array}{c} 1 \quad 1 \\ \diagdown \quad \diagup \\ n \quad n \\ \diagup \quad \diagdown \\ n \end{array} = \begin{array}{c} 1 \quad \text{---} \quad 1 \\ | \quad \boxed{} \quad | \\ n-1 \quad 1 \quad n-1 \\ | \quad \diagdown \quad \diagup \\ 1 \quad n-1 \end{array} \quad \text{for } n > 1.$$

Definition 2.8 (general type of the $\mathfrak{sl}_3$ -clasps) Let s_1 and s_2 be two sequences of signs which consist of m pluses and n minuses. We define an $\mathfrak{sl}_3$ -clasp in $\mathcal{TL}(s_1, \bar{s}_2)$ by gluing stair-step webs to the top and the bottom of JW_{-m+n} as follows:



Repeating these operations, we can arbitrarily exchange signs in sequences on the top and bottom of the disk, respectively, and we obtain an $\mathfrak{sl}_3$ -web in $\mathcal{TL}(s_1, \bar{s}_2)$. We denote it by $\text{JW}_{s_1}^{\bar{s}_2}$ and depict it as



The resulting $\mathfrak{sl}_3$ -clasp is independent of choice of sequences of stair-step webs; see [26]. Thus, $\text{JW}_{s_1}^{\bar{s}_2}$ is uniquely determined by s_1 and s_2 .

One can prove the following useful formulas for $\mathfrak{sl}_3$ -clasps by straightforward computation:

Lemma 2.9 Let s_1, s_2 and s_3 be sequences of signs. An arc labeled by a positive integer m (resp. n) denotes m -parallelization (resp. n -parallelization) of the arc.

$$\begin{aligned}
 (1) \quad & \begin{array}{c} s_3 \\ \hline s_2 \\ \hline s_1 \end{array} = \begin{array}{c} s_3 \\ \hline s_1 \end{array}, \quad \begin{array}{c} \cdots \\ \diagup \quad \diagdown \\ \cdots \end{array} = 0, \quad \begin{array}{c} \cdots \\ \text{cap} \\ \cdots \end{array} = 0, \\
 (2) \quad & \begin{array}{c} m \quad n \\ \diagdown \quad \diagup \\ \cdots \end{array} = \begin{array}{c} m \quad n \\ \diagup \quad \diagdown \\ \cdots \end{array}, \quad \begin{array}{c} m \quad n \\ \diagup \quad \diagdown \\ \cdots \end{array} = \begin{array}{c} m \quad n \\ \diagdown \quad \diagup \\ \cdots \end{array}, \\
 (3) \quad & \begin{array}{c} m \quad n \\ \text{cap} \\ \cdots \end{array} = (-1)^{mn} q^{mn/6} \begin{array}{c} m \quad n \\ \diagup \quad \diagdown \\ \cdots \end{array}, \quad \begin{array}{c} m \quad n \\ \text{cup} \\ \cdots \end{array} = (-1)^{mn} q^{-mn/6} \begin{array}{c} m \quad n \\ \diagdown \quad \diagup \\ \cdots \end{array}, \\
 (4) \quad & \begin{array}{c} m \quad n \\ \text{cap} \\ \cdots \end{array} = q^{mn/3} \begin{array}{c} m \quad n \\ \diagup \quad \diagdown \\ \cdots \end{array}, \quad \begin{array}{c} m \quad n \\ \text{cup} \\ \cdots \end{array} = q^{-mn/3} \begin{array}{c} m \quad n \\ \diagdown \quad \diagup \\ \cdots \end{array}, \\
 (5) \quad & \begin{array}{c} n \\ \diagdown \quad \diagup \\ \text{triangle} \\ \diagup \quad \diagdown \\ n \end{array} = \begin{array}{c} n \\ \diagdown \quad \diagup \\ \text{triangle} \\ \diagup \quad \diagdown \\ n \end{array}, \quad \begin{array}{c} n \\ \diagdown \quad \diagup \\ \text{triangle} \\ \diagup \quad \diagdown \\ n \end{array} = \begin{array}{c} n \\ \diagdown \quad \diagup \\ \text{triangle} \\ \diagup \quad \diagdown \\ n \end{array}, \\
 (6) \quad & \begin{array}{c} n \\ \text{loop} \\ \diagup \quad \diagdown \end{array} = q^{(n^2+3n)/3} \begin{array}{c} n \\ \diagup \quad \diagdown \end{array}, \quad \begin{array}{c} n \\ \text{loop} \\ \diagdown \quad \diagup \end{array} = q^{-(n^2+3n)/3} \begin{array}{c} n \\ \diagdown \quad \diagup \end{array}.
 \end{aligned}$$

Proof One can prove (1)–(6) by induction on labels and the skein relations. See [24; 26], for example. $\square$

We give a definition of the one-row colored $\mathfrak{sl}_3$ -Jones polynomial of oriented framed links via an $\mathfrak{sl}_3$ -web. First we introduce a normalization of a Laurent series by shifting the q -degree and changing the sign.

Definition 2.10 (minimum degree) We define the *minimum degree* $d_*: \mathcal{R} = \mathbb{Z}((q^{1/6})) \rightarrow \frac{1}{6}\mathbb{Z} \cup \infty$ by $d_*(f(q)) := \frac{1}{6}d$ for a nonzero series $f(q) = \sum_{i=d}^{\infty} a_i q^{i/6}$ in $\mathbb{Z}((q^{1/6}))$ such that $a_d \neq 0$. For the zero polynomial, we define its minimum degree as ∞ . We also define a normalization $\hat{f}(q)$ of a nonzero Laurent series $f(q)$ as

$$\hat{f}(q) := \pm q^{-d_*(f(q))} f(q) = \pm \sum_{i=0}^{\infty} a_{i+d} q^{i/6} \in \mathbb{Z}[[q^{1/6}]].$$

In the above, we choose the sign so that the constant term $\pm a_d$ is positive.

We note some properties of the minimum degree and useful examples:

Lemma 2.11 For any $f(q), g(q) \in \mathcal{R}$:

- (1) $d_*(f(q) + g(q)) \geq \min\{d_*(f(q)), d_*(g(q))\}$.
- (2) $d_*(f(q)g(q)) = d_*(f(q)) + d_*(g(q))$.

The equality in (1) holds if and only if $d_*(f(q)) \neq d_*(g(q))$ or $d_*(f(q)) = d_*(g(q)) =: d$ with $a_d + b_d \neq 0$, where $f(q) = \sum_{i=d}^{\infty} a_i q^{i/6}$ and $g(q) = \sum_{i=d}^{\infty} b_i q^{i/6}$.

Example 2.12 For any positive integer n and $1 \leq k \leq n$,

$$d_*([n]) = -\frac{1}{2}(n-1), \quad d_*([n]^{-1}) = \frac{1}{2}(n-1), \quad d_*\left(\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]\right) = -\frac{1}{2}k(n-k), \quad d_*\left(\left[\begin{smallmatrix} n \\ k \end{smallmatrix}\right]^{-1}\right) = \frac{1}{2}k(n-k).$$

These follow from the identity $(1 - q^m)^{-1} = 1 + q^m + q^{2m} + \cdots \in \mathcal{R} = \mathbb{Z}((q^{1/6}))$ and [Lemma 2.11](#).

Definition 2.13 Let L be a link diagram of a framed link whose framing is given by the blackboard framing. One can replace arcs of the link diagram with n parallelized arcs and put white boxes on the n parallelized arcs. The resulting diagram, denoted by $L^{(n)}$, represents an $\mathfrak{sl}_3$ -web in a disk D with no marked points.³ The *one-row colored $\mathfrak{sl}_3$ -Jones polynomial* $J_{L,n}^{\mathfrak{sl}_3}(q)$ with $(n, 0)$ -coloring (or n boxes) is defined by $L^{(n)} = J_{L,n}^{\mathfrak{sl}_3}(q)\emptyset$. We also define a variation of the one-row colored $\mathfrak{sl}_3$ -Jones polynomial as $\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)$, according to [Definition 2.10](#).

Remark 2.14 • Skein relations realize the Reidemeister moves (R1')–(R4) for arcs with one-row colored clasps because clasped arcs are expressed as linear combinations of $\mathfrak{sl}_3$ -webs. Hence, $J_{L,n}^{\mathfrak{sl}_3}(q)$ is an invariant of framed links.

- The choice of framing of L appears as multiplication by $\pm q^\bullet$; see [Lemma 2.9\(6\)](#). This difference is ignored in the normalization $\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)$; thus, it is an invariant of links.
- Lê [\[20\]](#) showed the integrality theorem for a quantum $\mathfrak{g}$ invariant of links. It says that $\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)$ belongs to $\mathbb{Z}[q]$.

We will discuss *zero stability* of $\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)$ for a certain class of links in the following sections. Let us recall the definition of zero stability and tails of the one-row colored $\mathfrak{sl}_3$ -Jones polynomials.

Definition 2.15 (one-row colored $\mathfrak{sl}_3$ -tail) The one-row colored $\mathfrak{sl}_3$ -Jones polynomial $\{\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)\}_n$ of a link L is *zero stable* if there exists a formal power series $\Phi_L^{\mathfrak{sl}_3}(q)$ in $\mathbb{Z}[[q]]$ such that

$$\Phi_L^{\mathfrak{sl}_3}(q) - \hat{J}_{L,n}^{\mathfrak{sl}_3}(q) \in q^{n+1}\mathbb{Z}[[q]]$$

for all $n \geq 1$. We call $\Phi_L^{\mathfrak{sl}_3}(q)$ the *one-row colored $\mathfrak{sl}_3$ -tail* of L , or simply the $\mathfrak{sl}_3$ -tail of L when $\{\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)\}_n$ is zero stable.

3 The minimum degree of clasped $\mathfrak{sl}_3$ -webs

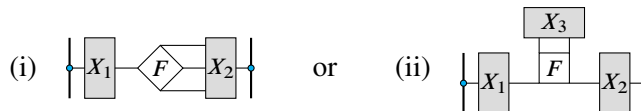
We will prove the existence of the $\mathfrak{sl}_3$ -tail of the one-row colored $\mathfrak{sl}_3$ -Jones polynomial by developing an $\mathfrak{sl}_3$ analog of Armond's argument [\[1\]](#) using the Kauffman bracket. In this section, we will discuss a lower bound on the minimum degree of a clasped $\mathfrak{sl}_3$ -web with no crossings in a disk.

First we prepare a lemma that studies the isomorphisms in [Example 2.3](#) in detail:

³Such an $\mathfrak{sl}_3$ -web space $\mathcal{W}(\emptyset; D)$ is spanned by the empty diagram $\emptyset$; see [Example 2.3](#).

Lemma 3.1 Let $G \in \mathcal{G}(-+; D)$ be a connected flat trivalent graph in a disk D with two marked points such that $G \neq 0$. There exists a sequence, composed only of 2- and 4-gon relations, that reduces G to an oriented simple arc γ without increasing the number of connected components of intermediate graphs.

Proof Let us assume that G has at least one elliptic face and $G \neq 0$. We only have to attend to the 4-gon relation because the 2-gon relation does not change the number of connected components of a graph. Let us prove the claim by induction on the number $v(G)$ of vertices. A connected flat trivalent graph G with $v(G) = 2$ has to be the diagram in the left-hand side of the 2-gon relation in Definition 2.1, and G with $v(G) = 3$ does not exist because a trivalent vertex is a sink or source. Let G be a connected flat trivalent graph with $v(G) \geq 4$ and assume that it has at least one internal 4-gon F . Example 2.3(2) claims that we cannot describe a circle in D which intersects G at a single edge. Example 2.3(5) claims that no circle intersects two incoming (resp. outgoing) edges and one outgoing (resp. incoming) edge of G . This fact requires that four edges incident to corners of F connect to other parts of G like either



where the subgraphs X_1 , X_2 , and X_3 of G are connected. We apply the 4-gon relation at F in (i):

$$\text{Diagram (i)} = \text{Diagram (left)} + \text{Diagram (right)}$$

The equation shows the 4-gon relation applied to case (i). The left side is the original graph with 4-gon F . The right side is the sum of two graphs: the first graph has X_1 and X_2 connected by a single edge, and the second graph has X_1 and X_2 connected by a different edge configuration.

Graphs after applying the 4-gon relation are divided into the left and right parts containing X_1 and X_2 , respectively, by cutting along a vertical line in D . Right and left subgraphs are considered $\mathfrak{sl}_3$ -web in a disk with two marked points whose number of vertices is smaller than $v(G)$. The right subgraph containing X_2 is connected due to Example 2.3(2). Hence these subgraphs satisfy the induction hypothesis. For case (ii), a sequence of the 2- and 4-gon relations changes X_3 into an arc without increasing the number of components because $v(X_3) < v(G)$. Then one can obtain a graph consisting of X_1 and X_2 by applying the 2-gon relation twice. The resulting graph also satisfies the induction hypothesis. $\square$

Proposition 3.2 Let G be a flat trivalent graph in a disk with no marked points, and we identify G with its value in $\mathcal{R}$ (see Example 2.3). Then

$$d_*(G) \geq -\frac{1}{4}v(G) - c(G),$$

where $v(G)$ is the number of trivalent vertices of G and $c(G)$ is the number of connected components of G . Moreover, $d_*(G) = -c(G)$ holds when $v(G) = 0$.

Proof We first prove $d_*(G) = -c(G)$ when $v(G) = 0$. If G has no trivalent vertices, then it consists only of loop components. By using an innermost argument and the trivial loop relation in Definition 2.1, it is easy to see that $G = [3]^{c(G)}$. We obtain $d_*(G) = c(G)d_*([3]) = -c(G)$ because $[3] = q + 1 + q^{-1}$. Let us consider when $G = \bigsqcup_{i=1}^{c(G)} G_i$ is a nontrivial flat trivalent graph with $v(G) > 0$, where the G_i are connected components of G . Choose a point p_i on the outermost edge of G_i and a small interval I_{p_i} for each $i = 1, 2, \dots, c(G)$. Then one can take disks $\{D_i\}_i^{c(G)}$ with two marked points all $i = 1, 2, \dots, c(G)$

such that $G_i \setminus \text{int}(I_{p_i}) \subset D_i$, ∂I_{p_i} is precisely the marked points of D_i , and $D_i \cap D_j = \emptyset$ for $i \neq j$. Apply Lemma 3.1 to $G_i \cap D_i$ for all $i \in \{1, 2, \dots, c(G)\}$ to obtain a disjoint union $\Gamma := \bigsqcup_{i=1}^{c(G)} \gamma_i$ of simple loops from G . Each component γ_i is obtained from G_i by a sequence of 2- and 4-gon relations preserving the number of connected components. Let $G = G' + G''$ be a 4-gon relation appearing in the above sequence; we can assume that $G' \neq 0$ and $d_*(G') \leq d_*(G'')$ without loss of generality. Then

- $d_*(G) \geq \min\{d_*(G'), d_*(G'')\} = d_*(G')$,
- $v(G) = v(G') + 4$, and
- $c(G) = c(G')$.

Thus,

$$d_*(G) + \frac{1}{4}v(G) + c(G) \geq d_*(G') + \frac{1}{4}v(G') + c(G') + 1.$$

Suppose instead that G' is obtained by a 2-gon relation, that is, $G = [2]G'$. Then

- $d_*(G) = d_*(G') - \frac{1}{2}$,
- $v(G) = v(G') + 2$, and
- $c(G) = c(G')$.

Thus,

$$d_*(G) + \frac{1}{4}v(G) + c(G) = d_*(G') + \frac{1}{4}v(G') + c(G').$$

As mentioned above, we can choose a reduction sequence from G to Γ so that flat trivalent graphs in this sequence satisfy the above inequality for the minimum degree. Note that $d_*(\Gamma) = -c(\Gamma)$ because $v(\Gamma) = 0$. Hence, G and Γ should satisfy

$$d_*(G) + \frac{1}{4}v(G) + c(G) \geq d_*(\Gamma) + \frac{1}{4}v(\Gamma) + c(\Gamma) = 0. \quad \square$$

Next we give a lower bound on the minimum degree of a flat trivalent graph with $\mathfrak{sl}_3$ -clasps. Let us consider the minimum degree of coefficients appearing in expansion formulas of $\mathfrak{sl}_3$ -clasps.

Lemma 3.3 (the single clasp expansion formula [17, Proposition 3.1]) *For any positive integer m ,*

$$\text{JW}_{-m} = \begin{array}{c} m \\ \updownarrow \\ \updownarrow \end{array} = \sum_{j=0}^{m-1} f_j^{(m)}(q) \begin{array}{c} j \quad 1 \quad m-j-1 \\ \updownarrow \quad \updownarrow \quad \updownarrow \\ \updownarrow \quad \updownarrow \quad \updownarrow \\ 1 \quad m-1 \end{array},$$

where $f_j^{(m)}(q) := (-1)^j [m-j]/[m]$.

One can obtain the following lemma from the single clasp expansion formula and induction on m :

Lemma 3.4 *The one-row colored $\mathfrak{sl}_3$ -clasp has an expansion*

$$\text{JW}_{-m} = \sum_M f_M(q) M$$

with $d_*(f_M(q)) = \frac{1}{4}v(M)$, where the sum runs over finitely many flat trivalent graphs M , and $v(M)$ is the number of trivalent vertices in M . We remark that M may contain 4-gons or 2-gons.

Proof It is obvious that the claim is true for $m = 1, 2$. We prove it by induction on m . A flat trivalent graph in the right-hand side of the single clasp expansion of JW_{-m} has a stair-step web with $2j$ vertices and JW_{-m-1} . We know that $d_*(f_j^{(m)}(q)) = \frac{1}{2}j$ by [Example 2.12](#), [Lemma 3.3](#) and the induction hypothesis derive an expansion

$$\text{JW}_{-m} = \sum_{j=0}^{m-1} \sum_M f_j^{(m)}(q) f_M(q) \begin{array}{c} j \quad 1 \quad m-j-1 \\ \uparrow \quad \uparrow \quad \uparrow \\ \text{---} \text{---} \text{---} \\ \downarrow \quad \downarrow \quad \downarrow \\ 1 \quad m-1 \end{array} \quad ,$$

where $d_*(f_M(q)) = \frac{1}{4}v(M)$. The flat trivalent graph in the right-hand side has $2j + v(M)$ vertices, and $d_*(f_j^{(m)}(q) f_M(q)) = d_*(f_j^{(m)}(q)) + d_*(f_M(q)) = \frac{1}{4}(2j + v(M))$ by [Lemma 2.11](#). This expansion satisfies the condition of our claim. $\square$

Remark 3.5 The proof of [Lemma 3.4](#) can be used to show that M constructed by composing webs

$$I_j := \begin{array}{c} \uparrow \quad \cdots \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\ \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \\ \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\ j \quad j+1 \end{array} \in \mathcal{TL}(-^m, -^m)$$

for $j = 1, 2, \dots, m-1$. Labels j and $j+1$ in the bottom mean the j^{th} and $(j+1)^{\text{st}}$ marked points, respectively.

Lemma 3.6 The two-row colored $\mathfrak{sl}_3$ -clasp has the expansion

$$\text{JW}_{-m+n} = \sum_{t=0}^{\min\{m,n\}} \sum_{M_1, M_2, M_3, M_4} f_{(m,n;t)}(M_1, M_2, M_3, M_4; q) \begin{array}{c} \uparrow M_4 \quad M_3 \\ \text{---} \text{---} \\ \downarrow M_1 \quad M_2 \end{array}_{(m,n;t)}$$

with $d_*(f_{(m,n;t)}(M_1, M_2, M_3, M_4; q)) = \frac{1}{2}t(t+1) + \sum_{i=1}^4 \frac{1}{4}v(M_i)$, where

$$\begin{array}{c} \uparrow M_4 \quad M_3 \\ \text{---} \text{---} \\ \downarrow M_1 \quad M_2 \end{array}_{(m,n;t)} := \begin{array}{c} \begin{array}{c} \uparrow m \\ \text{---} \\ \text{---} \\ \downarrow m-t \end{array} \quad \begin{array}{c} \uparrow n \\ \text{---} \\ \text{---} \\ \downarrow n-t \end{array} \\ \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \downarrow t \end{array} \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \downarrow t \end{array} \end{array}$$

Proof Apply [Lemma 3.4](#) to each one-row colored $\mathfrak{sl}_3$ -clasp in the right-hand side of [Definition 2.5](#). Then we obtain an expansion as in the statement such that

$$f_{(m,n;t)}(M_1, M_2, M_3, M_4; q) := (-1)^t \frac{\begin{bmatrix} m \\ t \end{bmatrix} \begin{bmatrix} n \\ t \end{bmatrix}}{\begin{bmatrix} m+n+1 \\ t \end{bmatrix}} f_{M_1}(q) f_{M_2}(q) f_{M_3}(q) f_{M_4}(q).$$

One can calculate the minimum degree as

$$\begin{aligned} d_*(f_{(m,n;t)}(M_1, M_2, M_3, M_4; q)) &= d_* \left((-1)^t \frac{\begin{bmatrix} m \\ t \end{bmatrix} \begin{bmatrix} n \\ t \end{bmatrix}}{\begin{bmatrix} m+n+1 \\ t \end{bmatrix}} \right) + \sum_{i=1}^4 d_*(f_{M_i}(q)) \\ &= \frac{1}{2}t(t+1) + \sum_{i=1}^4 \frac{1}{4}v(M_i). \end{aligned}$$

$\square$

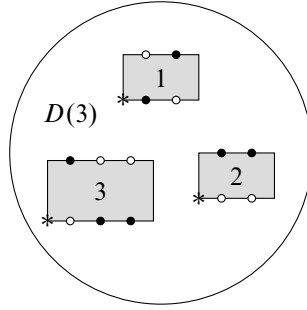


Figure 1: It is a k -holed disk $D(k)$ with $k = 3$. A shaded rectangle labeled by i is D_i . Marked points with sign $+$ (resp. $-$) are described as black (resp. white) dots. In this case, $s_1^{(0)} = \bar{s}_1^{(1)} = +- , s_2^{(0)} = \bar{s}_2^{(1)} = -- ,$ and $s_3^{(0)} = \bar{s}_3^{(1)} = -++$. A tangled trivalent graph G defines $\mathcal{TL}(-+, -+) \otimes \mathcal{TL}(++, ++) \otimes \mathcal{TL}(-+-, +--) \rightarrow \mathcal{R}$.

We introduce notation for a planar algebra specialized to our situation because it is useful for writing $\mathfrak{sl}_3$ -webs in the form of an equation. Let D be a disk, and D_i for $i = 1, 2, \dots, k$ disjoint rectangles in $D \setminus \partial D$. Each D_i is homeomorphic to $[0, 1] \times [0, 1]$ and it has a basepoint at $(0, 0)$ and marked points $P_i = P_i^{(0)} \sqcup P_i^{(1)}$. The set $P_i^{(j)}$ of marked points lies in the edge of D_i corresponding to $[0, 1] \times \{j\}$. We have a k -holed disk $D(k) := D \setminus \bigcup_{i=1}^k \text{int}(D_i)$ with marked points $\bigcup_{i=1}^k P_i$. A small disk D_i share P_i with $D(k)$ for $i = 1, 2, \dots, k$, see Figure 1. Let a sequence of signs $s_i^{(0)}$ (resp. $s_i^{(1)}$) be an assignment of signs to a set of marked points $P_i^{(0)}$ (resp. $P_i^{(1)}$) of $D(k)$ for each $i = 1, 2, \dots, k$. Then we consider the $\mathfrak{sl}_3$ -web spaces

- $\mathcal{W}(\bigcup_{i=1}^k s_i; D(k))$, where $s_i = s_i^{(0)} \cup s_i^{(1)}$;
- $\mathcal{TL}(\bar{s}_i^{(0)}, s_i^{(1)}) = \mathcal{W}(\bar{s}_i^{(0)} \cup s_i^{(1)}; D_i)$ for $i = 1, 2, \dots, k$.

As I mentioned above, $D(k)$ and D_i share the set of marked points $P_i = P_i^{(0)} \sqcup P_i^{(1)}$. Then the sequence of signs $s_i^{(j)}$ of $P_i^{(j)}$ in $D(k)$ is consistent with $\bar{s}_i^{(j)}$ of $P_i^{(j)}$ in D_i for $j = 1, 2$. For example, an edge terminating at $p \in P_i$ in $D(k)$ can be composed with an edge starting from p in D_i . Thus, the sign of p in $D(k)$ and D_i are different. For a tangled trivalent graph $G \in \mathcal{G}(\bigcup_{i=1}^k s_i; D(k))$, a linear map

$$G: \bigotimes_{i=1}^k \mathcal{TL}(\bar{s}_i^{(0)}, s_i^{(1)}) \rightarrow \mathcal{W}(\emptyset, D) \cong \mathbb{Z}((q^{1/6}))$$

is induced by a map $D(k) \sqcup (D_1 \sqcup D_2 \sqcup \dots \sqcup D_k) \rightarrow D$. This map composes $\mathfrak{sl}_3$ -webs in D_i for $i = 1, 2, \dots, k$ with G in $D(k)$.

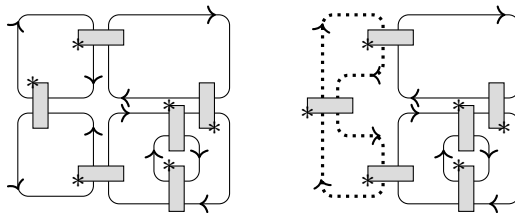
In this paper, we only consider segregated sign sequences $s_i^{(0)} = \epsilon^{m_i} \bar{\epsilon}^{n_i}$ and $s_i^{(1)} = \bar{\epsilon}^{m_i} \epsilon^{n_i}$, where ϵ is $+$ or $-$ and $m_i, n_i \in \mathbb{Z}_{\geq 0}$ satisfy $m_i + n_i = \#P_i^{(0)} = \#P_i^{(1)}$. The identity web $\mathbb{1}_{s_i^{(1)}}$ in $\mathcal{TL}(\bar{s}_i^{(0)}, s_i^{(1)})$ is $m_i + n_i$ parallel strands in D_i . The identity web $\mathbb{1}_{s_i^{(1)}}$ and the $\mathfrak{sl}_3$ -clasp $\text{JW}_{s_i^{(1)}}$ in D_i are simply denoted by $\mathbb{1}_{D_i}$ and JW_{D_i} , respectively.

Definition 3.7 Set $s_i^{(0)} = \bar{s}_i^{(1)} = \epsilon^{m_i} \bar{\epsilon}^{n_i}$ for all $i = 1, 2, \dots, k$. $G \in \mathcal{G}(\bigcup_{i=1}^k s_i; D(k))$ is *adequate* if

- G is a disjoint union of oriented simple arcs, and
- For every $j = 1, 2, \dots, k$, any pair of strands in $\mathbb{1}_{D_j}$ belong to different connected components of the graph $G(\bigotimes_{i=1}^k \mathbb{1}_{D_i})$ which is composed of oriented simple loops.

See [Example 3.8](#). We also call the clasped $\mathfrak{sl}_3$ -web $G(\bigotimes_{i=1}^k \text{JW}_{D_i})$ *adequate* when G is adequate.

Example 3.8 The left $\mathfrak{sl}_3$ -web below is adequate, but the right is not adequate because of the dotted arc:



Proposition 3.9 Let $D(k) = D \setminus \bigcup_{i=1}^k \text{int}(D_i)$ be a k -holed disk with signed marked point $s_i^{(0)} = \bar{s}_i^{(1)} = \epsilon^{m_i} \bar{\epsilon}^{n_i}$ on the i^{th} boundary component for $i = 1, 2, \dots, k$. For any flat trivalent graph G in $\mathcal{G}(\bigcup_{i=1}^k s_i; D(k))$,

$$d_* \left(G \left(\bigotimes_{i=1}^k \text{JW}_{D_i} \right) \right) \geq -\frac{1}{4}v(G) - c \left(G \left(\bigotimes_{i=1}^k \mathbb{1}_{D_i} \right) \right).$$

In particular, $d_*(G(\bigotimes_{i=1}^k \text{JW}_{D_i})) \geq d_*(G(\bigotimes_{i=1}^k \mathbb{1}_{D_i}))$ holds when G has no trivalent vertices due to [Proposition 3.2](#). Moreover, $d_*(G(\bigotimes_{i=1}^k \text{JW}_{D_i})) = d_*(G(\bigotimes_{i=1}^k \mathbb{1}_{D_i}))$ holds when G is adequate.

Proof [Lemma 3.6](#) expands all $\mathfrak{sl}_3$ -clasps in D_i for $i = 1, 2, \dots, k$ as

$$G \left(\bigotimes_{i=1}^k \text{JW}_{D_i} \right) = \sum_{t_1=0}^{\min\{m_1, n_1\}} \cdots \sum_{t_k=0}^{\min\{m_k, n_k\}} \sum_{M^{(1)}} \cdots \sum_{M^{(k)}} \left(\prod_{i=1}^k f_{t_i}(M^{(i)}; q) \right) G \left(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i} \right),$$

where

- $\sum_{M^{(i)}}$ means a summation over $M_1^{(i)}, M_2^{(i)}, M_3^{(i)}$ and $M_4^{(i)}$;
- $\uparrow M^{(i)} \downarrow_{t_i}$ is the $\mathfrak{sl}_3$ -web defined in [Lemma 3.6](#); and
- $f_{t_i}(M^{(i)}; q) := f_{(m_i, n_i; t_i)}(M_1^{(i)}, M_2^{(i)}, M_3^{(i)}, M_4^{(i)}; q)$ with $d_*(f_{t_i}(M^{(i)}; q)) = \frac{1}{2}t_i(t_i + 1) + \sum_{j=1}^4 \frac{1}{4}v(M_j^{(i)})$.

We remark that $v(G(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i})) = \sum_{i=1}^k \sum_{j=1}^4 v(M_j^{(i)}) + v(G)$ by definition. Now

$$\begin{aligned} (3-1) \quad d_* \left(\left(\prod_{i=1}^k f_{t_i}(M^{(i)}; q) \right) G \left(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i} \right) \right) \\ = \sum_{i=1}^k \frac{1}{2}t_i(t_i + 1) + \sum_{i=1}^k \sum_{j=1}^4 \frac{1}{4}v(M_j^{(i)}) + d_* \left(G \left(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i} \right) \right). \end{aligned}$$

Proposition 3.2 gives a lower bound on $d_*(G(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i}))$ using the number of vertices and connected components:

$$d_*\left(G\left(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i}\right)\right) \geq -\frac{1}{4}\left(\sum_{i=1}^k \sum_{j=1}^4 v(M_j^{(i)}) + v(G)\right) - c\left(G\left(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i}\right)\right).$$

Let $G(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i})$ be an $\mathfrak{sl}_3$ -web whose intersection with $D(k)$ is G and with each D_i is $\uparrow \mathbb{1}^{(i)} \downarrow_{t_i}$, where

$$\uparrow \mathbb{1}^{(i)} \downarrow_{t_i} := m_i - t_i \left\{ \begin{array}{c} \curvearrowright \\ t_i \\ \curvearrowleft \end{array} \right\} n_i - t_i.$$

In other words, $G(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i})$ is obtained by replacing all $M_j^{(i)}$ in $G(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i})$ with identity webs.

Remark 3.5 says that $G(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i})$ is obtained from $G(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i})$ by a sequence of *zip cobordisms* which replace parallel strands with the I_j . Then $c(G(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i})) \leq c(G(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}))$ holds because a zip cobordism reduces the number of connected components. Combining these inequalities with (3-1), we obtain

$$(3-2) \quad d_*\left(\left(\prod_{i=1}^k f_{t_i}(M^{(i)}; q)\right) G\left(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i}\right)\right) \geq \sum_{i=1}^k \frac{1}{2} t_i (t_i + 1) - \frac{1}{4} v(G) - c\left(G\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right).$$

One can make a similar argument when the JW_{D_i} have one-row colored $\mathfrak{sl}_3$ -clasps by using **Lemma 3.4**. If the i^{th} disk D_i has a one-row colored $\mathfrak{sl}_3$ -clasp, then we read $\uparrow M^{(i)} \downarrow_{t_i}$ as $M^{(i)}$ and $f_{t_i}(M^{(i)}; q)$ as $f_{M^{(i)}}(q)$, and replace $\sum_{j=1}^4 \frac{1}{4} v(M_j^{(i)})$ with $\frac{1}{4} v(M^{(i)})$.

Finally, we observe how the right-hand side of (3-2) changes by a single *orientable saddle cobordism* which transforms t_i to $t_i + 1$. One can see that the single orientable saddle cobordism changes $c(G(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}))$ by ± 1 by considering the orientation of strands in $\uparrow \mathbb{1}^{(i)} \downarrow_{t_i}$. Hence, the right-hand side of (3-2) is a monotonically increasing function on $0 \leq t_i \leq \min\{m_i, n_i\}$. Consequently,

$$\begin{aligned} d_*\left(\left(\prod_{i=1}^k f_{t_i}(M^{(i)}; q)\right) G\left(\bigotimes_{i=1}^k \uparrow M^{(i)} \downarrow_{t_i}\right)\right) &\geq -\frac{1}{4} v(G) - c\left(G\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_0\right)\right) \\ &= -\frac{1}{4} v(G) - c\left(G\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right), \end{aligned}$$

because $\uparrow \mathbb{1}^{(i)} \downarrow_0 = \mathbb{1}_{D_i}$ by definition.

When $v(G) = 0$, the right-hand side becomes $d_*(G(\bigotimes_{i=1}^k \mathbb{1}_{D_i}))$ by **Proposition 3.2**. Moreover, if G is adequate, $c(G(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}))$ strictly decreases by the orientable saddle cobordism changing t_i from 0 to 1 for some i . Hence, we obtain $d_*(G(\bigotimes_{i=1}^k JW_{D_i})) = d_*(G(\bigotimes_{i=1}^k \mathbb{1}_{D_i}))$. One can make a similar argument when all the JW_{D_i} are one-row colored $\mathfrak{sl}_3$ -clasps and G is adequate. $\square$

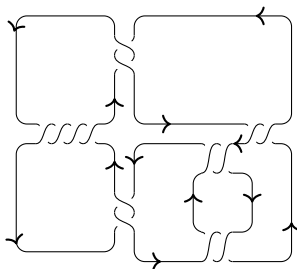
4 Zero stability of the one-row colored $\mathfrak{sl}_3$ -Jones polynomial

We show the zero stability of the one-row colored $\mathfrak{sl}_3$ -tail for a certain class of B -adequate oriented links.

Definition 4.1 Let D be a disk equipped with pairwise disjoint twist regions D_i for $i = 1, \dots, k$ in $\text{int}(D)$ such that each D_i is isomorphic to $[0, 1] \times [0, 1]$ with a basepoint at $(0, 0)$ and an assignment $I: \{D_i\}_{i=1}^k \rightarrow 2\mathbb{Z}_{>0}$. An *antiparallel B -adequate link* is an oriented link represented by an oriented link diagram L in D satisfying the following conditions:

- $L \cap D(k)$ is an adequate graph G , where $D(k) := D \setminus \bigsqcup_{i=1}^k \text{int}(D_i)$.
- $L \cap D_i$ is a twist region R_{l_i} with negative $l_i := I(D_i)$ half twists of antiparallel strands for each i ; see Figure 2.

Example 4.2 The diagram below represents an antiparallel B -adequate link:



Example 4.3 (plumbed-like links) Let X be a planar embedded graph equipped with a weight $I: E(X) \rightarrow 2\mathbb{Z}_{\geq 0}$ for the edge set $E(X)$. Then we obtain an antiparallel B -adequate link diagram from X by replacing all vertices with positively oriented circles and then adding a twist $R_{I(e)}$ between two circles connected by an edge $e \in E(X)$.

Before we prove the zero stability for antiparallel B -adequate links, let us introduce some symbols for values of special $\mathfrak{sl}_3$ -webs and coefficients. We can describe the one-row colored $\mathfrak{sl}_3$ -Jones polynomial by using these values.

Lemma 4.4 We have

$$\Delta^{(n)}(j) = \frac{[n-j+1]^2[2n-2j+2]}{[2]}, \quad \Theta^{(n)}(j) = \frac{[2n-j+2]}{[n-j]^2} \Delta^{(n)}(j),$$

$$\gamma^{(n)}(j) = (-1)^{n-j} q^{-n^2/6} q^{-j^2/2 + (n+1)j},$$

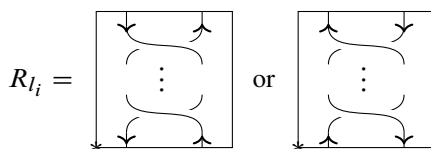


Figure 2: The diagram R_{l_i} has l_i crossings, where $l_i \in 2\mathbb{Z}_{>0}$.

where

Proof It is well known that the value of a closure of JW_{-m+n} is obtained by quantum dimension $[m+1][n+1][m+n+2]/[2]$, and it is just $\Delta_j^{(n)}$. One can compute the value γ by using [Lemma 2.9](#)(3), (4) and (6). $\Theta^{(n)}(j)$ was computed in [\[25\]](#). $\square$

Lemma 4.5 We have

$$\begin{aligned} d_*(\Delta^{(n)}(j)) &= d_*(\Delta^{(n)}(j-1)) + 2, & d_*(\Theta^{(n)}(j)) &= d_*(\Theta^{(n)}(j-1)) + 1, \\ d_*(\gamma^{(n)}(j)) &= d_*(\gamma^{(n)}(j-1)) + (n-j + \tfrac{3}{2}). \end{aligned}$$

Proof From [Lemma 2.11](#) and [Example 2.12](#), one can compute the minimum degrees of $\Delta^{(n)}(j)$ and $\Theta^{(n)}(j)$ as $d_*(\Delta^{(n)}(j)) = -2n + 2j$ and $d_*(\Theta^{(n)}(j)) = -2n + j$. $\square$

An n -parallelization L_n of the antiparallel B -adequate link diagram L defines an adequate graph $G_n := L_n \cap D(k) \in (\bigcup_{i=1}^k s_i; D(k))$, where $s_i^{(0)} = \bar{s}_i^{(1)} = \epsilon_i^n \bar{\epsilon}_i^n$ for some $\epsilon_i \in \{\pm\}$, and an n -parallelization $(R_{l_i})_n := L_n \cap D_i$ of l_i half twists for any i . Then we define a clasped $\mathfrak{sl}_3$ -web $R_{l_i}^{(n)}$ in D_i by inserting one-row colored $\mathfrak{sl}_3$ -clasps for each n parallelized strands of $(R_{l_i})_n$. The one-row colored $\mathfrak{sl}_3$ -Jones polynomial $J_{L,n}^{\mathfrak{sl}_3}(q)$ of the antiparallel B -adequate link L is obtained by replacing $(R_{l_i})_n$ of L_n with $R_{l_i}^{(n)}$. Using a linear map defined by G_n , this replacement is described as $G_n(\bigotimes_{i=1}^k (R_{l_i}^{(n)}))$.

Lemma 4.6 For an antiparallel B -adequate link diagram L with twist regions D_i ,

$$L^{(n)} = G_n\left(\bigotimes_{i=1}^k (R_{l_i}^{(n)})\right) = \sum_{t_1, t_2, \dots, t_k=0}^n \prod_{i=1}^k \Gamma^{(n)}(t_i; l_i) G_n\left(\bigotimes_{i=1}^k M(t_i; n)\right),$$

where

$$\Gamma^{(n)}(t_i; l_i) = \gamma^{(n)}(t_i)^{l_i} \frac{\Delta^{(n)}(t_i)}{\Theta^{(n)}(t_i)} \quad \text{and} \quad M(t_i; n) = \begin{array}{c} \text{[Diagram 1]} \end{array} \quad \text{or} \quad \begin{array}{c} \text{[Diagram 2]} \end{array}.$$

Proof Apply the formula

$$\begin{array}{c} \text{[Diagram]} \end{array} = \sum_{t=0}^n \frac{\Delta^{(n)}(t)}{\Theta^{(n)}(t)} M(t; n)$$

shown in [\[25\]](#) to all twist regions, and resolve twists by definition of $\gamma^{(n)}(j)$ in [Lemma 4.4](#). We obtain the desired formula. $\square$

Lemma 4.7 $d_*(\Gamma^{(n)}(t_i; l_i)) = d_*(\Gamma^{(n)}(t_i - 1; l_i)) + l_i(n - t_i + \frac{3}{2}) + 1.$

Proof By Lemmas 2.11 and 4.5,

$$\begin{aligned} d_*(\Gamma^{(n)}(t_i; l_i)) &= l_i d_*(\gamma^{(n)}(t_i)) + d_*(\Delta^{(n)}(t_i)) - d_*(\Theta^{(n)}(t_i)) \\ &= l_i d_*(\gamma^{(n)}(t_i - 1)) + d_*(\Delta^{(n)}(t_i - 1)) - d_*(\Theta^{(n)}(t_i - 1)) + l_i(n - t_i + \frac{3}{2}) + 1 \\ &= d_*(\Gamma^{(n)}(t_i - 1; l_i)) + l_i(n - t_i + \frac{3}{2}) + 1. \end{aligned} \quad \square$$

Proposition 4.8 Let L be an antiparallel B -adequate link diagram with disjoint twist regions D_i , $l: \{D_i\}_{i=1}^k \rightarrow 2\mathbb{Z}_{>0}$ and $G = L \cap D(k)$. Then

$$G_n \left(\bigotimes_{i=1}^k (R_{l_i}^{(n)}) \right) - \prod_{i=1}^k \gamma^{(n)}(0)^{l_i} G_n \left(\bigotimes_{i=1}^k \text{JW}_{D_i} \right) \in q^{2(n+2)} + d_*(L^{(n)})\mathbb{Z}[q].$$

To prove Proposition 4.8, we prepare several lemmas. First of all, let us introduce an operation S_j corresponding to the single orientable saddle cobordism at D_j for $i = 1, 2, \dots, k$. More precisely, S_j acts on the set $\{\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i} \mid 0 \geq t_i \geq n\}$ of an $\mathfrak{sl}_3$ -web in $\bigsqcup_{i=1}^k D_i$ as follows: S_j replaces $\uparrow \mathbb{1}^{(j)} \downarrow_{t_j}$ with $\uparrow \mathbb{1}^{(j)} \downarrow_{t_j-1}$, and acts as the identity on $\uparrow \mathbb{1}^{(j)} \downarrow_0$ in D_j or elements in D_i with $i \neq j$. Note that $\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}$ changes to $\bigotimes_{i=1}^k \mathbb{1}_{D_i}$ by a composition of orientable saddle cobordisms. For instance, $S_k^{t_k} \dots S_2^{t_2} S_1^{t_1}$ realizes this deformation.

Lemma 4.9 For an adequate graph G of the antiparallel B -adequate link diagram L in Proposition 4.8 and any fixed tuple $(t_1, \dots, t_k) \in \{0, \dots, n\}^k$,

$$d_* \left(G_n \left(\bigotimes_{i=1}^k M(t_i; n) \right) \right) \geq d_* \left(G_n \left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i} \right) \right).$$

Proof A clasped $\mathfrak{sl}_3$ -web $M(t_i; n)$ has five $\mathfrak{sl}_3$ -clasps in D_i . For all $i = 1, 2, \dots, k$, one can choose five small disks $D_{(i,1)}, \dots, D_{(i,5)}$ in D_i so that each small disk surrounds a single $\mathfrak{sl}_3$ -clasp. The intersection of L and $D \setminus \bigsqcup \{D_{(i,j)} \mid 1 \leq i \leq k, 1 \leq j \leq 5\}$ is a graph G' with no trivalent vertices. Then $G_n(\bigotimes_{i=1}^k M(t_i; n)) = G'_n(\bigotimes_{i=1}^k \bigotimes_{j=1}^5 \text{JW}_{D_{(i,j)}})$ by the construction of G' , where $\text{JW}_{D_{(i,j)}}$ is a one- or two-row colored clasp. Apply Proposition 3.9 to $G'_n(\bigotimes_{i=1}^k \bigotimes_{j=1}^5 \text{JW}_{D_{(i,j)}})$ to obtain

$$\begin{aligned} d_* \left(G_n \left(\bigotimes_{i=1}^k M(t_i; n) \right) \right) &= d_* \left(G'_n \left(\bigotimes_{i=1}^k \bigotimes_{j=1}^5 \text{JW}_{D_{(i,j)}} \right) \right) \\ &\geq d_* \left(G'_n \left(\bigotimes_{i=1}^k \bigotimes_{j=1}^5 \mathbb{1}_{D_{(i,j)}} \right) \right) = d_* \left(G_n \left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i} \right) \right). \end{aligned} \quad \square$$

Lemma 4.10 For an adequate graph G of the antiparallel B -adequate link diagram L in Proposition 4.8 and any fixed tuple $(t_1, \dots, t_k) \in \{0, \dots, n\}^k$ with $0 < t_j \leq n$,

$$d_* \left(\Gamma^{(n)}(t_j; l_j) G_n \left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i} \right) \right) > d_* \left(\Gamma^{(n)}(t_j - 1; l_j) G_n \left(S_j \left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i} \right) \right) \right).$$

Proof Note that $d_*(G_n(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i})) = -c(G_n(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}))$ and $d_*(G_n(S_j(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}))) = -c(G_n(S_j(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i})))$ by [Proposition 3.2](#). The orientable saddle cobordism S_j changes the number of connected components by ± 1 or 0. Hence $d_*(G_n(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i})) \geq d_*(G_n(S_j(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}))) - 1$. This inequality and [Lemma 4.7](#) imply

$$\begin{aligned} d_*\left(\Gamma^{(n)}(t_j; l_j) G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right) &= d_*(\Gamma^{(n)}(t_j; l_j)) + d_*\left(G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right) \\ &\geq d_*(\Gamma^{(n)}(t_j; l_j)) + d_*\left(G_n\left(S_j\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right)\right) - 1 \\ &= d_*(\Gamma^{(n)}(t_j - 1; l_j)) + d_*\left(G_n\left(S_j\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right)\right) + l_j(n - t_j + \frac{3}{2}) \\ &= d_*\left(\Gamma^{(n)}(t_j - 1; l_j) G_n\left(S_j\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right)\right) + l_j(n - t_j + \frac{3}{2}). \end{aligned}$$

One can easily see that $l_j(n - t_j + \frac{3}{2}) \geq 3$ because $l_j \in 2\mathbb{Z}_{>0}$ and $0 < t_j \leq n$. □

Lemma 4.11 For an adequate graph G of the antiparallel B -adequate link diagram L in [Proposition 4.8](#) and any $0 < j \leq k$,

$$d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(\delta_{ij}; l_i)\right) G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}}\right)\right) - d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right) \geq 2n + 3,$$

where δ_{ij} is the Kronecker delta function.

Proof The adequacy of G_n says that $c(G_n(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}})) - c(G_n(\bigotimes_{i=1}^k \mathbb{1}_{D_i})) = -1$. We know $d_*(\Gamma^{(n)}(1; l_j)) - d_*(\Gamma^{(n)}(0; l_j)) = l_j(n + \frac{1}{2}) + 1$ by [Lemma 4.7](#), while $d_*(G_n(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}})) = -c(G_n(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}}))$ and $d_*(G_n(\bigotimes_{i=1}^k \mathbb{1}_{D_i})) = -c(G_n(\bigotimes_{i=1}^k \mathbb{1}_{D_i}))$ by [Proposition 3.2](#). Hence,

$$\begin{aligned} d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(\delta_{ij}; l_i)\right) G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}}\right)\right) &- d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right) \\ &= d_*(\Gamma^{(n)}(1; l_j)) - d_*(\Gamma^{(n)}(0; l_j)) + d_*\left(G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}}\right)\right) - d_*\left(G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right) \\ &= l_j(n + \frac{1}{2}) + 1 - c\left(G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}}\right)\right) + c\left(G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right) \\ &= l_j(n + \frac{1}{2}) + 2 \geq 2n + 3. \end{aligned}$$

The last inequality holds because l_j is a positive even integer. □

Lemma 4.12 For an adequate graph G of the antiparallel B -adequate link diagram L in [Proposition 4.8](#),

$$d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right) = d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k \text{JW}_{D_i}\right)\right).$$

Proof This assertion comes from the adequacy of G and [Proposition 3.9](#). □

Proof of Proposition 4.8 By [Lemmas 2.11](#) and [4.9](#), we obtain

$$\begin{aligned} d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k M(t_i; n)\right)\right) &= \left(\sum_{i=1}^k d_*(\Gamma^{(n)}(t_i; l_i))\right) + d_*\left(G_n\left(\bigotimes_{i=1}^k M(t_i; n)\right)\right) \\ &\geq \left(\sum_{i=1}^k d_*(\Gamma^{(n)}(t_i; l_i))\right) + d_*\left(G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right) \\ &= d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right). \end{aligned}$$

Choose a sequence of orientable saddle cobordisms that changes $G_n(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i})$ to

$$G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_0\right) = G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right),$$

and apply [Lemma 4.10](#) to $d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right)$ along the sequence until just before the last step. We can apply [Lemma 4.11](#) to the last orientable saddle cobordism given by $\mathcal{S}_j$. This operation gives

$$\begin{aligned} d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{t_i}\right)\right) &> d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(\delta_{ij}; l_i)\right) G_n\left(\bigotimes_{i=1}^k \uparrow \mathbb{1}^{(i)} \downarrow_{\delta_{ij}}\right)\right) \\ &\geq d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right) + 2n + 3. \end{aligned}$$

The above two inequalities and [Lemma 4.12](#) imply

$$\begin{aligned} (4-1) \quad d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k M(t_i; n)\right)\right) &> d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k \mathbb{1}_{D_i}\right)\right) + 2n + 3 \\ &= d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k \text{JW}_{D_i}\right)\right) + 2n + 3 \end{aligned}$$

for any $(t_1, t_2, \dots, t_k) \neq (0, 0, \dots, 0)$.

Finally, we will compare $d_*(G_n(\bigotimes_{i=1}^k (R_{l_i}^{(n)})))$ to $d_*(\prod_{i=1}^k \gamma^{(n)}(0)^{l_i} G_n(\bigotimes_{i=1}^k M(0; n)))$ by using the expansion in [Lemma 4.6](#) and (4-1). By [Lemma 4.6](#),

$$\begin{aligned} G_n\left(\bigotimes_{i=1}^k (R_{l_i}^{(n)})\right) &= \left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k M(0; n)\right) + \sum_{(t_1, t_2, \dots, t_k) \neq (0, 0, \dots, 0)} \left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k M(t_i; n)\right) \\ &= \left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k JW_{D_i}\right) + \sum_{(t_1, t_2, \dots, t_k) \neq (0, 0, \dots, 0)} \left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k M(t_i; n)\right). \end{aligned}$$

By [Lemma 2.11](#) and (4-1), one can obtain

$$\begin{aligned} d_*\left(G_n\left(\bigotimes_{i=1}^k (R_{l_i}^{(n)})\right) - \left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k JW_{D_i}\right)\right) &= \min_{(t_1, t_2, \dots, t_k) \neq (0, 0, \dots, 0)} \left\{ d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(t_i; l_i)\right) G_n\left(\bigotimes_{i=1}^k M(t_i; n)\right)\right) \right\} \\ &> d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k JW_{D_i}\right)\right) + 2n + 3. \end{aligned}$$

Note that

$$d_*\left(G_n\left(\bigotimes_{i=1}^k (R_{l_i}^{(n)})\right)\right) = d_*\left(\left(\prod_{i=1}^k \Gamma^{(n)}(0; l_i)\right) G_n\left(\bigotimes_{i=1}^k JW_{D_i}\right)\right)$$

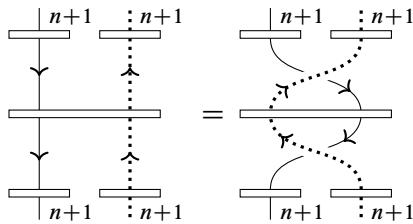
by (4-1) and $\Gamma^{(n)}(0; l_i) = \prod_{i=1}^k \gamma^{(n)}(0)^{l_i}$ by definition. □

Definition 4.13 For $f(q)$ and $g(q)$ in $\mathbb{Z}((q))$, we define $f(q) \equiv_n g(q)$ if $d_*(\hat{f}(q) - \hat{g}(q)) \geq n + 1$.

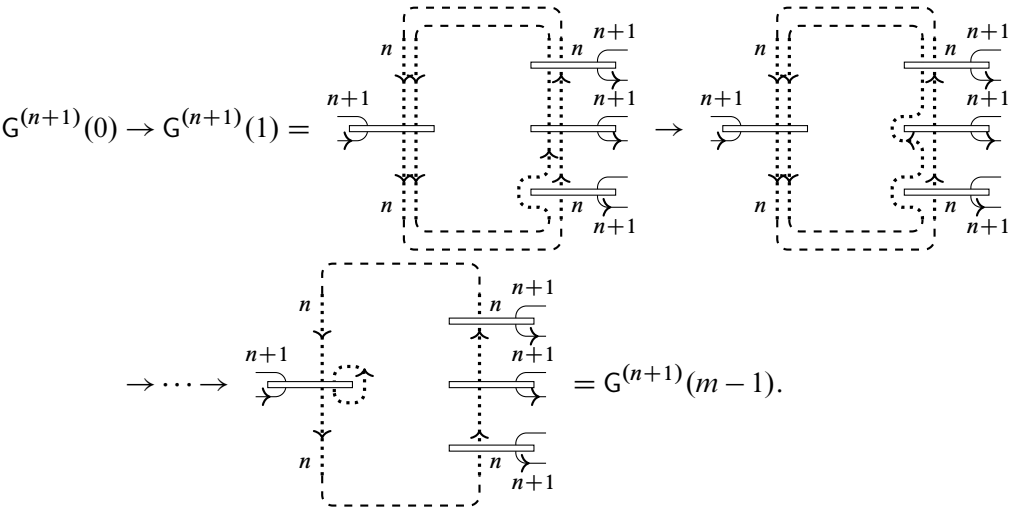
Proposition 4.14 Let JW_{D_i} represent $M(0; n)$ in D_i , as defined in [Lemma 4.6](#). If G is adequate, then

$$G_{n+1}\left(\bigotimes_{i=1}^k JW_{D_i}\right) \equiv_{n+1} G_n\left(\bigotimes_{i=1}^k JW_{D_i}\right).$$

We will prove this proposition using a similar strategy to [1]. Let us explain it in our situation. Choose a one-row colored $\mathfrak{sl}_3$ -clasp with $n + 1$ strands in $G_{n+1}(\bigotimes_{i=1}^k \mathbb{1}_{D_i})$. It corresponds to $n + 1$ parallel circles in $G(\bigotimes_{i=1}^k \mathbb{1}_{D_i})$. First we move the $n + 1$ parallel strands to the left side of the two-row colored $\mathfrak{sl}_3$ -clasps at the center of D_i . The “left side” is determined by the orientation of $n + 1$ parallel strands at each $\mathfrak{sl}_3$ -clasp:



In the above, the chosen $n + 1$ parallel strands are expressed as a dotted arc labeled by $n + 1$. We remark that this deformation of $\mathfrak{sl}_3$ -webs does not change the coefficients. We assume that the chosen $n + 1$ parallel strands pass through m two-row colored $\mathfrak{sl}_3$ -clasps $\text{JW}_{D_1}, \text{JW}_{D_2}, \dots, \text{JW}_{D_m}$ in this order by replacing labels of twist regions if necessary. We denote the initial $\mathfrak{sl}_3$ -web by $G^{(n+1)}(0)$, and a clasped $\mathfrak{sl}_3$ -web obtained by unclasp the leftmost strand of the $n + 1$ strands from $\text{JW}_{D_1}, \dots, \text{JW}_{D_{j-1}}$, and JW_{D_j} in $G^{(n+1)}(0)$ by $G^{(n+1)}(j)$ for $j = 1, 2, \dots, m - 1$. If one could unclasp the leftmost strand from $\text{JW}_{D_1}, \dots, \text{JW}_{D_{m-1}}$, then the $\mathfrak{sl}_3$ -web would become $G^{(n+1)}(m - 1)$. One can shrink the unclasp strand in $G^{(n+1)}(m - 1)$ to JW_{D_m} as follows:



We will see that the above sequence of $\mathfrak{sl}_3$ -webs can be realized by computing $G^{(n+1)}(j)$ modulo $q^{n+1} \mathbb{Z}[[q]]$.

Lemma 4.15

$$\begin{array}{c} n+1 \\ | \\ \text{strand} \\ | \\ n+1 \end{array} \begin{array}{c} \nwarrow k_j \\ \nearrow k_j \end{array} = \begin{array}{c} n+1 \\ | \\ \text{strand} \\ | \\ n+1 \end{array} \begin{array}{c} \nwarrow k_j \\ \nearrow k_j \end{array} + (-1)^{n+1} \frac{[k_j]}{[n+k_j+2]} \begin{array}{c} n+1 \\ | \\ \text{strand} \\ | \\ n+1 \end{array} \begin{array}{c} \nwarrow k_j \\ \nearrow k_j \end{array}.$$

We defer the proof of this lemma to the [appendix](#); see [Proposition A.6](#).

Proof of Proposition 4.14 Let us do the unclasp operation that we explained. Choose $n + 1$ parallel circles passing the left side of $\mathfrak{sl}_3$ -clasps $\text{JW}_{D_1}, \dots, \text{JW}_{D_m}$. We apply [Lemma 4.15](#) to the j^{th} $\mathfrak{sl}_3$ -clasp

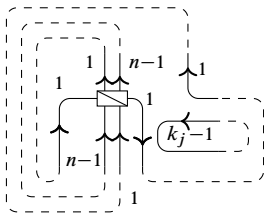


Figure 3: A flat $\mathfrak{sl}_3$ -web $\tilde{H}^{(n+1)}(j)$ obtained from $H^{(n+1)}(j)$.

JW_{D_j} in $G^{(n+1)}(j-1)$ for $j = 1, 2, \dots, m-2$. Then we obtain

$$G^{(n+1)}(j-1) = G^{(n+1)}(j) + (-1)^{n+1} \frac{[k_j]}{[n+k_j+2]} H^{(n+1)}(j),$$

where $H^{(n+1)}(j)$ is a clasped $\mathfrak{sl}_3$ -web corresponding to the second term in Lemma 4.15. We use k_j above although $k_j = n+1$ in this situation because it will be useful for later discussion. We compare $d_*(G^{(n+1)}(j))$ to $d_*(H^{(n+1)}(j))$. Let $\tilde{G}^{(n+1)}(j)$ and $\tilde{H}^{(n+1)}(j)$ denote flat trivalent graphs obtained by replacing all $\mathfrak{sl}_3$ -clasps in $G^{(n+1)}(j)$ and $H^{(n+1)}(j)$, respectively, with identity webs. Proposition 3.9 says that the lower bound on the minimum degree $d_*(H^{(n+1)}(j))$ is given by the number of vertices and connected components of $\tilde{H}$. By tracing strands of $\tilde{H}^{(n+1)}(j)$ as in Figure 3, one can see that $c(\tilde{G}^{(n+1)}(j-1)) - c(\tilde{H}^{(n+1)}(j)) = n+1$ and $v(\tilde{H}^{(n+1)}(j)) = 2n$. By Proposition 3.9,

$$\begin{aligned} d_*(H^{(n+1)}(j)) &\geq -\frac{1}{4}v(\tilde{H}^{(n+1)}(j)) - c(\tilde{H}^{(n+1)}(j)) = -\frac{1}{4}(2n) - (c(\tilde{G}^{(n+1)}(j)) - (n+1)) \\ &= \frac{1}{2}(n+2) - c(\tilde{G}^{(n+1)}(j)). \end{aligned}$$

Proposition 3.2 and $v(G^{(n+1)}(j)) = 0$ give $d_*(\tilde{G}^{(n+1)}(j)) = -c(\tilde{G}^{(n+1)}(j))$. Moreover, $d_*(G^{(n+1)}(j)) = d_*(\tilde{G}^{(n+1)}(j))$ holds by adequacy of $G^{(n+1)}(j)$ and Proposition 3.9. Using Lemma 2.11 and Example 2.12, the above facts lead to the inequality

$$d_*\left((-1)^{n+1} \frac{[k_j]}{[n+k_j+2]} \tilde{H}^{(n+1)}(j)\right) \geq \frac{1}{2}(n+2) + \left(\frac{1}{2}(n+2) + d_*(G^{(n+1)}(j))\right) = (n+2) + d_*(G^{(n+1)}(j)).$$

It also holds that $d_*(G^{(n+1)}(j-1)) = d_*(G^{(n+1)}(j))$ due to Propositions 3.2 and 3.9. In fact, the adequacy of these clasped $\mathfrak{sl}_3$ -webs and $\tilde{G}^{(n+1)}(j-1) = \tilde{G}^{(n+1)}(j)$ imply

$$\begin{aligned} d_*(G^{(n+1)}(j-1)) &= d_*(\tilde{G}^{(n+1)}(j-1)) = -c(\tilde{G}^{(n+1)}(j-1)) \\ &= -c(\tilde{G}^{(n+1)}(j)) = d_*(\tilde{G}^{(n+1)}(j)) = d_*(G^{(n+1)}(j)). \end{aligned}$$

Thus we obtain

$$G^{(n+1)}(j) - G^{(n+1)}(j-1) = (-1)^{n+1} \frac{[k_j]}{[n+k_j+2]} \tilde{H}^{(n+1)}(j) \in q^{(n+2)+d_*(G^{(n+1)})} \mathbb{Z}[[q^{1/6}]],$$

where $d_*(G^{(n+1)}) := d_*(G^{(n+1)}(j-1)) = d_*(G^{(n+1)}(j))$.

This result holds independently of the number $m-j$ of $\mathfrak{sl}_3$ -clasps that the $n+1$ parallel strands pass through, and also the number k_j of the oppositely oriented strands adjacent to the $n+1$ parallel strands.

We repeatedly apply [Lemma 4.15](#) to $\text{JW}_{D_1}, \dots, \text{JW}_{D_{m-1}}$ and obtain

$$G^{(n+1)}(0) \equiv_{n+1} G^{(n+1)}(1) \equiv_{n+1} \dots \equiv_{n+1} G^{(n+1)}(m-1).$$

Let $G^{(n+1)}(m)$ be an $\mathfrak{sl}_3$ -web removing the small circle from $G^{(n+1)}(m-1)$:

$$G^{(n+1)}(m-1) = \text{web with small circle} \quad , \quad G^{(n+1)}(m) = \text{web without small circle}.$$

Then we obtain

$$G^{(n+1)}(m-1) = \frac{[n+2][n+k_m+3]}{[n+1][n+k_m+2]} G^{(n+1)}(m)$$

by [Proposition A.1](#). From the above equality, it is easily seen that

$$G^{(n+1)}(1) \equiv_{n+1} G^{(n+1)}(0)$$

holds for any k_m . Next, we consider the leftmost strand of the other $n+1$ parallel circles. One can unclasp the leftmost strand from $\mathfrak{sl}_3$ -clasps exactly in the same way. The label k_j in this argument might be n . However, it works independently of k_j as I mentioned above. We repeatedly apply this argument until all $n+1$ parallel circles passing through $\mathfrak{sl}_3$ -clasps become n parallel strands. Consequently, we obtain $G_{n+1}(\bigotimes_{i=1}^k \mathbb{1}_{D_i}) \equiv_{n+1} G_n(\bigotimes_{i=1}^k \mathbb{1}_{D_i})$. $\square$

Theorem 4.16 *Let L be an antiparallel B -adequate link. Then*

$$\hat{J}_{L,n+1}^{\mathfrak{sl}_3}(q) - \hat{J}_{L,n}^{\mathfrak{sl}_3}(q) \in q^{n+1} \mathbb{Z}[[q]].$$

In other words, the one-row colored $\mathfrak{sl}_3$ -Jones polynomial $\{\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)\}_n$ of L is zero stable.

Proof Let us take a link diagram $G(\bigotimes_{i=1}^k R_{l_i})$ of L with an adequate graph G and twist regions $I: \{D_i\}_{i=1}^k \rightarrow 2\mathbb{Z}_{>0}$. The one-row colored $\mathfrak{sl}_3$ -Jones polynomial $\hat{J}_{L,n}^{\mathfrak{sl}_3}(q)$ is given by the normalization in [Definition 2.10](#) of the clasped $\mathfrak{sl}_3$ -web $G_n(\bigotimes_{i=1}^k (R_{l_i}^{(n)}))$. [Lemma 4.6](#) and [Proposition 4.8](#) claim that

$$G_n\left(\bigotimes_{i=1}^k (R_{l_i}^{(n)})\right) \equiv_{2n+1} \prod_{i=1}^k \gamma^{(n)}(0)^{l_i} G_n\left(\bigotimes_{i=1}^k \text{JW}_{D_i}\right),$$

$$G_{n+1}\left(\bigotimes_{i=1}^k (R_{l_i}^{(n+1)})\right) \equiv_{2n+3} \prod_{i=1}^k \gamma^{(n+1)}(0)^{l_i} G_{n+1}\left(\bigotimes_{i=1}^k \text{JW}_{D_i}\right),$$

and [Proposition 4.14](#) claims

$$G_n\left(\bigotimes_{i=1}^k \text{JW}_{D_i}\right) \equiv_{n+1} G_{n+1}\left(\bigotimes_{i=1}^k \text{JW}_{D_i}\right).$$

It is easy to see that $f(q) \equiv_n g(q)$ if $f(q) \equiv_N g(q)$ for some $N \geq n$, and $(-1)^{k_1} q^{k_2} f(q) \equiv_n (-1)^{l_1} q^{l_2} g(q)$ if $f(q) \equiv_n g(q)$ for any k_1, k_2, l_1 and l_2 . Hence, the above equivalence relations derive

$$\begin{aligned} G_n \left(\bigotimes_{i=1}^k (R_{l_i}^{(n)}) \right) &\equiv_{n+1} \prod_{i=1}^k \gamma^{(n)}(0)^{l_i} G_n \left(\bigotimes_{i=1}^k \text{JW}_{D_i} \right) \equiv_{n+1} \prod_{i=1}^k \gamma^{(n+1)}(0)^{l_i} G_{n+1} \left(\bigotimes_{i=1}^k \text{JW}_{D_i} \right) \\ &\equiv_{n+1} G_{n+1} \left(\bigotimes_{i=1}^k (R_{l_i}^{(n+1)}) \right). \end{aligned}$$

This means $\hat{J}_{L,n+1}^{\mathfrak{sl}_3}(q) - \hat{J}_{L,n}^{\mathfrak{sl}_3}(q) \in q^{n+1} \mathbb{Z}[[q^{1/6}]]$. □

Appendix Formulas for clasped $\mathfrak{sl}_3$ -webs

It is well known that the closure of JW_{-m+n} is given by

$$\Delta(m, n) = \text{Diagram of a closed web with } m \text{ and } n \text{ strands} = \frac{[m+1][n+1][m+n+2]}{[2]}.$$

Proposition A.1

$$\text{Diagram of a clasped web with } m, n, l \text{ strands} = \frac{\Delta(m+l, n)}{\Delta(m, n)} \text{Diagram of a web with } m, n \text{ strands}.$$

Proof It is known that this clasped $\mathfrak{sl}_3$ -web space is one-dimensional and it is spanned by JW_{-m+n} . Thus, we set

$$\text{Diagram of a clasped web with } m, n, l \text{ strands} = C \text{Diagram of a web with } m, n \text{ strands}.$$

The closures of the diagrams in the left- and right-hand sides are given by $\Delta(m+l, n)$ and $\Delta(m, n)$, respectively. Hence, $C = \Delta(m+l, n)/\Delta(m, n)$. □

In order to prove [Lemma 4.15](#), we prepare some lemmas.

Lemma A.2 (the bubble skein expansion formula [\[24\]](#))

$$\text{Diagram of a web with } l-a, l, l-b, a, b, k-a, k, k-b \text{ strands} = \sum_{t=\max\{a,b\}}^{\min\{a+b,k,l\}} \frac{[k]_t [l]_t [t]_a [t]_b [k+l-t+2]_{k+l-a-b+2}}{[a]_t [a]_t [b]_t [b]_t} \text{Diagram of a web with } l-a, l-t, l-b, t-a, t, t-b, k-a, k-t, k-b \text{ strands}.$$

Lemma A.3 [17, Theorem 3.3]

$$\begin{array}{c} 1 \quad k \quad l \\ \text{---} \\ 1 \quad k \quad l \end{array} = \begin{array}{c} 1 \quad k \quad l \\ \text{---} \\ 1 \quad k \quad l \end{array} - \frac{[k]}{[k+1]} \begin{array}{c} 1 \quad k \quad l \\ \text{---} \\ 1 \quad k \quad l \end{array} - \frac{[l]}{[k+1][k+l+2]} \begin{array}{c} 1 \quad k \quad l \\ \text{---} \\ 1 \quad k \quad l \end{array}.$$

Lemma A.4

$$\begin{array}{c} k \quad 1 \\ \text{---} \\ 1 \quad k \end{array} = \frac{(-1)^k}{[k+1]} \begin{array}{c} k \\ \text{---} \\ 1 \end{array}.$$

Proof Apply Lemma 3.3 to an $\mathfrak{sl}_3$ -clasp in the left side above. Then one can see that the diagrams in the expansion vanish except for the last term, due to the bottom $\mathfrak{sl}_3$ -clasp. This becomes the right-hand side by Lemma 2.9(2). $\square$

Lemma A.5

$$\begin{array}{c} k+1 \quad l \\ \text{---} \\ 1 \quad k \quad l-1 \\ \text{---} \\ k \quad l-1 \end{array} = (-1)^k \frac{[l+1]}{[k+l+1]} \begin{array}{c} k+1 \quad l \\ \text{---} \\ k \quad 1 \quad l-1 \\ \text{---} \\ k \quad l-1 \end{array}.$$

Proof It is known that the $\mathfrak{sl}_3$ -web space on a disk with clasped endpoints JW_{-l} , JW_{+k+1} and JW_{-k+l-1} is spanned by one clasped $\mathfrak{sl}_3$ -web in the right-hand side. See for example [18; 17] for details. Hence we only have to determine the coefficient C such that

$$\begin{array}{c} k+1 \quad l \\ \text{---} \\ 1 \quad k \quad l-1 \\ \text{---} \\ k \quad l-1 \end{array} = C \begin{array}{c} k+1 \quad l \\ \text{---} \\ k \quad 1 \quad l-1 \\ \text{---} \\ k \quad l-1 \end{array}.$$

Attach an $\mathfrak{sl}_3$ -web

$$\begin{array}{c} k \quad 1 \quad l-1 \\ \text{---} \\ 1 \end{array}$$

to the top of the $\mathfrak{sl}_3$ -webs in both sides. The left-hand side becomes $(-1)^k \Delta(k, l) / ([k+1] \Delta(k, l-1)) \cdot \text{JW}_{-k+l+1}$ by Lemma A.4 and Proposition A.1. The right-hand side becomes $C[k+l+2] / ([k+1][l]) \cdot \text{JW}_{-k+l+1}$ by Lemma A.2. We obtain the value C in the assertion by solving this equation. $\square$

Proposition A.6

$$\begin{array}{c} 1 \quad k \quad l \\ \text{---} \\ 1 \quad k \quad l \end{array} = \begin{array}{c} k+1 \quad l \\ \text{---} \\ 1 \quad k \quad l \end{array} + (-1)^{k+1} \frac{[l]}{[k+l+2]} \begin{array}{c} k+1 \quad l \\ \text{---} \\ 1 \quad k \quad l \end{array}.$$

Proof Let us denote the second coefficient in the assertion by $a_k := (-1)^{k+1} [l] / [k+l+2]$. We first attach an $\mathfrak{sl}_3$ -clasp

$$\begin{array}{c} k+1 \\ \text{---} \\ 1 \end{array}$$

$$(A-1) \quad -\frac{[k]}{[k+1]} \text{diagram}_1 - \frac{[l]}{[k+1][k+l+2]} \text{diagram}_2 = a_k \text{diagram}_3$$
$$\begin{aligned} & -\frac{[k]}{[k+1]}a_{k-1} + \left(-\frac{[l]}{[k+1][k+l+2]}\right)\left((-1)^k \frac{[l+1]}{[k+l+1]}\right) \\ & = (-1)^{k+1} \frac{[l]}{[k+1][k+l+1][k+l+2]}([k][k+l+2] + [l+1]) \\ & = (-1)^{k+1} \frac{[l]}{[k+l+2]} = a_k. \end{aligned}$$

(A-2)

$$= a_{k-1}$$

References

- Algebraic & Geometric Topology, Volume 25 (2025)

- [5] **O T Dasbach, X-S Lin**, *On the head and the tail of the colored Jones polynomial*, Compos. Math. 142 (2006) 1332–1342 [MR](#) [Zbl](#)
- [6] **O T Dasbach, X-S Lin**, *A volumish theorem for the Jones polynomial of alternating knots*, Pacific J. Math. 231 (2007) 279–291 [MR](#) [Zbl](#)
- [7] **M Elhamdadi, M Hajij**, *Pretzel knots and q -series*, Osaka J. Math. 54 (2017) 363–381 [MR](#) [Zbl](#)
- [8] **C Frohman, A S Sikora**, *$SU(3)$ -skein algebras and webs on surfaces*, Math. Z. 300 (2022) 33–56 [MR](#) [Zbl](#)
- [9] **S Garoufalidis, T T Q Lê**, *Nahm sums, stability and the colored Jones polynomial*, Res. Math. Sci. 2 (2015) art. id. 1 [MR](#) [Zbl](#)
- [10] **S Garoufalidis, H Morton, T Vuong**, *The SL_3 colored Jones polynomial of the trefoil*, Proc. Amer. Math. Soc. 141 (2013) 2209–2220 [MR](#) [Zbl](#)
- [11] **S Garoufalidis, S Norin, T Vuong**, *Flag algebras and the stable coefficients of the Jones polynomial*, European J. Combin. 51 (2016) 165–189 [MR](#) [Zbl](#)
- [12] **S Garoufalidis, T Vuong**, *A stability conjecture for the colored Jones polynomial*, Topology Proc. 49 (2017) 215–253 [MR](#) [Zbl](#)
- [13] **M Hajij**, *The tail of a quantum spin network*, Ramanujan J. 40 (2016) 135–176 [MR](#) [Zbl](#)
- [14] **K Kawasoe**, *The one-row-colored $\mathfrak{sl}_3$ Jones polynomials for pretzel links*, J. Knot Theory Ramifications 32 (2023) art. id. 2250105 [MR](#) [Zbl](#)
- [15] **A Keilthy, R Osburn**, *Rogers–Ramanujan type identities for alternating knots*, J. Number Theory 161 (2016) 255–280 [MR](#) [Zbl](#)
- [16] **D Kim**, *Trihedron coefficients for $\mathcal{U}_q(\mathfrak{sl}(3, \mathbb{C}))$* , J. Knot Theory Ramifications 15 (2006) 453–469 [MR](#) [Zbl](#)
- [17] **D Kim**, *Jones–Wenzl idempotents for rank 2 simple Lie algebras*, Osaka J. Math. 44 (2007) 691–722 [MR](#) [Zbl](#)
- [18] **G Kuperberg**, *Spiders for rank 2 Lie algebras*, Comm. Math. Phys. 180 (1996) 109–151 [MR](#) [Zbl](#)
- [19] **R Lawrence**, *The $PSU(3)$ invariant of the Poincaré homology sphere*, Topology Appl. 127 (2003) 153–168 [MR](#) [Zbl](#)
- [20] **T T Q Lê**, *Integrality and symmetry of quantum link invariants*, Duke Math. J. 102 (2000) 273–306 [MR](#) [Zbl](#)
- [21] **T Ohtsuki, S Yamada**, *Quantum $SU(3)$ invariant of 3-manifolds via linear skein theory*, J. Knot Theory Ramifications 6 (1997) 373–404 [MR](#) [Zbl](#)
- [22] **M Rosso, V Jones**, *On the invariants of torus knots derived from quantum groups*, J. Knot Theory Ramifications 2 (1993) 97–112 [MR](#) [Zbl](#)
- [23] **A S Sikora, B W Westbury**, *Confluence theory for graphs*, Algebr. Geom. Topol. 7 (2007) 439–478 [MR](#) [Zbl](#)
- [24] **W Yuasa**, *The $\mathfrak{sl}_3$ colored Jones polynomials for 2-bridge links*, J. Knot Theory Ramifications 26 (2017) art. id. 1750038 [MR](#) [Zbl](#)
- [25] **W Yuasa**, *A q -series identity via the $\mathfrak{sl}_3$ colored Jones polynomials for the $(2, 2m)$ -torus link*, Proc. Amer. Math. Soc. 146 (2018) 3153–3166 [MR](#) [Zbl](#)
- [26] **W Yuasa**, *Twist formulas for one-row colored A_2 webs and $\mathfrak{sl}_3$ tails of $(2, 2m)$ -torus links*, Acta Math. Vietnam. 46 (2021) 369–387 [MR](#) [Zbl](#)

*Graduate School of Science, Division of Mathematics and Mathematical Sciences, Kyoto University
Kyoto, Japan*

*Current address: International Institute for Sustainability with Knotted Chiral Meta Matter, Hiroshima University
Hiroshima, Japan*

wyuasa@hiroshima-u.ac.jp

<https://wataruyuasa.github.io/math/>

Received: 27 August 2020 Revised: 7 October 2023

Quillen homology of spectral Lie algebras with application to mod p homology of labeled configuration spaces

ADELA YIYU ZHANG

We provide a general method computing the mod p Quillen homology of algebras over the monad that parametrizes the structure of mod p homology of spectral Lie algebras, including the construction of a May-type spectral sequence when $p = 2$. This is the E^2 -page of the bar spectral sequence converging to the mod p topological Quillen homology of spectral Lie algebras. As an application, we study the mod p homology of the labeled configuration space $B_k(M; X)$ of k points in a manifold M with labels in a spectrum X , which is the mod p topological Quillen homology of a certain spectral Lie algebra by a result of Knudsen. We obtain general upper bounds for the mod p homology of $B_k(M; X)$, as well as explicit computations for $k \leq 3$. When p is odd, we observe that the mod p homology of $B_k(M^n; S^r)$ for small k depends only on the cohomology ring of the one-point compactification of M when $n + r$ is even. This supplements and contrasts with the result of Bökigheimer, Cohen and Taylor when $n + r$ is odd.

17B56, 55N20, 55R80

1. Introduction	1945
2. Preliminaries	1952
3. Computing the Quillen homology of spectral Lie algebras	1961
4. Application to the Knudsen spectral sequence	1975
5. Upper bounds and low weight computations	1980
6. Odd primary homology	1990
References	1994

1 Introduction

Spectral Lie algebras generalize the notion of Lie algebras over a field k to the $(\infty-)$ category of spectra. They are parametrized by the spectral Lie operad $s\mathcal{L}$, whose underlying symmetric sequence $\{\partial_n(\text{Id})\}_n$ is given by the Goodwillie derivatives of the identity functor on the category of pointed spaces. The spectral Lie operad is Koszul dual to the nonunital $\mathbb{E}_\infty$ -operad (Ching [23]). The homology operad $\{H_*(\partial_n(\text{Id}); k)\}_n$ of the spectral Lie operad recovers the ordinary Lie operad over k up to a shift (Ginzburg and Kapranov [31]; Fresse [30]; Ching [23]).

A natural next step is to study the mod p topological Quillen homology

$$\mathrm{TQ}_*^{s\mathcal{L}}(L; \mathbb{F}_p) := \pi_*(|\mathrm{Bar}_\bullet(\mathrm{id}, s\mathcal{L}, L)| \otimes \mathbb{F}_p)$$

of a spectral Lie algebra L , defined in analogy to the mod p topological André–Quillen homology of nonunital $\mathbb{E}_\infty$ -algebras introduced by Basterra [7]. One approach is to use the classical bar spectral sequence

$$(1) \quad E_{s,t}^2 = \pi_s \pi_t \mathrm{Bar}_\bullet(\mathrm{id}, s\mathcal{L}, L \otimes \mathbb{F}_p) \Rightarrow \mathrm{TQ}_{s+t}^{s\mathcal{L}}(L; \mathbb{F}_p),$$

obtained by skeletal filtration of the geometric realization.

To compute the E^2 -page, it is necessary to understand the structure of the mod p homology of spectral Lie algebras. Behrens [9] constructed Dyer–Lashof-type unary operations $\bar{Q}^j$ of degree $j - 1$ on the mod 2 homology of connective spectral Lie algebras and determined the relations among these operations. Building on the work of Behrens, Antolín Camarena [2] identified the monad $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ that parametrizes natural operations on the mod 2 homology of spectral Lie algebras. An algebra over $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ is an unstable module over the algebra $\bar{\mathcal{R}}$ of Behrens’ operations, along with a shifted Lie algebra structure such that brackets of operations always vanish and the self-bracket on an element x is identified with the bottom nonvanishing operation $\bar{Q}_0 := \bar{Q}^{|x|}$ on x . Following their approach, Kjaer [37] constructed Dyer–Lashof-type unary operations $\bar{\beta}^\epsilon \bar{Q}^j$ on the mod p homology of spectral Lie algebras for $p > 2$ and proved that brackets of operations always vanish. Konovalov [40] computed the relations among unary operations by studying differentials in an algebraic Goodwillie spectral sequence and thus determined the entire structure of operations in the odd primary case.

Hence the E^2 -page of the bar spectral sequence (1) associated to a spectral Lie algebra L is equivalent to the Quillen homology

$$\mathrm{HQ}_*^{\mathrm{Lie}_{\bar{\mathcal{R}}}^s}(H_*(L; \mathbb{F}_p)) := \pi_*(\mathrm{Bar}_\bullet(\mathrm{id}, \mathrm{Lie}_{\bar{\mathcal{R}}}^s, H_*(L; \mathbb{F}_p)))$$

of the $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $H_*(L; \mathbb{F}_p)$ when $p = 2$. Equivalently, the E^2 -page is obtained by applying to $H_*(L; \mathbb{F}_p)$ the total left derived functor $\pi_*(\mathbb{L} Q_{\mathrm{Mod}_{\mathbb{F}_p}}^{\mathrm{Lie}_{\bar{\mathcal{R}}}^s}(-))$ of the indecomposable functor

$$Q_{\mathrm{Mod}_{\mathbb{F}_p}}^{\mathrm{Lie}_{\bar{\mathcal{R}}}^s} : \mathrm{Lie}_{\bar{\mathcal{R}}}^s \rightarrow \mathrm{Mod}_{\mathbb{F}_p}$$

on the category of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras which heuristically kills the $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebra structure on a $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebra.

The main challenge in computing the Quillen homology of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras when $p = 2$ arises from the identification of the self-bracket with the bottom operation $\bar{Q}_0$, which precludes a factorization of the free $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebra functor as a composite of the free $\mathrm{Lie}_{\mathbb{F}_2}^s$ -algebra functor followed by the free $\bar{\mathcal{R}}$ -algebra functor. Furthermore, since the category of $\mathrm{Lie}_{\mathbb{F}_2}^s$ -algebras is nonabelian, we cannot resort to a Grothendieck spectral sequence.

The method we use involves bounding the Quillen homology of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras by the Quillen homology of variants of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras whose unary and binary operations are disentangled. An upper bound is obtained

by constructing a May-type spectral sequence with respect to the length filtration of the homogeneous algebra $\bar{\mathcal{R}}$ in [Theorem 3.6](#). To compute the May E^1 -page we further construct an algebraic Bockstein spectral sequence, whose E^1 -page is given by the Quillen homology of a variant of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras ([Definition 3.1](#)). A lower bound can be produced by mapping into the Quillen homology of another variant of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras ([Definition 3.21](#)). Then we apply a general result about factoring bar constructions against a composite monad ([Lemma 2.13](#)), which generalizes [\[17, Proposition 4.19\]](#) by Brantner, Hahn and Knudsen, in that we replace the bar construction computing the Quillen homology of these variant algebras with a smaller complex obtained as the total complex of a double complex in [Corollary 3.10](#) and [Lemma 3.24](#).

To compute the homotopy groups of these total complexes, we utilize the machinery of Koszul duality for additive Koszul algebras (Priddy [\[51\]](#)) and Lie algebras ([\[17\]](#); Chevalley and Eilenberg [\[22\]](#); May [\[45\]](#); [\[51\]](#)), as well as explicit understanding of the Bousfield–Cartan–Dwyer operations

$$\gamma_i : \pi_{h+r,i}(\Lambda^h(V_\bullet)) \rightarrow \pi_{2h+1+r+i,2t-1}(\Lambda^{2h+1}(V_\bullet)), \quad 1 \leq i \leq r$$

on the homotopy groups of the free simplicial shifted graded exterior algebra $\Lambda^\bullet(V_\bullet)$ on a simplicial $\mathbb{F}_2$ -module $V_\bullet$ (Bökstedt and Ottosen [\[13\]](#); Bousfield [\[14\]](#); Cartan [\[20\]](#); Dwyer [\[27\]](#); Haugseng and Miller [\[34\]](#)). Thus we obtain general upper bounds for the Quillen homology of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras via [Theorem 3.6](#) and [Corollary 3.16](#), as well as precise formulae in low weights in [Corollary 3.7](#).

Furthermore, we are able to provide a full computation of the Quillen homology of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras in universal cases. Denote by $\mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Mod}_{\bar{\mathcal{R}}}}$ the free allowable $\bar{\mathcal{R}}$ -module functor. The category $\mathrm{Mod}_{\bar{\mathcal{R}}}$ is stable under the desuspension functor Σ^{-1} of $\mathbb{F}_2$ -modules. Then for $1 \leq n \leq \infty$, the $\bar{\mathcal{R}}$ -module $\Sigma^{-n} \mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2)$ is an $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebra whose $\mathrm{Lie}_{\mathbb{F}_2}^s$ -structure is trivial. Note that when $n = \infty$, this is the trivial $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\mathrm{colim}_{i \rightarrow \infty} \Sigma^{-n} \mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2) \simeq \Sigma^k \mathbb{F}_2$.

Theorem 1.1 ([Theorem 3.28](#)) *The Quillen homology*

$$\mathrm{HQ}_{*,*}^{\mathrm{Lie}_{\bar{\mathcal{R}}}^s}(\Sigma^{-n} \mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2)) \cong \pi_{*,*} \mathrm{Bar}_\bullet(\mathrm{id}, \mathrm{Lie}_{\bar{\mathcal{R}}}^s, \Sigma^{-n} \mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2))$$

of the $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\Sigma^{-n} \mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2)$, $1 \leq n \leq \infty$ is isomorphic as a bigraded vector space to the shifted graded exterior algebra over $\mathbb{F}_2$ on generators $\gamma_I \bar{Q}_J(x_k)$ satisfying the following conditions:

- (1) $I = (i_1, \dots, i_m)$ satisfies $i_l \geq 2i_{l+1}$ for $l < m$, $i_m \geq 2$, and $i_1 - i_2 - \dots - i_m \leq r$;
- (2) $J = (j_1, \dots, j_r)$ satisfies $0 \leq j_l \leq j_{l+1} + 1$ for $l < r$, $0 \leq j_r < n$, and if $j_1 = 0$ then either $r = 1$ or $i_m = 2$.

Note in particular that natural operations on a class of degree k in the Quillen homology of $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras are given by the Quillen homology of the trivial $\mathrm{Lie}_{\bar{\mathcal{R}}}^s$ -algebras $\Sigma^k \mathbb{F}_2$, and the above theorem gives us a dimension count.

Application to labeled configuration spaces The second half of the paper makes use of the computation of the Quillen homology of $\mathrm{Lie}_{\mathcal{R}}^s$ -algebras to study the mod p homology of the labeled configuration spectrum

$$B_k(M, X) := \Sigma_+^\infty \mathrm{Conf}_k(M) \otimes_{h\Sigma_k} X^{\otimes k}$$

of k points in a parallelizable manifold M with labels in a spectrum X .

The study of labeled configuration spaces dates back to as early as Segal [54] and McDuff [48] as generalizations of the unordered configuration space of k points in M . The rational homology groups of labeled configuration spaces are relatively well understood in cases of interests via classical methods; see for instance Bökigheimer and Cohen [10]; Bökigheimer, Cohen and Taylor [12]; Kříž [42]; Totaro [55]; Félix and Thomas [29].

In contrast, the mod p homology groups of these objects have remained mostly intractable. Classically, the only known cases are the following:

- $M = \mathbb{R}^\infty$ with arbitrary labeling spectra by May [46; 47] and Steinberger [18, III] who built on the work of Adem [1], Kudo and Araki [43], Cartan [19], Dyer and Lashof [28], and $M = \mathbb{R}^n$ by Cohen [24, III]. Then $\bigoplus_{k \geq 0} B_k(M; X)$ is the free $\mathbb{E}_n$ -algebra on X . Its mod p homology is captured by the Dyer–Lashof operations, a polynomial product, and the Browder brackets as a functor of $H_*(X; \mathbb{F}_p)$.
- Arbitrary manifold M with labeling spectrum $X = \Sigma^\infty S^r$, where either $p = 2$ or $p > 2$ and $n + r$ is odd (Bökigheimer, Cohen and Taylor [12]; Milgram and Löffler [49]; Bökigheimer, Cohen and Milgram [11]). In these cases, there is a homology decomposition

$$(2) \quad H_*\left(\bigoplus_{k \geq 0} B_k(M; S^r); \mathbb{F}_p\right) \cong \bigotimes_i H_*(\Omega^i S^{n+r}; \mathbb{F}_p)^{\otimes \dim H_i(M)}.$$

In particular, the homology depends only on the $\mathbb{F}_p$ -module $H_*(M; \mathbb{F}_p)$.

The most recent developments in the computation of the homology of labeled configuration spaces originate from a result of Knudsen [39]. Using the machinery of factorization homology, he established an equivalence of spectra

$$(3) \quad \bigoplus_{k \geq 1} B_k(M; V) \simeq |\mathrm{Bar}_\bullet(\mathrm{id}, s\mathcal{L}, \mathrm{Free}^{s\mathcal{L}}(\Sigma^n V)^{M^+})|.$$

Here M is a parallelizable n -manifold, $s\mathcal{L}$ is the monad associated to the free spectral Lie algebra functor $\mathrm{Free}^{s\mathcal{L}}$, and $(-)^{M^+}$ the cotensor with the one-point compactification of M in the ∞ -category of spectral Lie algebras. A rational version of this equivalence was proved by Ayala and Francis [6].

Knudsen’s result opens up a path for extracting information about the homology of labeled configuration spaces. Knudsen [38] provided a general formula for the Betti numbers of unordered configuration spaces by observing that the bar spectral sequence with rational coefficients for the bar construction (3) collapses

at the E^2 -page. Building on Knudsen's work, Drummond-Cole and Knudsen [25] computed the Betti numbers of unordered configuration spaces of surfaces, vastly improving on earlier works including the case of once-punctured orientable surface by Bökigheimer and Cohen [10]. Brantner, Hahn, and Knudsen [17] studied the Knudsen spectral sequence with coefficients in Morava E -theory at an odd prime using Brantner's results [16] on the structure of the Morava E -theory of spectral Lie algebras. They computed the weight p part of the labeled configuration spaces in $\mathbb{R}^n$ and punctured genus g surfaces $\Sigma_{g,1}$ for $g \geq 1$ with coefficient in a sphere. By letting the height go to infinity, they observed that the integral homology of $B_p(\Sigma_{g,1})$ is p -power-torsion free for any odd prime p .

In this paper, we adapt their approach and study the mod p homology of $B_k(M, X)$ for M a parallelizable n -manifold and X any spectrum by examining the mod p Knudsen spectral sequence, ie the bar spectral sequence (1) with coefficients in $\mathbb{F}_p$ applied to the bar construction (3).

When $p = 2$, our general understanding of the E^2 -page, ie the Quillen homology of $\mathrm{Lie}_{\mathcal{R}}^s$ -algebras, allows us to obtain an upper bound for $H_*(B_k(M, X); \mathbb{F}_2)$ in Theorem 5.5 for arbitrary parallelizable manifold M and spectrum X . In the universal case $M = \mathbb{R}^\infty$ and $X = \mathbb{S}^r$, the bar spectral sequence has E^2 -page given by Theorem 3.28. Comparing with the computation of the homology of free $\mathbb{E}_\infty$ -algebras (Adem [1]; Dyer and Lashof [28]; May [46]; Bruner, May, McClure and Steinberger [18]), we see that there are infinitely many higher differentials and conjecture the following universal pattern, which can be verified in low weights by sparsity arguments:

Conjecture 1.2 (Conjecture 4.5) *Each page of the spectral sequence*

$$E_{s,t}^2 = \mathrm{HQ}_{s,t}^{\mathrm{Lie}_{\mathcal{R}}^s}(\Sigma^k \mathbb{F}_2) \Rightarrow \pi_{s+t} \mathrm{Bar}_\bullet(\mathrm{id}, s\mathcal{L}, \Sigma^k \mathbb{F}_2)$$

is an exterior algebra. The higher differentials act on the exterior generators of the E^2 -page as follows, see Figure 1:

(1) For an exterior generator $\alpha = \bar{Q}_{j_1} \cdots \bar{Q}_{j_m}(x_k)$ on the E^2 -page, we have

$$d^{r+1} \gamma_{r+1}(\alpha) = \bar{Q}_r(\alpha)$$

for $r < m$ and $r \leq j_1 + 1$.

(2) For an exterior generator $\beta = \gamma_{n+1} \bar{Q}_{j_1} \cdots \bar{Q}_{j_{m-1}}(x_k)$ on the E^2 -page, we have

$$(a) \quad d^{n+1}(\beta) = \bar{Q}_n \bar{Q}_{j_1} \cdots \bar{Q}_{j_{m-1}}(x_k),$$

$$(b) \quad d^{n+1} \gamma_{m+n}(\beta) = d^{n+1}(\beta) \otimes \beta,$$

$$(c) \quad \gamma_l d^{n+1}(\beta) = d^{2n+1} \gamma_{n+l-1}(\beta) \text{ for } n+3 \leq l \leq m.$$

These generate all higher differentials under further applications of the γ_i operations in accordance with (2)(b) and (2)(c), as well as the exterior product.

Remark 1.3 While the pattern of universal differentials is similar to classical ones studied by Dwyer [26] and Turner [56], the operations $\bar{Q}_j$ on coalgebras over the comonad $\pi_{*,*} \mathrm{Bar}_\bullet(\mathrm{id}, \mathrm{Lie}_{\mathcal{R}}^s, -)$ increase

filtration and hence cannot be constructed directly at the chain level, see [Remark 4.6](#). In forthcoming work with Robert Burklund and Andrew Senger, we construct a suitable deformation of the comonad associated to the bar construction $|\text{Bar}_\bullet(\text{id}, s\mathcal{L}, -)|$ to the ∞ -category of Postnikov-connective filtered $\mathbb{F}_2$ -modules and use the comonad structure to prove the conjectured pattern of differentials.

For small k , we use sparsity arguments to show that the weight k part of the Knudsen spectral sequence with $\mathbb{F}_2$ coefficients always collapses on the E^2 -page. Thus we obtain an $\mathbb{F}_2$ -bases of $H_*(B_k(M; X); \mathbb{F}_2)$ for any parallelizable manifold M and spectrum X when $k = 2$ in [Corollary 5.6](#), and for closed parallelizable M when $k = 3$ in [Corollary 5.8](#). In particular, we observe that the $\mathbb{F}_2$ -module $H_*(B_k(M; X))$ depends on and only on the cohomology ring $H^*(M^+; \mathbb{F}_2)$ when $H_*(X; \mathbb{F}_2)$ has at least two generators. This is in contrast to the case when $X = \mathbb{S}^r$, in that the equivalence (2) depends only on the $\mathbb{F}_2$ -module $H^*(M; \mathbb{F}_2)$ [12]. As examples, we produce explicit bases for $H_*(B_k(M, X); \mathbb{F}_2)$, $k = 2, 3$ when X is an arbitrary spectrum, and M is a closed torus or a punctured genus g surface in [Section 5.3](#), as well as the (punctured) real projective space $\mathbb{R}\mathbb{P}^3$ in [Section 5.4](#).

When $p > 2$, we compute the weight $k \leq p$ part of the E^2 -page of the Knudsen spectral sequence with $\mathbb{F}_p$ -coefficients in terms of certain $\text{Lie}_{\mathbb{F}_p}^s$ -algebra homology in [Proposition 6.7](#). We deduce the existence of a single d_{p-2} -differential in the Knudsen spectral sequence when $M = \mathbb{R}^n$ with $n \geq 1$, $X = \mathbb{S}^{2l}$ and $k = p \geq 5$ in [Proposition 6.8](#). Then we show that the mod p Knudsen spectral sequence collapses when $k = 2$ or $k = 3$ and $p \geq 5$.

Corollary 1.4 ([Corollary 6.9](#)) *Let M^n be a parallelizable manifold and X any spectrum. Let $\mathfrak{g}$ be the $\text{Lie}_{\mathbb{F}_p}^s$ -algebra $\tilde{H}^*(M^+; \mathbb{F}_p) \otimes \text{Lie}_{\mathcal{R}}^s(\Sigma^n H_*(X; \mathbb{F}_p))$ with brackets given by*

$$[y \otimes x, y' \otimes x'] := (y \cup y') \otimes [x, x'],$$

and $\text{CE}(\mathfrak{g})$ the shifted Chevalley–Eilenberg complex ([Definition 6.6](#)). Denote by $\text{wt}_n(H_{s,t}(\text{CE}(\mathfrak{g})))$ the weight n part of the $\text{Lie}_{\mathbb{F}_p}^s$ -algebra homology of $\mathfrak{g}$, whose weight grading is induced by regarding $H_*(X; \mathbb{F}_p)$ as an $\mathbb{F}_p$ -module in weight 1.

(1) *For all i , there is an isomorphism of $\mathbb{F}_p$ -modules*

$$H_i(B_2(M; X); \mathbb{F}_p) \cong \bigoplus_{s+t=i} \text{wt}_2(H_{s,t}(\text{CE}(\mathfrak{g}))).$$

(2) *If $p \geq 5$, then for all i*

$$H_i(B_3(M; X); \mathbb{F}_p) \cong \bigoplus_{s+t=i} \text{wt}_3(H_{s,t}(\text{CE}(\mathfrak{g}))).$$

Remark 1.5 When $X = \mathbb{S}^r$ and $k = 2, 3$, we observe that the $\mathbb{F}_p$ -module $H_*(B_k(M; \mathbb{S}^r); \mathbb{F}_p)$ depends on and only on the cohomology ring $H^*(M^+; \mathbb{F}_p)$ when $r + l$ is even; see [Remark 6.10](#). This is again in contrast to the case when $r + l$ is odd in the equivalence (2) [12].

In follow-up work with Matthew Chen [21], we build on the odd primary method in this paper and the Drummond-Cole–Knudsen computation [25] of the rational homology of the unordered configurations space $B_k(\Sigma_g; \mathbb{S})$ to identify the higher differentials in the Knudsen spectral sequence for $H_*(B_k(\Sigma_g; \mathbb{S}); \mathbb{F}_p)$. As a result, we show that for $p \geq 3$, the integral homology of $B_k(M; \mathbb{S})$ has no p -power torsion for M a closed torus or a punctured genus $g \geq 1$ surface when $k \leq p$. In particular, the latter case serves as a simpler proof to Brantner, Hahn and Knudsen’s [17, Theorem 1.10].

1.1 Outline

In Section 2, we recall the definition of spectral Lie algebras and the structure of their mod 2 homology as algebras over the monad $\mathrm{Lie}_{\mathcal{R}}^s$. Then we define the Quillen homology of $\mathrm{Lie}_{\mathcal{R}}^s$ -algebras and the mod p topological Quillen homology of spectral Lie algebras. The two are related by a bar spectral sequence.

In Section 3, we provide general upper bounds for the Quillen homology of $\mathrm{Lie}_{\mathcal{R}}^s$ -algebras and a precise formula in low weights by comparing with the Quillen homology of two variant algebras when $p = 2$. Then we explicitly compute the Quillen homology of the $\mathrm{Lie}_{\mathcal{R}}^s$ -algebras $\Sigma^k \mathbb{F}_2$ and $\Sigma^{-n} \mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Mod}_{\mathcal{R}}}(\Sigma^{n+k} \mathbb{F}_2)$.

In Section 4, we review Knudsen’s result that expresses labeled configuration spaces in parallelizable manifolds as topological Quillen objects of certain spectral Lie algebras. In the universal case $M = \mathbb{R}^\infty$, we conjecture patterns of higher differentials.

In Section 5, we apply our understanding of the Quillen homology of $\mathrm{Lie}_{\mathcal{R}}^s$ -algebras to extract explicit information about the mod 2 homology of labeled configuration spaces, including general upper bounds and low weight computations. Then we extend the methods to $p > 2$ to study the odd primary homology of labeled configuration spaces in Section 6.

1.2 Conventions

We assume that every object is graded and weighted whenever it makes sense. For instance, $\mathrm{Mod}_{\mathbb{F}_p}$ stands for the ordinary category of weighted graded $\mathbb{F}_p$ -modules. A weighted graded $\mathbb{F}_2$ -module $M_\bullet$ is an $\mathbb{N}$ -indexed collection of $\mathbb{Z}$ -graded $\mathbb{F}_p$ -modules $\{M(w)_\bullet\}_{w \in \mathbb{N}}$. The weight grading of an element $x \in M(w)_n$ is w , and the internal grading is $|x| = n$. Morphisms are weight preserving morphisms of graded $\mathbb{F}_p$ -modules. The Day convolution $\otimes$ makes $\mathrm{Mod}_{\mathbb{F}_p}$ a symmetric monoidal category. The Koszul sign rule $x \otimes y = (-1)^{|x||y|} y \otimes x$ for the symmetric monoidal product $\otimes$ depends only on the internal grading and not the weight grading.

Similarly, a shifted Lie algebra L over $\mathbb{F}_p$ is a weighted graded $\mathbb{F}_p$ -module equipped with a shifted Lie bracket $[-, -]: L_m \otimes L_n \rightarrow L_{m+n-1}$ that adds weights, as well as satisfying graded commutativity $[x, y] = (-1)^{|x||y|} [y, x]$ and the graded Jacobi identity

$$(-1)^{|x||z|} [x, [y, z]] + (-1)^{|y||x|} [y, [z, x]] + (-1)^{|z||y|} [z, [x, y]] = 0.$$

When $p = 3$, we further require that $[[x, x], x] = 0$ for all $x \in L$. Denote by $\mathrm{Lie}_{\mathbb{F}_p}^s$ the category of shifted weighted graded Lie algebras over $\mathbb{F}_p$, as well as the monad associated to the free $\mathrm{Lie}_{\mathbb{F}_p}^s$ -algebra functor.

When $p = 2$, we use the abbreviation $\mathrm{Lie}^s = \mathrm{Lie}_{\mathbb{F}_2}^s$. We further consider the category $\mathrm{Lie}^{s, \mathrm{ti}}$ of totally isotropic Lie^s -algebras, ie Lie^s -algebras that have vanishing self-brackets. We use the notation $\langle -, - \rangle$ exclusively for $\mathrm{Lie}^{s, \mathrm{ti}}$ brackets.

We mean by shifted graded exterior algebra over $\mathbb{F}_p$ a graded $\mathbb{F}_p$ -module $M_\bullet$ together with a graded commutative product $M_m \wedge M_n \rightarrow M_{m+n-1}$ such that $x \wedge x = 0$ for all $x \in M_\bullet$. We will often omit the adjectives shifted graded for the exterior algebra.

We use $\mathbb{F}_p$ for both the field $\mathbb{F}_p$ and its Eilenberg–MacLane spectrum. The coefficients for the homology group $H_*(-)$ is $\mathbb{F}_2$ unless specifically stated.

We use $\pi_n(-)$ to denote the following functors: the functor taking the n^{th} homotopy group of a spectrum, an $\mathbb{F}_p$ -module spectrum, or a simplicial $\mathbb{F}_p$ -module, as well as the functor taking the n^{th} homology group of a chain complex over $\mathbb{F}_p$.

We use $\pi_{*,*}(-)$ to denote the functor taking the bigraded homotopy groups of a (weighted graded) bisimplicial $\mathbb{F}_p$ -module, which is equivalent to taking the homology groups of the total complex of the associated double complex via the generalized Eilenberg–Zilber theorem. The bidegree (s, t) is given by the pair (simplicial degree, internal degree).

Acknowledgements

The author would like to thank Jeremy Hahn and Haynes Miller for many discussions and encouragement, Lukas Brantner, Nir Gadish, Mike Hopkins, Ben Knudsen, Nikolai Konovalov, and Andrew Senger for helpful conversations, as well as the anonymous referees for their patience and detailed comments.

The author was partially supported by NSF Grant DMS-1906072 and the Danish National Research Foundation through the Copenhagen Centre for Geometry and Topology (DNRF151) during the course of this work.

2 Preliminaries

2.1 The spectral Lie operad

We begin with a brief review of the spectral Lie operad. Ching [23] and Salvatore [53] showed that the Goodwillie derivatives $\partial_n(\mathrm{Id})$ of the identity functor $\mathrm{Id}: \mathrm{Top}_* \rightarrow \mathrm{Top}_*$ form an operad $s\mathcal{L} := \{\partial_n(\mathrm{Id})\}_n$ in Spectra. This operad is Koszul dual to the nonunital commutative operad $\mathbb{E}_\infty^{\mathrm{nu}}$ via the operadic bar construction

$$s\mathcal{L} \simeq \mathbb{D} \mathrm{Bar}(1, \mathbb{E}_\infty^{\mathrm{nu}}, 1).$$

For a description of the operadic bar construction, see [23] for a topological model using trees and [16, Appendix D] for an ∞ -categorical construction along with a comparison with the topological model.

The n^{th} -derivative $\partial_n(\mathrm{Id})$ admits an explicit description due to Arone and Mahowald [5], following the work of Johnson [35]. Let $\mathcal{P}_n$ be the poset of partitions of the set $\underline{n} = \{1, 2, \dots, n\}$ ordered by

refinements, equipped with a Σ_n -action induced from that on $\underline{n}$. Denote by $\hat{0}$ the discrete partition and $\hat{1}$ the partition $\{\underline{n}\}$. Set $\Pi_n = \mathcal{P}_n - \{\hat{0}, \hat{1}\}$. Regarding a poset $\mathcal{P}$ as a category, we obtain via the nerve construction a simplicial set $N_\bullet(\mathcal{P})$. The *partition complex* $\Sigma|\Pi_n|^\diamond$, the reduced-unreduced suspension of the realization $|\Pi_n|$, is modeled by the simplicial set

$$N_\bullet(\mathcal{P}_n)/(N_\bullet(\mathcal{P}_n - \hat{0}) \cup N_\bullet(\mathcal{P}_n - \hat{1}))$$

for $n \geq 2$ and the simplicial 0-circle S^0 for $n = 1$. Then there is an equivalence

$$\partial_n(\text{Id}) \simeq \mathbb{D}(\Sigma|\Pi_n|^\diamond)$$

of spectra with Σ_n -action, where $\mathbb{D}$ denotes the Spanier–Whitehead dual of a spectrum.

2.2 Operations on the mod 2 homology of spectral Lie algebras

Next, we describe the structure on the mod 2 homology of an algebra L over the spectral Lie operad. It consists of a Lie^s -algebra structure along with Dyer–Lashof like unary operations.

The second structure map of a spectral Lie algebra L is given by

$$\xi: \partial_2(\text{Id}) \otimes_{h\Sigma_2} L^{\otimes 2} \simeq \partial_2(\text{Id}) \otimes L_{h\Sigma_2}^{\otimes 2} \simeq \mathbb{S}^{-1} \otimes L_{h\Sigma_2}^{\otimes 2} \rightarrow L.$$

At the level of homology, this gives rise to a shifted Lie bracket

$$[-, -]: H_m(L) \otimes H_n(L) \rightarrow H_{m+n-1}(L),$$

making $H_*(L)$ a graded shifted Lie algebra [2, Proposition 5.2].

For L a connective spectral Lie algebra, Behrens defined unary operations of weight 2,

$$\bar{Q}^j: H_d(L) \rightarrow H_{d+j-1}(L),$$

on the mod 2 homology of L via $x \mapsto \xi_* \sigma^{-1} Q^j(x)$, where $Q^j: H_d(L) \rightarrow H_{d+j}(L_{h\Sigma_2}^{\otimes 2})$ is an extended Dyer–Lashof operation $x \mapsto e_{j-d} \otimes x \otimes x$, $\sigma^{-1}: H_*(L_{h\Sigma_2}^{\otimes 2}) \rightarrow H_{*-1}(\partial_2(\text{Id}) \otimes L_{h\Sigma_2}^{\otimes 2})$ is the desuspension isomorphism, and ξ is the second structure map [9, Section 1.5; 2, Definition 5.4]. Furthermore, Behrens showed that the quadratic relations

$$(4) \quad \bar{Q}^r \bar{Q}^s = \sum_{l=0}^{r-s-1} \binom{r-2l-1}{s-l} \bar{Q}^{r+s-l} \bar{Q}^l$$

for $s < r \leq 2s$ generate all the relations among the unary operations on a class in some positive degree [9, Theorem 1.5.1]. By definition, for x a homogeneous class $\bar{Q}^i(x) = 0$ whenever $i < |x|$. Hence $\bar{Q}^r \bar{Q}^s(x) = 0$ for $|x| \geq 1$ and $r \leq s$.

Since the extended Dyer–Lashof operations are defined on the mod 2 homology of all nonconnective spectra, the operations $\bar{Q}^i$ for all $i \in \mathbb{Z}$ can be defined on the mod 2 homology of any spectral Lie algebra

L with $\bar{Q}^i(x) = 0$ for any homogeneous class $x \in H_*(L)$ and $i < |x|$. Let $\bar{\mathcal{R}}$ be the quotient algebra of the free algebra over $\mathbb{F}_2$ on generators $\{\bar{Q}^j\}_{j \in \mathbb{Z}}$ by the two sided ideal generated by the relations

$$(5) \quad \bar{Q}^r \bar{Q}^s = \sum_{l \leq r-s-1} \binom{r-2l-1}{s-l} \bar{Q}^{r+s-l} \bar{Q}^l$$

for all $r \leq 2s$.

The Lie^s-bracket interacts with the unary operations in the following way.

Proposition 2.1 [2, Lemma 6.4, 6.5] *For any $j \in \mathbb{Z}$ and x, y homogeneous classes in the mod 2 homology of a spectral Lie algebra, we have $[\bar{Q}^j(x), y] = 0$ and $\bar{Q}^{|x|}(x) = [x, x]$.*

Remark 2.2 It follows that $\bar{Q}^{2|x|-1} \bar{Q}^{|x|}(x) = [[x, x], [x, x]] = 0$. This is guaranteed by the Behrens' relations, since $r = 2|x| - 1 \leq s = 2|x|$ and the right hand side of (5) vanished due to instability of the extended Dyer–Lashof operations.

Sometimes it is more convenient to switch to the lower indexing $\bar{Q}_j(x) := \bar{Q}^{|x|+j}(x)$, which automatically takes into account the instability condition.

Definition 2.3 The lower indexed $\bar{\mathcal{R}}$ -algebra is generated by symbols $\bar{Q}_j$ for $j \geq 0$ and relations

$$(6) \quad \bar{Q}_a \bar{Q}_b = \sum_{0 \leq c < (a+2b-1)/3} \binom{a+b-2c-2}{b-c} \bar{Q}_{a+2b-2c} \bar{Q}_c$$

for $0 \leq a \leq b + 1$. When $j < 0$ we set $\bar{Q}_j = 0$.

Definition 2.4 An $\mathbb{F}_2$ -module $M_\bullet$ over $\bar{\mathcal{R}}$ is *allowable* if for any homogeneous element $x \in M_\bullet$ we have $\bar{Q}^{j_1} \bar{Q}^{j_2} \dots \bar{Q}^{j_m}(x) = 0$ whenever $j_1 < j_2 + \dots + j_m + |x|$. Alternatively, an allowable $\bar{\mathcal{R}}$ -module M is a module over the lower-indexed $\bar{\mathcal{R}}$ -algebra.

Now we extend Behrens' results to all spectral Lie algebras.

Proposition 2.5 *For L any spectral Lie algebra, its mod 2 homology $H_*(L)$ is an allowable module over $\bar{\mathcal{R}}$. Furthermore, for all $k \geq 0$ and $n \in \mathbb{Z}$ there is an isomorphism of $\mathbb{F}_2$ -modules,*

$$\begin{aligned} H_*(\partial_{2^k}(\text{Id}) \otimes_{h\Sigma_{2^k}} (\mathbb{S}^n)^{\otimes 2^k}) &\cong \mathbb{F}_2\{\bar{Q}^{j_1} \dots \bar{Q}^{j_k}(x_n), j_l > 2j_{l+1} \forall l < k, j_k \geq n\} \\ &\cong \mathbb{F}_2\{\bar{Q}_{i_1} \dots \bar{Q}_{i_k}(x_n), \forall l, i_l \geq 0, i_l > i_{l+1} + 1\}. \end{aligned}$$

Proof The connectedness assumption in Behrens' proof of [9, Theorem 1.5.1] is necessary only because of the connectedness assumption on the following two inputs to the proof. Kuhn [44, Example 7.6] (see

also [9, Lemma 1.4.3]) showed that for Y a connected space, the transfer $\tau: H_*(Y_{h\Sigma_4}^{\otimes 4}) \rightarrow H_*(Y_{h\Sigma_2 \wr \Sigma_2}^{\otimes 4})$ is given by

$$(7) \quad Q^r Q^s \mapsto Q^r \wr Q^s + \sum_t \left[\binom{s-r+t}{s-t} + \binom{s-r+t}{2t-r} \right] Q^{r+s-t} \wr Q^t.$$

On the other hand, Arone and Mahowald's computation [5, Theorem 3.16]

$$H_*(\partial_{2^k}(\text{Id}) \otimes_{h\Sigma_{2^k}} (\mathbb{S}^n)^{\otimes 2^k}) \cong \mathbb{F}_2\{\bar{Q}^{j_1} \dots \bar{Q}^{j_k}(x_n), j_l > 2j_{l+1} \forall l < k, j_k \geq n\}$$

works for any odd integer n , and extends to positive even integers via the fiber sequence

$$\partial_{2m}(\text{Id}) \otimes_{h\Sigma_{2m}} (\mathbb{S}^n)^{\otimes 2m} \xrightarrow{E} \Sigma^{-1} \partial_{2m}(\text{Id}) \otimes_{h\Sigma_{2m}} (\mathbb{S}^{n+1})^{\otimes 2m} \xrightarrow{H} \Sigma^{-1} \partial_m(\text{Id}) \otimes_{h\Sigma_m} (\mathbb{S}^{2n+1})^{\otimes m},$$

which was obtained by differentiating the EHP sequence [5, Proposition 4.7; 9, Corollary 2.1.4]. Behrens proved the relations by using the transfer formula and inductively checking that they are compatible with the operadic composition; then he provided a basis by comparing with the Arone–Mahowald answer. Hence we only need to remove the connectedness assumption on both inputs.

Note that Kuhn's transfer formula can be obtained as a consequence of the computation of the transfer map $\tau_0: H_*(B\Sigma_4) \rightarrow H_*(B\Sigma_2 \wr \Sigma_2)$ on group homology by Priddy [52, Section 4]. For any $j, n \in \mathbb{Z}$, the Dyer–Lashof operation Q^j on a class x in degree n is defined via the canonical isomorphism $H_{n+j}((\Sigma^n \mathbb{F}_2)_{h\Sigma_2}^{\otimes 2}) \cong H_{j-n}(B\Sigma_2)[2n]$ [46], where $[k]$ denotes a shift in homological degree by k . Similarly, the wreath product $Q^r \wr Q^s$ and the weight 4 operation $Q^r Q^s$ are defined in $H_{j-n}(B\Sigma_2 \wr \Sigma_2)[4n]$ and $H_{j-n}(B\Sigma_4)[4n]$ respectively, so the transfer map τ on a class in degree n of any spectrum Y is a shift of τ_0 by $4n$. Hence the formula (7) holds for the transfer map τ on any spectrum Y .

Next we extend the Arone–Mahowald computation to nonpositive spheres. We make use of the long exact sequence

$$\begin{aligned} \dots \rightarrow H_*(\Sigma^{-2} \partial_m(\text{Id}) \otimes_{h\Sigma_m} (\mathbb{S}^{2n+1})^{\otimes m}) &\xrightarrow{P_*} H_*(\partial_{2m}(\text{Id}) \otimes_{h\Sigma_{2m}} (\mathbb{S}^n)^{\otimes 2m}) \\ &\xrightarrow{E_*} H_*(\Sigma^{-1} \partial_{2m}(\text{Id}) \otimes_{h\Sigma_{2m}} (\mathbb{S}^{n+1})^{\otimes 2m}) \xrightarrow{H_*} \dots \end{aligned}$$

and isomorphisms

$$H_*(\partial_{2m-1}(\text{Id}) \otimes_{h\Sigma_{2m-1}} (\mathbb{S}^{2n})^{\otimes (2m-1)}) \cong H_*(\Sigma^{-1}(\partial_{2m-1}(\text{Id}) \otimes_{h\Sigma_{2m-1}} (\mathbb{S}^{2n+1})^{\otimes (2m-1)}))$$

for all n obtained by Brantner in [16, 4.1.3]; see [37, Lemma 4.4]. There is an equivalence of $\mathbb{F}_2$ -module spectra with Σ_m -action

$$\partial_m(\text{Id}) \otimes_{h\Sigma_m} (\Sigma^n \mathbb{F}_2)^{\otimes m} \simeq \Sigma^{2mn} \partial_m(\text{Id}) \otimes_{h\Sigma_m} (\Sigma^{-n} \mathbb{F}_2)^{\otimes m}$$

for any integers $m, n \geq 0$, where the action on Σ^{2mn} is trivial. Hence we obtain an isomorphism

$$H_*(\partial_m(\text{Id}) \otimes_{h\Sigma_m} (\mathbb{S}^n)^{\otimes m}) \cong H_*(\Sigma^{2mn} \partial_m(\text{Id}) \otimes_{h\Sigma_m} (\mathbb{S}^{-n})^{\otimes m})$$

sending $\bar{Q}_{j_1} \dots \bar{Q}_{j_k}(\iota_n)$ to $\sigma^{2^{k+1}n} \bar{Q}_{j_1} \dots \bar{Q}_{j_k}(\iota_{-n})$ when $m = 2^k$, and both vanish when $m \neq 2^k$ for some $k \geq 0$. This addresses the case of the negative spheres.

For $n = 0$, we use the long exact sequence above. It follows from the case $n = 1$ that

$$H_*(\partial_m(\text{Id}) \otimes_{h\Sigma_m} (\mathbb{S}^0)^{\otimes m}) = 0$$

when m is not a power of 2. Now suppose that $m = 2^k$. By the [9, Proposition 2.2.5] (see the remark after [37, Proposition 4.3]), the maps E_* and P_* preserve the $\bar{Q}$ operations, sending the class $\bar{Q}^J(x_n)$ to $\sigma^{-1}\bar{Q}^J(x_{n+1})$ and $\sigma^{-2}\bar{Q}^J x_{2n+1}$ to $\bar{Q}^J \bar{Q}^n(x_n)$, respectively. This addresses the case $n = 0$. $\square$

Denote by $\text{Mod}_{\bar{\mathcal{R}}}$ the category of allowable $\bar{\mathcal{R}}$ -modules and $\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}$ the free allowable $\bar{\mathcal{R}}$ -module functor, which is left adjoint to the underlying functor $U_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}} : \text{Mod}_{\bar{\mathcal{R}}} \rightarrow \text{Mod}_{\mathbb{F}_2}$. We will suppress the adjective allowable from here on. Then there is an additive monad associated with the free $\bar{\mathcal{R}}$ -module functor, which we denote by $\mathcal{A}_{\bar{\mathcal{R}}}$.

Definition 2.6 [2, Definition 6.1] An $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra is a graded $\mathbb{F}_2$ -module $L_\bullet$ with a shifted Lie bracket and an (allowable) $\bar{\mathcal{R}}$ -module structure on $L_\bullet$ such that

- (1) $\bar{Q}_0(x) = \bar{Q}^k(x) = [x, x]$ if $x \in L_k$, and
- (2) $[x, \bar{Q}^k(y)] = 0$ for all $x, y \in L_\bullet$.

Denote by $\text{Lie}_{\bar{\mathcal{R}}}^s$ the category of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras. To describe the free $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra functor, we recall the construction of *Lyndon words* on a set S , which provides a basis for the free $\text{Lie}^{s, \text{ti}}$ -algebra on an $\mathbb{F}_2$ -module with $\mathbb{F}_2$ -basis S .

Construction 2.7 [33] The Lyndon words on a set S is defined recursively as follows: The elements of S are Lyndon words of length one and given an arbitrary fixed total ordering. Suppose that we have defined Lyndon words of length less than k with a total ordering. Then a Lyndon word of length k is a formal bracket $\langle w_1, w_2 \rangle$ such that:

- (1) w_1, w_2 are Lyndon words whose lengths add up to k .
- (2) $w_1 < w_2$ in the order defined thus far.
- (3) To take into account the Jacobi identity, if $w_2 = \langle w_3, w_4 \rangle$ for some Lyndon words w_3, w_4 , then we require $w_3 \leq w_1$.

To extend the total order to Lyndon words of weight at most k , we first impose an arbitrary total ordering on Lyndon words of length k , and then declare that they are greater than all Lyndon words of lower weights.

The free $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra functor can be computed explicitly as follows:

Proposition 2.8 [2, Proposition 7.4] Let $V_\bullet$ be an $\mathbb{F}_2$ -module with an ordered basis B of $V_\bullet$. First take the free totally isotropic Lie-algebra with $\langle -, - \rangle$ the free $\text{Lie}^{s, \text{ti}}$ bracket. Denote by B' the set of Lyndon words on the letters B , which is an $\mathbb{F}_2$ -basis of $\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^{s, \text{ti}}}(V_\bullet)$. Then we take the free $\bar{\mathcal{R}}$ -module on the underlying $\mathbb{F}_2$ -module of $\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^{s, \text{ti}}}(V_\bullet)$ and obtain a basis consisting of elements of the form

$\bar{Q}^I w$ with $w \in B'$. Equip the free $\bar{\mathcal{R}}$ -module $\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\text{Lie}^{s, \text{ti}}(V_\bullet))$ with a Lie^s bracket $[-, -]$ defined on the induced basis by requiring $[\bar{Q}^I w_1, \bar{Q}^J w_2] = 0$ if $I \neq \emptyset$ or $J \neq \emptyset$, and setting recursively along the ordering on B' .

- (1) If $\langle w_1, w_2 \rangle$ is a Lyndon word, then $[w_1, w_2] = \langle w_1, w_2 \rangle$.
- (2) $[w, w] := \bar{Q}^{|w|} w$.
- (3) $[w_1, w_2] := [w_2, w_1]$ if $w_1 > w_2$.
- (4) $[w_1, w_2] := [w_3, [w_1, w_4]] + [w_4, [w_1, w_3]]$ if $w_1 < w_2$ and $w_2 = [w_3, w_4]$ with $w_1 < w_3$.

Antolín Camarena showed that the monad $\text{Lie}_{\bar{\mathcal{R}}}^s$ parametrizes natural operations on the mod 2 homology of connected spectral Lie algebras. The connectivity assumption can be removed in view of [Proposition 2.5](#). Denote by $\text{Free}^{s, \mathcal{L}}$ the free spectral Lie algebra functor on Spectra given explicitly by

$$X \mapsto \bigoplus_{n \geq 1} \partial_n(\text{Id}) \otimes_{h\Sigma_n} X^{\otimes n}.$$

Theorem 2.9 [\[2, Theorem 7.1\]](#) *The canonical map*

$$\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(H_*(X; \mathbb{F}_2)) \rightarrow H_*(\text{Free}^{s, \mathcal{L}}(X); \mathbb{F}_2)$$

of Lie^s _{$\bar{\mathcal{R}}$} -algebras is an isomorphism for any spectrum X .

Proof Behrens proved the theorem in the case when $X = \mathbb{S}^k$, $k > 0$. Antolín Camarena proved the isomorphism for X a connected spectrum follows: To extend Behrens' theorem to a finite wedge of spheres, he made use of a result of Arone and Kankaarinta that applies Goodwillie calculus to the Hilton–Milnor theorem [\[4, Theorem 0.1\]](#). To extend to all connected spectra, note that $X \otimes \mathbb{F}_2$ can be written as a filtered colimit of finite wedges of $\mathbb{S}^m \otimes \mathbb{F}_2$ in the category of $\mathbb{F}_2$ -module spectra. The same arguments work to extend the isomorphism in [Proposition 2.5](#) to all spectra. $\square$

The category $\text{Mod}_{\bar{\mathcal{R}}}$ is stable under the desuspension functor $\Omega := \Sigma^{-1}$ of $\mathbb{F}_2$ -modules since the extended Dyer–Lashof operations are. Namely, for $M \in \text{Mod}_{\bar{\mathcal{R}}}$, the $\mathbb{F}_2$ -module ΩM has an $\bar{\mathcal{R}}$ -module structure given by $\bar{Q}^j(\sigma^{-1}x) = \sigma^{-1}\bar{Q}^j(x)$ for any $x \in M$. As a result, for any Lie^s _{$\bar{\mathcal{R}}$} -algebra, there is an Lie^s _{$\bar{\mathcal{R}}$} -structure on $\Omega \mathfrak{g}$ such that the bracket is trivial.

Proposition 2.10 *There is a natural Lie^s _{$\bar{\mathcal{R}}$} -module structure on $\Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2)$ for $1 \leq n \leq \infty$, where the bracket is trivial and $\bar{Q}^j$ acts by $x \mapsto \sigma^{-n}\bar{Q}^j(\sigma^n x)$. The canonical map*

$$\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^k \mathbb{F}_2) \cong \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\Sigma^k \mathbb{F}_2) \rightarrow \Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2)$$

is surjective.

Proof There is a canonical colimit-to-limit comparison map

$$(8) \quad \text{Free}^{s, \mathcal{L}}(\Sigma^k \mathbb{F}_2) \rightarrow \Omega \text{Free}^{s, \mathcal{L}}(\Sigma^{k+1} \mathbb{F}_2)$$

of spectral Lie algebras over $\mathbb{F}_2$, which after taking homotopy groups is the composite of the top and right arrows of the diagram

$$\begin{array}{ccc} \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^k \mathbb{F}_2) & \longrightarrow & \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^s_{\bar{\mathcal{R}}}}(\Omega \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{k+1} \mathbb{F}_2)) \\ \downarrow \text{ev} & & \downarrow \text{ev} \\ \Sigma^k \mathbb{F}_2 & \xrightarrow{i} & \Omega \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{k+1} \mathbb{F}_2) \end{array}$$

Let x be the generator of $\Sigma^k \mathbb{F}_2$. By naturality of the $\bar{Q}^j$ operation, the class $\bar{Q}^j(x)$ on the top left corner is mapped to $\bar{Q}^j(i(x))$, which is sent to $\sigma^{-1} \bar{Q}^j \sigma(x)$ under evaluation. In general $\bar{Q}^J(x)$ is mapped to $\sigma^{-1} \bar{Q}^J \sigma(x)$ for any sequence J . Since the Lie^s bracket of operations always vanishes and

$$[i(x), i(x)] = \bar{Q}^{|x|}(i(x)) = \sigma^{-1} Q^{|x|}(\sigma x) = 0,$$

the Lie^s -bracket is trivial on $\Omega \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{k+1} \mathbb{F}_2)$. Applying [Theorem 2.9](#), we see that the composite is surjective since $|\sigma(x)| = |x| + 1$. Iterating the construction yields the claim. $\square$

2.3 Quillen homology of spectral Lie algebras

Now we introduce the main object of interest. The inclusion of trivial $\text{Lie}^s_{\bar{\mathcal{R}}}$ -algebras admits a left adjoint $Q_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^s_{\bar{\mathcal{R}}}}$ called the indecomposable functor, ie we have an adjunction

$$\text{Lie}^s_{\bar{\mathcal{R}}} \begin{array}{c} \xrightarrow{Q_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^s_{\bar{\mathcal{R}}}}} \\ \xleftarrow{T_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^s_{\bar{\mathcal{R}}}}} \end{array} \text{Mod}_{\mathbb{F}_2}.$$

Denote again by $\text{Lie}^s_{\bar{\mathcal{R}}}$ the monad associated to the free $\text{Lie}^s_{\bar{\mathcal{R}}}$ -algebra functor.

We would like to understand the left derived functor of this left adjoint, and we take a small detour to deal with the model structure. We mainly follow Sections 3.1 and 3.2 of [\[36\]](#) and Section 4 of [\[17\]](#).

2.3.1 The derived indecomposable functor Let T be an augmented monad on the category Mod_k of weighted graded k -modules, where k is a field. Denote by $\text{Alg}_T(\text{Mod}_{\mathbb{F}_2})$ the category of T -algebras. The forgetful functor $U : \text{Alg}_T(\text{Mod}_k) \rightarrow \text{Mod}_k$ admits a left adjoint, the free functor $\text{Free}^T : \text{Mod}_k \rightarrow \text{Alg}_T(\text{Mod}_k)$.

Denote by $s \text{Mod}_k$ the category of simplicial weighted graded k -modules. Levelwise application of the adjunction $\text{Free}^T \dashv U$ gives rise to an adjunction between the corresponding categories of simplicial objects

$$\text{Free}^T \dashv U : \text{Alg}_T(s \text{Mod}_k) \rightarrow s \text{Mod}_k,$$

as well as a monad T on $s \text{Mod}_k$. We equip $s \text{Mod}_k$ with the standard cofibrantly generated model structure. Suppose that the path objects of $s \text{Mod}_k$ lifts to $s \text{Alg}_T$, the category of simplicial T -algebras. Then this adjunction induces a right transferred model structure on the category of simplicial T -algebras,

with weak equivalences and fibrations defined on the underlying simplicial weighted graded k -modules by [36, Theorem 3.2, Remark 3.3]. In particular, this is true for $T = \bar{\mathcal{R}}, \text{Lie}^s, \text{Lie}^{s, \text{ti}}, \text{Lie}_{\bar{\mathcal{R}}}^s$, see [17, Proposition 3.4, 4.14].

Denote by $T^T : \text{Mod}_k = \text{Alg}_{\text{Id}}(\text{Mod}_k) \rightarrow \text{Alg}_T(\text{Mod}_k)$ the inclusion of trivial T -algebras, which is induced by the augmentation. It has a left adjoint $Q^T : \text{Alg}_T(\text{Mod}_k) \rightarrow \text{Mod}_k$, the indecomposable functor with respect to the T -algebra structure, which satisfies $Q^T \circ \text{Free}^T \simeq \text{id}$. Applying this adjunction levelwise to the corresponding categories of simplicial objects, we obtain a Quillen adjunction

$$Q^T \dashv T^T : s\text{Alg}_T \rightarrow s\text{Mod}_k.$$

The left derived functor $\mathbb{L}Q^T$ of Q^T can be computed by the following standard recipe.

Construction 2.11 Given a right module $R : \text{Mod}_k \rightarrow \mathcal{D}$ over T , and a simplicial object A in $\text{Alg}_T(\text{Mod}_k)$, one can apply the two-sided bar construction $\text{Bar}_\bullet(R, T, -)$ levelwise to A . The diagonal of the resulting bisimplicial complex is a simplicial object in $\mathcal{D}$, denoted by $\text{Bar}_\bullet(R, T, A)$.

In particular, if we regard a T -algebra A as the constant simplicial object on $U(A)$ equipped with a simplicial T -algebra structure, denoted also as A by abuse of notation, then $\text{Bar}_\bullet(R, T, A)$ agrees with the usual two-sided bar construction.

Since the free resolution $\text{Bar}_\bullet(\text{Free}^T, T, A)$ is a cofibrant replacement of A in the category of simplicial T -algebras, the left derived functor of a functor F can be computed by applying F levelwise to a cofibrant replacement, so

$$\mathbb{L}Q^T(A) \simeq Q^T \text{Bar}_\bullet(\text{Free}^T, T, A) = \text{Bar}_\bullet(\text{id}, T, A).$$

Now suppose that we have a composite monad $R \circ L$ in Mod_k with distributive law the natural transformation $L \circ R \Rightarrow R \circ L$ in the sense of Beck [8, Section 1]. Suppose in addition that L, R and $R \circ L$ are all compatibly augmented and each admit a cofibrant replacement given by the free resolution. Let $\text{Alg}_L, \text{Alg}_R, \text{Alg}_{R \circ L}$ be the respective categories of algebras. Then an $R \circ L$ -algebra A is an R -algebra via the forgetful map $U_{R \circ L}^R : \text{Alg}_{R \circ L} \rightarrow \text{Alg}_R$ induced by the augmentation of L , and an L -algebra via the augmentation of R . Furthermore, we have adjunctions

$$\text{Mod}_k \xrightleftharpoons[T^L]{Q^L} \text{Alg}_L \xrightleftharpoons[T_L^{R \circ L}]{Q_L^{R \circ L}} \text{Alg}_{R \circ L}.$$

Construction 2.12 For A an algebra over $R \circ L$, the free resolution $\text{Bar}_\bullet(\text{Free}^R, R, A)$ has the structure of a simplicial $R \circ L$ -algebra given as follows. Levelwise, the $R \circ L$ -algebra structure map is given by

$$(R \circ L) \circ R^{\text{on}}(A) \rightarrow R \circ (R \circ L) \circ R^{\circ(n-1)}(A) \rightarrow \cdots \rightarrow R^{\text{on}} \circ (R \circ L)(A) \rightarrow R^{\text{on}}(A),$$

where the rightmost arrow is the $R \circ L$ -algebra structure map on A and the other arrows are induced from the distributive law $L \circ R \Rightarrow R \circ L$. The face and degeneracy maps are structure maps of the monad R and hence compatible with the levelwise $R \circ L$ -algebra structure maps by naturality of the distributive law.

Levelwise application of $Q_L^{R \circ L}$ to $\text{Bar}_\bullet(\text{Free}^R, R, A)$ yields a simplicial L -algebra structure on the bar construction $\text{Bar}_\bullet(\text{id}, R, A) = Q_L^{R \circ L} \text{Bar}_\bullet(\text{Free}^R, R, A)$.

We record the following result about factoring the left derived functor of the indecomposable functor of a composite monad, which generalizes [17, Proposition 4.19].

Lemma 2.13 *Let A be an $R \circ L$ -algebra. The homotopy groups of $\text{Bar}_\bullet(\text{id}, R \circ L, A)$ are computed by the homotopy groups of the bisimplicial object $\text{Bar}_\bullet(\text{id}, L, \text{Bar}_\bullet(\text{id}, R, A))$.*

Recall that the homotopy groups of a bisimplicial k -module can be computed via the Eilenberg–Zilber theorem, ie by first taking associated chain complexes in both directions and then forming the total complex of the double complex. See for instance [32, Chapter 4].

Proof The augmentation $R \circ L \rightarrow R$ induces a map of simplicial $R \circ L$ -algebras

$$\Psi: \text{Bar}_\bullet(\text{Free}^{R \circ L}, R \circ L, A) \rightarrow \text{Bar}_\bullet(\text{Free}^R, R, A),$$

where the simplicial $R \circ L$ -algebra structure on the target is given by Construction 2.12. This is an equivalence since both are free resolutions of A as an $R \circ L$ -algebra and an R -algebra, respectively, and weak equivalences in $s\text{Alg}_{R \circ L}$ are detected by the underlying simplicial k -modules. We want to show that $Q_L^{R \circ L}$ preserves this weak equivalence. Since U^L preserves weak equivalences, it suffices to show that $U^L \circ Q_L^{R \circ L} \circ \Psi$ is a weak equivalence.

Note that there is an isomorphism

$$Q^R \circ U_R^{R \circ L} \cong U^L \circ Q_L^{R \circ L}.$$

Hence $U^L \circ Q_L^{R \circ L} \circ \Psi$ is the map

$$\begin{aligned} Q^R \circ U_R^{R \circ L} \text{Bar}_\bullet(\text{Free}^{R \circ L}, R \circ L, A) &\rightarrow Q^R \circ U_R^{R \circ L} \text{Bar}_\bullet(\text{Free}^R, R, A) \\ &\simeq Q^R \text{Bar}_\bullet(\text{Free}^R, R, A) \simeq \text{Bar}_\bullet(\text{id}, R, A). \end{aligned}$$

Since both $U_R^{R \circ L} \text{Bar}_\bullet(\text{Free}^{R \circ L}, R \circ L, A)$ and $\text{Bar}_\bullet(\text{Free}^R, R, A)$ are free resolutions of A in $s\text{Alg}_R$ and Q^R is a left Quillen functor, this is indeed a weak equivalence. Hence

$$\begin{aligned} \text{Bar}_\bullet(\text{id}, L, \text{Bar}_\bullet(\text{id}, R, A)) &\simeq Q^L \circ Q_L^{R \circ L} \text{Bar}_\bullet(\text{Free}^{R \circ L}, R \circ L, A) \\ &\simeq Q^L \circ Q_L^{R \circ L} \text{Bar}_\bullet(\text{Free}^R, R, A) \\ &\simeq Q^L \text{Bar}_\bullet(\text{id}, R, A), \end{aligned}$$

and the lemma follows from Construction 2.11. $\square$

2.3.2 Quillen homology of $s\text{Lie}_R^s$ -algebras Since the path object of $s\text{Mod}_{\mathbb{F}_2}$ lifts to $s\text{Lie}_R^s$, the discussion in the previous subsection guarantees that any $\text{Lie}_{\overline{R}}^s$ -algebra $\mathfrak{g}$ admits a free resolution $\text{Bar}_\bullet(\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\overline{R}}^s}, \text{Lie}_{\overline{R}}^s, \mathfrak{g})$ in $\text{Lie}_{\overline{R}}^s$. Thus the left derived functor of $Q_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\overline{R}}^s}$ is computed by

$$\mathbb{L}Q_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\overline{R}}^s}(\mathfrak{g}) \simeq Q_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\overline{R}}^s} \text{Bar}_\bullet(\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\overline{R}}^s}, \text{Lie}_{\overline{R}}^s, \mathfrak{g}) \simeq \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{R}}^s, \mathfrak{g}),$$

where $\text{id}: \text{Mod}_{\mathbb{F}_2} \rightarrow \text{Mod}_{\mathbb{F}_2}$ is the identity functor considered as the trivial right module over the monad $\text{Lie}_{\bar{\mathcal{R}}}^s$ with structure map the augmentation.

Definition 2.14 The *Quillen homology* of a $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\mathfrak{g}$, denoted by $\text{HQ}_{*}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\mathfrak{g})$, is the total left derived functor

$$\text{HQ}_{*,*}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\mathfrak{g}) := H_{*,*} \mathbb{L} Q_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\mathfrak{g}) \simeq \pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g}).$$

We are interested in computing the Quillen homology of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras, since it helps to understand the spectral Lie analog of the mod p topological André–Quillen homology of nonunital $\mathbb{E}_{\infty}$ -algebras introduced by Kříž [41] and Basterra [7].

Definition 2.15 For L a spectral Lie algebra, its *topological Quillen object* is the bar construction

$$\text{TQ}^{s, \mathcal{L}}(L) := |\text{Bar}_{\bullet}(\text{id}, s\mathcal{L}, L)|.$$

We define its mod p *topological Quillen homology* to be

$$\text{TQ}_{*}^{s, \mathcal{L}}(L; \mathbb{F}_p) := \pi_{*}(|\text{Bar}_{\bullet}(\text{id}, s\mathcal{L}, L)| \otimes \mathbb{F}_p).$$

Using the skeletal filtration of the geometric realization of the bar construction, we obtain a bar spectral sequence

$$E_{s,t}^2 = \pi_s \pi_t \text{Bar}_{\bullet}(\text{id}, s\mathcal{L}, L \otimes \mathbb{F}_p) \Rightarrow \text{TQ}_{s+t}^{s, \mathcal{L}}(L; \mathbb{F}_p)$$

converging to the mod p topological Quillen homology. When $p = 2$, we can apply [Theorem 2.9](#) repeatedly and deduce that:

Proposition 2.16 *There is a bar spectral sequence*

$$E_{s,t}^2 = \pi_{s,t} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \pi_{*}(L \otimes \mathbb{F}_2)) \cong \text{HQ}_{s,t}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(H_{*}(L; \mathbb{F}_2)) \Rightarrow \text{TQ}_{s+t}^{s, \mathcal{L}}(L; \mathbb{F}_2).$$

3 Computing the Quillen homology of spectral Lie algebras

In this section, we study the Quillen homology of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras when $p = 2$ via comparison with two smaller double complexes that are easy to compute via Koszul duality arguments.

3.1 May-type spectral sequence and an upper bound

First we find an upper bound for $\pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$ by constructing a May-type spectral sequence. The dimensions of its E^1 -page is bounded above by the homotopy groups of the bar construction of the following variant of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras whose unary and binary operations do not intertwine.

Definition 3.1 Define a $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebra to be an $\mathbb{F}_2$ -module L with an allowable $\bar{\mathcal{R}}$ -module structure and a $\text{Lie}^{s, \text{ti}}$ -bracket $\langle -, - \rangle$ such that $\langle x, \bar{Q}^i(y) \rangle = 0$ for all $x, y \in L$. Denote by $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ the category of $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebras and also the monad associated to the free $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebra functor.

The underlying $\mathbb{F}_2$ -module of the free $\text{Lie}^{s, \text{ti}}_{\bar{\mathcal{R}}}$ -algebra on an $\mathbb{F}_2$ -module V is given by that of $\mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^{s, \text{ti}}(V)$. Hence $\text{Lie}^{s, \text{ti}}_{\bar{\mathcal{R}}}$ admits an alternative description as the category of algebras over the composite monad $\mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^{s, \text{ti}}$, with distributive law the natural transformation $\text{Lie}^{s, \text{ti}} \circ \mathcal{A}_{\bar{\mathcal{R}}} \Rightarrow \mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^{s, \text{ti}}$ determined by $\langle -, \bar{Q}^i(-) \rangle = 0$ for all i .

Remark 3.2 Comparing with Proposition 2.8, we see that the underlying $\bar{\mathcal{R}}$ -modules of the free $\text{Lie}^s_{\bar{\mathcal{R}}}$ and $\text{Lie}^{s, \text{ti}}_{\bar{\mathcal{R}}}$ -algebra on any $\mathbb{F}_2$ -module agree. The only difference between the two free functors is that in the latter we do not change the $\text{Lie}^{s, \text{ti}}$ -algebra to a Lie^s -algebra via the identification $\bar{Q}_0(x) = [x, x]$.

In particular, the bar construction $\text{Bar}_\bullet(\text{id}, \text{Lie}^s_{\bar{\mathcal{R}}}, \mathfrak{g})$ is levelwise isomorphic to $\text{Bar}_\bullet(\text{id}, \text{Lie}^{s, \text{ti}}_{\bar{\mathcal{R}}}, \mathfrak{g})$. The latter has simpler face maps in the sense that the face maps preserve the unary and binary structures respectively, whereas in the former, a Lie bracket that is not a self-bracket can be mapped to a self-bracket. To deal with these face maps, we draw inspiration from the May spectral sequence: suppose that we want to compute the Ext groups over a Hopf algebroid (A, Γ) . In good cases, there exists a filtration on Γ such that the associated graded is a Hopf algebra (A, Γ') , ie the left and right unit are equal. Then we obtain a May spectral sequence with E^1 -page the Ext group over the Hopf algebra Γ' , whose cochain complex has differentials simpler than the cobar complex for Γ . The higher differentials are determined by the difference between the left and the right unit.

To construct a filtration on $\text{Bar}_\bullet(\text{id}, \text{Lie}^s_{\bar{\mathcal{R}}}, \mathfrak{g})$ so that the associated graded assembles to $\text{Bar}_\bullet(\text{id}, \text{Lie}^{s, \text{ti}}_{\bar{\mathcal{R}}}, \mathfrak{g})$, first we need to construct a filtration on any $\text{Lie}^s_{\bar{\mathcal{R}}}$ -algebra so that the two sides of the identification $\bar{Q}_0(x) = [x, x]$ live in different filtrations.

Construction 3.3 (length filtration) Consider the complete filtration

$$\cdots \rightarrow \bar{\mathcal{R}}(n) \rightarrow \bar{\mathcal{R}}(n-1) \rightarrow \cdots \rightarrow \bar{\mathcal{R}}(1) \rightarrow \bar{\mathcal{R}}$$

of the homogeneous algebra $\bar{\mathcal{R}}$, where $\bar{\mathcal{R}}(n)$ is the ideal generated by monomials $\bar{Q}^I$ with $|I| = n$. Thus we obtain functors $\mathcal{A}_{\bar{\mathcal{R}}(n)}$ on $\text{Mod}_{\mathbb{F}_2}$, sending M to the submodule of $\mathcal{A}_{\bar{\mathcal{R}}}(M)$ consisting of $\bar{Q}^I(x)$ for $x \in M$ and $|I| \geq n$. In other words, this assembles to a filtered monad $\hat{\mathcal{A}}_{\bar{\mathcal{R}}}$. The images of the induced evaluation maps $\mathcal{A}_{\bar{\mathcal{R}}(q)}(M) \xrightarrow{\text{ev}_q} M$ form a complete decreasing filtration for any $\bar{\mathcal{R}}$ -module M . Taking cokernels yields a complete increasing filtration

$$F_l^q(M) = \text{coker}(\mathcal{A}_{\bar{\mathcal{R}}(q)}(M) \xrightarrow{\text{ev}_q} M).$$

We call this the *length filtration* of M , which gives rise to a filtered object $\hat{M}$ as an algebra over the filtered monad $\hat{\mathcal{A}}_{\bar{\mathcal{R}}}$ whose underlying $\bar{\mathcal{R}}$ -module is M .

Given an arbitrary $\text{Lie}^s_{\bar{\mathcal{R}}}$ -algebra $\mathfrak{g}$, we would like to equip $\mathfrak{g}$ with the structure of an $\text{Lie}^{s, \text{ti}}_{\bar{\mathcal{R}}}$ -algebra. This boils down to producing a method that equips any Lie^s -algebra with a $\text{Lie}^{s, \text{ti}}$ -algebra structure.

Construction 3.4 ($\text{Lie}^{s, \text{ti}}$ -structure on Lie^s -algebras) For $\mathfrak{g}$ is Lie^s -algebra with bracket $[-, -]$, let V' be the ideal of self-brackets. Thus we obtain a two-step filtration $V' \rightarrow \mathfrak{g}$ of $\mathfrak{g}$. Denote by $\langle -, - \rangle$ the canonical $\text{Lie}^{s, \text{ti}}$ -bracket on the quotient $V = \mathfrak{g}/V' = Q_{\text{Lie}^{s, \text{ti}}}^{\text{Lie}^s}(\mathfrak{g})$ and consider V' as a trivial $\text{Lie}^{s, \text{ti}}$ -algebra.

Thus we obtain a $\text{Lie}^{s, \text{ti}}$ -structure on the associated graded of $\mathfrak{g}$ as the product of V and V' with the above $\text{Lie}^{s, \text{ti}}$ -structures. Denote by $\tilde{\mathfrak{g}}$ the resulting $\text{Lie}^{s, \text{ti}}$ -algebra with $\langle -, - \rangle$ the $\text{Lie}^{s, \text{ti}}$ -bracket.

Therefore, any $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\mathfrak{g}$ admits a $\text{Lie}^{s, \text{ti}}$ -structure that is unique up to a choice of splitting of $\mathfrak{g} \rightarrow V$. Denote by $\tilde{\mathfrak{g}}$ the resulting $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebra, which has the same underlying $\bar{\mathcal{R}}$ -module structure as $\mathfrak{g}$; see [Remark 3.2](#).

Remark 3.5 If we fix a choice of splitting for $\mathfrak{g} \rightarrow V$, then any Lie^s bracket $[x, y]$ in $\mathfrak{g}$ is equal to a sum of self-brackets and the $\text{Lie}^{s, \text{ti}}$ -brackets $\langle x, y \rangle$ in $\tilde{\mathfrak{g}}$.

Now we compute the E^2 -page of the bar spectral sequence by constructing a May-type spectral sequence in the sense that the filtration comes from the length filtration of $\bar{\mathcal{R}}$ -modules in [Construction 3.3](#).

Theorem 3.6 Let $\mathfrak{g}$ be a $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra and $\tilde{\mathfrak{g}}$ the associated $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebra via [Construction 3.4](#). Then there is a May-type spectral sequence with respect to the $\bar{\mathcal{R}}$ -module structure converging to $\pi_{s, t} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$. The E^1 -page $\bigoplus_q E_{q, s, t}^1$ of the May-type spectral sequence has dimensions bounded above by

$$\pi_{s, t} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}, \tilde{\mathfrak{g}}),$$

in the sense that there is another algebraic spectral sequence converging to the May E^1 -page whose E^1 -page is $\pi_{s, t} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}, \tilde{\mathfrak{g}})$.

Proof We start by inductively constructing a filtration on $(\text{Lie}_{\bar{\mathcal{R}}}^s)^{\circ n}(\mathfrak{g})$ that heuristically count the number of $\bar{Q}$ symbols in a given element.

Since any $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\mathfrak{g}$ is an $\bar{\mathcal{R}}$ -module, it admits a length filtration. The filtration is compatible with the bracket since brackets of operations always vanish ([Definition 2.6](#)). Furthermore, since any self-bracket $[x, x] = \bar{Q}_0(x)$ is in $F_l^1(\mathfrak{g})$ and the right hand side is zero in $F_l^0(\mathfrak{g})$, we deduce that $\text{Gr}_l^0(\mathfrak{g})$ is a $\text{Lie}^{s, \text{ti}}$ -algebra, and the $\text{Lie}^{s, \text{ti}}$ -structure can be extended to $\bigoplus \text{Gr}^q(\mathfrak{g})$ via trivial extension to positive q . On the other hand, $\bigoplus \text{Gr}^q(\mathfrak{g})$ is an $\bar{\mathcal{R}}$ -module since $\bar{\mathcal{R}}$ is homogeneous. Hence $\tilde{\mathfrak{g}} = \bigoplus \text{Gr}_l^q(\mathfrak{g})$ equipped with the $\text{Lie}^{s, \text{ti}}$ -bracket in [Construction 3.4](#) is an algebra over the composite monad $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}} = \mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^{s, \text{ti}}$.

Now we define a new filtration $F^{\bullet}$ on $\text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g})$ that combines the length filtration on $\mathfrak{g}$, the length filtration on $\text{Lie}_{\bar{\mathcal{R}}}^s(M)$ for any $\mathbb{F}_2$ -module M , and the effect of Lie^s brackets. We extend the length filtration on $\mathfrak{g}$ to $\text{Lie}^{s, \text{ti}}(\mathfrak{g})$ via the Day convolution, ie for $x \in F_l^q(\mathfrak{g})$, $y \in F_l^r(\mathfrak{g})$, we have $\langle x, y \rangle \in F^{q+r}(\text{Lie}^{s, \text{ti}}(\mathfrak{g}))$, so on and so forth. Then we extend it to a new filtration on $\text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g})$ by combining with the length filtration on $\text{Lie}_{\bar{\mathcal{R}}}^s(M)$ for M an $\mathbb{F}_2$ -module, using the fact that when $\mathfrak{g} = M$ is an $\mathbb{F}_2$ -module, $\text{Gr}_l^0(\text{Lie}_{\bar{\mathcal{R}}}^s(M)) = \text{Lie}^{s, \text{ti}}(M)$. In particular, after passing to the associated graded, the evaluation map $\text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g}) \rightarrow \mathfrak{g}$ assembles to the $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebra structure map $\text{ev}: \mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^{s, \text{ti}}(\tilde{\mathfrak{g}}) \rightarrow \tilde{\mathfrak{g}}$. This is because for $x \in \mathfrak{g}$, $[x, x] = \bar{Q}_0|x \in \text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g})$ is mapped to a nonzero element only if $x \in F_l^0(\mathfrak{g})$, in which case $\bar{Q}_0|x \in F^1 \text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g})$ and $\bar{Q}_0(x) \in F_l^1(\mathfrak{g})$ while $[x, x] \in F^0 \text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g})$.

Iterating this process, we obtain a filtration $F^{\bullet}$ on $\text{Lie}_{\bar{\mathcal{R}}}^s \circ (\text{Lie}_{\bar{\mathcal{R}}}^s)^{\circ n}(\mathfrak{g})$ for all $n > 0$ by combining the filtration $F^{\bullet}$ on $(\text{Lie}_{\bar{\mathcal{R}}}^s)^{\circ n}(\mathfrak{g})$ with the length filtration on $\text{Lie}_{\bar{\mathcal{R}}}^s$. This is the n^{th} simplicial level of

$\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$, with associated graded assembling to $(\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}})^{\circ n}(\tilde{\mathfrak{g}})$. Explicitly, $F^q((\text{Lie}_{\bar{\mathcal{R}}}^s)^{\circ n}(\mathfrak{g}))$ is the collection of elements $\alpha|x$ in simplicial degree n satisfying the following condition: if we rewrite $\alpha|x$ as an element in $(\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}})^{\circ n}(\mathfrak{g})$ via Remarks 3.2 and 3.5, so any Lie^s bracket in $\alpha|x$ is written as a linear combination of $\text{Lie}^{s, \text{ti}}$ brackets and $\bar{Q}_0$ applies to other elements, then the sum of the filtration degree of $x \in \mathfrak{g}$ times the number of times x appears and the number of symbols $\bar{Q}^j$ in any term of $\alpha|x$ coming from applications of the monad $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ is at most q .

Since $\bar{\mathcal{R}}$ is a homogeneous algebra and evaluation of brackets do not increase the number of $\bar{Q}^j$ in the expression, the structure map $\text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g}) \rightarrow \mathfrak{g}$ is compatible with this filtration, and so are the face maps and degeneracy maps in $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$. The induced filtration $F^\bullet$ on the normalized complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$ gives rise to a May-type spectral sequence

$$\bigoplus_q E_{q,s,t}^1 = \bigoplus_q \pi_{s,t} \text{Gr}^q \text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g}) \Rightarrow \pi_{s,t} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g}).$$

Note that the face maps

$$(\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}})^{\circ n}(\tilde{\mathfrak{g}}) = \bigoplus_q \text{Gr}^q \text{Bar}_n(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g}) \rightarrow (\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}})^{\circ(n-1)}(\tilde{\mathfrak{g}}) = \bigoplus_q \text{Gr}^q \text{Bar}_{n-1}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$$

do not change the associated graded degree unless the differential creates self-brackets — evaluating the $\bar{\mathcal{R}}$ -module structure or a $\text{Lie}^{s, \text{ti}}$ -bracket is either zero or does not change the number of $\bar{Q}$ symbols. Hence they assemble to the $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebra structure maps $(\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}})^{\circ j}(\tilde{\mathfrak{g}}) \rightarrow (\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}})^{\circ(j-1)}(\tilde{\mathfrak{g}})$ except in the following situation: for $x \in \text{Gr}_l^0(\mathfrak{g})$, the classes $\gamma_1(\bar{Q}^i|x) := [\bar{Q}^i|1|x, 1|\bar{Q}^i|x]$ and $\bar{Q}_0|\bar{Q}^i|x$ are both in the second associated graded piece. Hence the total differential ∂ of the normalized complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$ sends

$$\gamma_1(\bar{Q}^i|x) := [\bar{Q}^i|1|x, 1|\bar{Q}^i|x]$$

to the element

$$[\bar{Q}^i|x, \bar{Q}^i|x] + [\bar{Q}^i|x, 1|\bar{Q}^i(x)] = \bar{Q}_0|\bar{Q}^i|x + [\bar{Q}^i|x, 1|\bar{Q}^i(x)]$$

in $E_{2,*,*}^0$, ie the self-bracket has not been filtered away. Similarly, any class containing $\gamma_1(\bar{Q}^i|x)$ with $x \in \text{Gr}_l^0(\mathfrak{g})$ has a face map whose target has at least one self-bracket term. Whereas when $x \in F_l^1(\mathfrak{g})$, the self-brackets in the target of such differentials are not visible in the associated graded because the number of $\bar{Q}^j$ in the term decrease after we rewrite the self-brackets in terms of $\bar{Q}_0$.

Hence we need to construct another spectral sequence to compute the E^1 -page of the May-type spectral sequence. To further filter away the self-brackets in such differentials, we assign weight 1 to $\gamma_1(\bar{Q}^i|x)$ and $[\bar{Q}^i|x, 1|\bar{Q}^i(x)]$ for all i and $x \in \text{Gr}_l^0(\mathfrak{g})$, weight 0 to everything else in $\mathfrak{g}$, $\text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g})$, and $\text{Lie}_{\bar{\mathcal{R}}}^s \circ \text{Lie}_{\bar{\mathcal{R}}}^s(\mathfrak{g})$, including $\bar{Q}_0|\bar{Q}^i|x$. Then we propagate the weight to $(\text{Lie}_{\bar{\mathcal{R}}}^s)^{\circ n}(\mathfrak{g})$ for $n > 2$ by stipulating that further applying operations $\bar{Q}$ does not change weight and brackets add weights. The associated graded of this weight filtration is precisely $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}, \tilde{\mathfrak{g}})$, since the only face or degeneracy maps that are altered are the ones involving $\gamma_1(\bar{Q}^i|x)$ for $x \in \text{Gr}_l^0(\mathfrak{g})$, whose target no longer contains the self-bracket term $\bar{Q}_0|\bar{Q}^i|x$. Therefore we obtain an algebraic spectral sequence that converges to the E^1 -page of the

May-type spectral sequence. Then its E^1 -page has dimensions precisely those of $\pi_{s,t} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}, \tilde{\mathfrak{g}})$. Therefore we obtain an upper bound of the dimension of the E^1 -page of the May-type spectral sequence $\bigoplus_q E_{q,s,t}^1$. $\square$

We will call the spectral sequence above computing the E^1 -page of the May-type spectral sequence the γ_1 -Bockstein spectral sequence. Since differentials preserve weights and the γ_1 operation on $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}, L)$ appears in weight at least four, we immediately deduce the following from [Theorem 3.6](#).

Corollary 3.7 *For any $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\mathfrak{g}$, the homotopy groups of $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$ and $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}, \tilde{\mathfrak{g}})$ are isomorphic in weight less than four.*

Proof In the γ_1 -Bockstein spectral sequence, the differentials do not appear until weight 4 since $\gamma_1(\bar{Q}_j | x)$ has weight 4. By construction, differentials in the May-type spectral sequence occur when the source and target of a face map in $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$ have different number of self-brackets. In weight two and three this cannot happen. Hence both spectral sequences collapse in weight less than four. $\square$

Remark 3.8 In the case where the $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\mathfrak{g}$ has vanishing Lie brackets, we can combine the two spectral sequences in [Theorem 3.6](#) into a single spectral sequence that converges to the E^2 -page of the bar spectral sequence by simply shifting the filtration of any element in $\mathfrak{g}$ up by 2. Since $\bar{Q}$ operations are additive and the Lie structure is trivial, the resulting filtered object $\mathfrak{g}[2]$ is a module over the filtered monad $\hat{\mathcal{A}}_{\bar{\mathcal{R}}}$, whose underlying object is still the $\bar{\mathcal{R}}$ -module $\mathfrak{g}$. Then in the bar construction $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$, the class $\bar{Q}_0 | \alpha$ always lives in a lower filtration than $[\alpha, \alpha]$ for any class α , since the filtration of α is at least 2. Hence the May-type spectral sequence has E^1 -page given precisely by $\pi_{s,t} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}, \tilde{\mathfrak{g}})$.

Since $\text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}} = \mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^{s,\text{ti}}$ is a composite monad, we apply [Construction 2.12](#) and [Lemma 2.13](#) to compute the homotopy groups of $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}, L)$ for L a $\text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}$ -algebra.

Construction 3.9 For L a $\text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}$ -algebra with $\text{Lie}^{s,\text{ti}}$ -bracket $\langle -, - \rangle$, denote by $\text{AR}_\bullet(L)$ the bar construction $\text{Bar}_\bullet(\text{id}, \mathcal{A}_{\bar{\mathcal{R}}}, L)$ equipped with a $\text{Lie}^{s,\text{ti}}$ -bracket $\langle -, - \rangle$ given levelwise by

$$\langle \alpha_1 | \alpha_2 | \dots | \alpha_n | x, \beta_1 | \beta_2 | \dots | \beta_n | y \rangle = \begin{cases} 1 | \dots | 1 | \langle x, y \rangle & \text{if } \alpha_i = \beta_i = 1, 1 \leq i \leq n, \\ 0 & \text{otherwise,} \end{cases}$$

where $\alpha_i, \beta_j \in \bar{\mathcal{R}}$ and $x, y \in L$. Here we use L to mean the underlying $\bar{\mathcal{R}}$ -module $U_{\text{Mod}_{\bar{\mathcal{R}}}}^{\text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}}(L)$.

Corollary 3.10 *For L a $\text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}$ -algebra with $\text{Lie}^{s,\text{ti}}$ -bracket $\langle -, - \rangle$, there is an isomorphism of bigraded homotopy groups*

$$\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}, L) \cong \pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, \text{AR}_\bullet(L)).$$

3.2 Homology groups of simplicial $\text{Lie}^{s,\text{ti}}$ -algebras

The homotopy groups of $\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, V_\bullet)$ for $V_\bullet$ a simplicial $\text{Lie}^{s,\text{ti}}$ -algebra can be computed via a shifted version of the classical Chevalley–Eilenberg complex.

Recall from [22; 45, Section 5; 51] that given a Lie^{ti} -algebra L , ie an unshifted totally isotropic ungraded Lie algebra over $\mathbb{F}_2$, its Lie^{ti} -algebra homology is computed by

$$H_*^{\text{Lie}^{\text{ti}}}(L) := H_*(\mathbb{L} Q_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^{\text{ti}}_{\mathbb{F}_2}}(L)[1] \oplus \mathbb{F}_2) = H_*(\text{CE}(L)).$$

Here $\text{CE}(L)$ is the standard Chevalley–Eilenberg complex, defined to be the exterior algebra on $L[1]$ with differential δ given by

$$\delta(\sigma x_1 \otimes \cdots \otimes \sigma x_n) = \sum_{1 \leq i < j \leq n} [\sigma x_i, \sigma x_j] \otimes \sigma x_1 \otimes \cdots \otimes \widehat{\sigma x_i} \otimes \cdots \otimes \widehat{\sigma x_j} \otimes \cdots \otimes \sigma x_n.$$

There is no divided power part at $p = 2$. Since we are working with shifted, graded totally isotropic Lie algebras, we use a modified version for ease of notation. First we note that given a $\text{Lie}^{s,\text{ti}}$ -algebra L , there are weak equivalences

$$(9) \quad N(\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, L)) \simeq N(\Sigma \text{Bar}_\bullet(\text{id}, \text{Lie}^{\text{ti}}_{\mathbb{F}_2}, \Sigma^{-1}L)) \simeq \Sigma \overline{\text{CE}}(\Sigma^{-1}L[1])[-1],$$

where $\overline{\text{CE}}$ is the reduced complex.

Definition 3.11 The *Chevalley–Eilenberg complex* for a $\text{Lie}^{s,\text{ti}}$ -algebra L is $\text{CE}(L) = (\Lambda^\bullet(L), \delta)$, where $\Lambda^\bullet(L)$ is the free shifted graded exterior algebra on L (placed in homological degree 0) with a shifted graded exterior product $\Sigma^{-1} \otimes [1]$, which we continue to denote by $\otimes$, that increases homological degree by one and decreases internal degree by one, reflecting the behavior of shifted graded Lie brackets. The differential δ is given by

$$\delta(x_1 \otimes \cdots \otimes x_n) = \sum_{1 \leq i < j \leq n} [x_i, x_j] \otimes x_1 \otimes \cdots \otimes \hat{x}_i \otimes \cdots \otimes \hat{x}_j \otimes \cdots \otimes x_n.$$

Then the $\text{Lie}^{s,\text{ti}}$ -algebra homology of L is given by the bigraded $\mathbb{F}_2$ -module

$$H_{*,*}^{\text{Lie}^{s,\text{ti}}}(L) := \pi_{*,*}(\mathbb{L} Q_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^{s,\text{ti}}}(L) \oplus \mathbb{F}_2) \cong H_{*,*}(N(\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, L) \oplus \mathbb{F}_2)) \cong H_{*,*}(\text{CE}(L)),$$

where the last isomorphism follows from rearranging the right hand side in (9).

In the case where L is a simplicial $\text{Lie}^{s,\text{ti}}$ -algebra, its Chevalley–Eilenberg complex $\text{CE}(L)$ is the simplicial chain complex obtained by applying the Chevalley–Eilenberg complex levelwise. Then Dold–Kan correspondence says that the homotopy groups of $\text{CE}(L)$ are isomorphic to the homology groups of its total complex. A simplicial version of May’s result is recorded in [17, Section 3]. Here we state the shifted version.

Theorem 3.12 [17, Theorem 3.13] *Let L be a simplicial $\text{Lie}^{s,\text{ti}}$ -algebra. Then there is a natural isomorphism of bigraded $\mathbb{F}_2$ -modules*

$$H_{*,*}^{\text{Lie}^{s,\text{ti}}}(L) := \pi_{*,*}(\mathbb{L} Q_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^{s,\text{ti}}}(L) \oplus \mathbb{F}_2) \cong H_{*,*}(\text{CE}(L)).$$

In the total complex of $\mathrm{CE}(L)$, the differential in the homological direction is given by δ in Definition 3.11. The differential d in the simplicial direction is obtained by applying the shifted graded exterior algebra functor $\Lambda^\bullet$ to each simplicial differential d_i of L and taking the alternating sum, ie

$$d = d_0 \otimes d_0 \otimes \cdots \otimes d_0 + \cdots + d_r \otimes d_r \otimes \cdots \otimes d_r.$$

Both differentials preserve weights.

If the $\mathrm{Lie}^{s, \mathrm{ti}}$ -bracket on a simplicial $\mathrm{Lie}^{s, \mathrm{ti}}$ -algebra L is trivial, then the differential δ in the homological direction vanishes and $H_{*,*}(\mathrm{CE}(L)) \cong \pi_{*,*}(\Lambda^\bullet(L))$. The natural operations on the homotopy groups of simplicial exterior algebras are well-understood by the work of Cartan, Bousfield, and Dwyer. We only state their results in the case of free algebras, and modify the grading to take into account the fact that we work with shifted, graded exterior algebras.

Theorem 3.13 [27, Theorem 2.1, Remark 4.4; 14; 20; 34, Theorem 3.9] *Let $V_\bullet$ be a simplicial graded $\mathbb{F}_2$ -module. There are natural operations*

$$\gamma_i : \pi_{h,r,t}(\Lambda^h(V_\bullet)) \rightarrow \pi_{2h+1,r+i,2t-1}(\Lambda^{2h+1}(V_\bullet)), \quad 1 \leq i \leq r$$

for all $r \geq 1$, satisfying the relations

$$\gamma_i \gamma_j(x) = \sum_{(i+1)/2 \leq l \leq (i+j)/3} \binom{j-i+l-1}{j-l} \gamma_{i+j-l} \gamma_l(x) \quad \text{for all } i < 2j.$$

Here in the trigrading (h, r, t) records the number of exterior products h , the simplicial degree r in $V_\bullet$, and the internal degree t .

Furthermore, they computed the homotopy groups of the free exterior algebra on a simplicial $\mathbb{F}_2$ -module.

Definition 3.14 A sequence $I = (i_1, \dots, i_m)$ is γ -admissible if $i_l \geq 2i_{l+1}$ for $1 \leq l \leq m-1$. The excess of I is $e(I) = i_1 - i_2 - \cdots - i_m$.

Theorem 3.15 [14, Theorem 8.6; 34, Theorem 3.19] *Let A be a graded $\mathbb{F}_2$ -basis for $\pi_*(V_\bullet)$. Then $\pi_{*,*}(\Lambda^\bullet(V_\bullet))$ is the (shifted graded) exterior algebra on generators $\gamma_I(\alpha)$, where $\alpha \in A$ and $I = (i_1, \dots, i_m)$ is γ -admissible with $e(I) \leq s(\alpha)$, where $s(\alpha)$ is the simplicial degree of the basis element α .*

The following is immediate by combining Theorems 3.12 and 3.15.

Corollary 3.16 *Suppose that L is a $\mathrm{Lie}_{\mathcal{R}}^{s, \mathrm{ti}}$ -algebra with trivial Lie brackets. Then*

$$\pi_{*,*} \mathrm{Bar}_\bullet(\mathrm{id}, \mathrm{Lie}^{s, \mathrm{ti}}, \mathrm{AR}_\bullet(L))$$

is isomorphic as a bigraded $\mathbb{F}_2$ -vector space to the (shifted graded) exterior algebra over $\mathbb{F}_2$ on generators $\gamma_I(\alpha)$, where α is a basis element of $\pi_{r,*}(\mathrm{AR}_\bullet(L))$ (see Construction 3.9) and I is γ -admissible with $e(I) \leq r$.

Now we can compute the homotopy groups of $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, L)$ when the $\text{Lie}^{s, \text{ti}}$ structure on L is trivial. First we recall the following result of Priddy that computes the Ext and Tor groups of a homogeneous Koszul algebra, which we make use of to compute the Tor groups over $\overline{\mathcal{R}}$.

Theorem 3.17 [51, Theorem 2.5] *Let R be a homogeneous Koszul algebra over $\mathbb{F}_2$ on generators $a_i, i \in J$ in weight 1 and quadratic relations r_j . Let B be a subset of the set S of nonempty sequences on J such that there is a basis of R consisting of monomials $\{a_I\}_{I \in S}$. Then the cohomology algebra $H^{*,*}(A) = \text{Ext}_R^{*,*}(\mathbb{F}_2, \mathbb{F}_2)$ is isomorphic to the tensor algebra on $a_i^\vee$ subject to relations that are linear dual to the r_j .*

We record an explicit description of the procedure of cycle completion that produces a given class in the Tor groups, which will be useful later.

Remark 3.18 Call $a_i a_j$ allowable if $(i, j) \in B$ and unallowable otherwise. Since we are working over $\mathbb{F}_2$ and the cohomology of $\overline{\mathcal{R}}$ as a bigraded $\mathbb{F}_2$ -module is finite in each bidegree, we are allowed identify the bigraded $\mathbb{F}_2$ -modules $\text{Tor}_{m,n}^R(\mathbb{F}_2, \mathbb{F}_2)$ with the $\mathbb{F}_2$ -linear dual of $\text{Ext}_R^{m,n}(\mathbb{F}_2, \mathbb{F}_2)$. To simplify notation, we will name classes in $\text{Tor}_{m,n}^R(\mathbb{F}_2, \mathbb{F}_2)$ by its the name of its dual in $\text{Ext}_R^{m,n}(\mathbb{F}_2, \mathbb{F}_2)$. A cycle corresponding to the class

$$a_{i_1}^\vee a_{i_2}^\vee \cdots a_{i_m}^\vee \in \text{Tor}_{m,*}^R(\mathbb{F}_2, \mathbb{F}_2)$$

with (i_k, i_{k+1}) unallowable for all k in the reduced bar complex over R is a sum $\sum_j [a_{j_1} | a_{j_2} | \cdots | a_{j_m}] \in R^{\otimes m}$ that contains the term $[a_{i_1} | a_{i_2} | \cdots | a_{i_m}]$ with nonzero coefficient. We call this the cycle completion of the monomial $[a_{i_1} | a_{i_2} | \cdots | a_{i_m}]$. To find the cycle explicitly, we start with $\alpha_0 = [a_{i_1} | a_{i_2} | \cdots | a_{i_m}]$. The differential ∂ is a sum of face maps composing adjacent terms $a_{i_k} a_{i_{k+1}}$. We use the relation $a_{i_k} a_{i_{k+1}} = \sum b_{j_k} b_{j_{k+1}}$ to cancel out the terms $[a_{i_1} | \cdots | a_{i_k} a_{i_{k+1}} | \cdots | a_{i_m}]$ in the differential by adding $\sum [a_{i_1} | \cdots | a_{j_{k-1}} | b_{j_k} b_{j_{k+1}} | a_{j_{k+2}} | \cdots | a_{i_m}]$ to α_0 for all k and denote the resulting sum α_1 . Then we pair off the differential for every term in $\alpha_1 - \alpha_0$, ie for each nonzero term in $\partial(\alpha_1 - \alpha_0)$ obtained by composing an unallowable 2-tuple via the k^{th} face map, we use the relations in R to find a sum in $R^{\otimes m}$ whose image under the k^{th} face map cancel out that term. Thus we obtain a new sum α_2 such that all terms in the differential on α_1 are paired off. Now we repeat the process again. It has to terminate since the number of unallowable adjacent pairs is nonincreasing for any term at each step and there are finitely many monomials with a given number of unallowable adjacent pairs. In other words, $a_{i_1} a_{i_2} \cdots a_{i_m}$ can be written as a unique sum of basis monomials through this iterative process in finite steps.

Lemma 3.19 (1) *Suppose that $L = \Sigma^k \mathbb{F}_2$ is a trivial $\text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}$ -algebra. Then $\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \text{AR}_\bullet(L))$ is the exterior algebra over $\mathbb{F}_2$ on generators $\gamma_I \overline{Q}^J(x_k)$, where x_k is the generator of $\pi_*(L)$, $J = (j_1, \dots, j_r)$ satisfies*

$$j_{l+1} + \cdots + j_r + k - (r - l) \leq j_l \leq 2j_{l+1}$$

for $1 \leq l < r$ and $j_r > k$, and I is γ -admissible with $e(I) \leq r$. In lower indexing, the generators are $\gamma_I \overline{Q}_J(x_k)$, where $J = (j_1, \dots, j_r)$ satisfies $0 \leq j_l \leq j_{l+1} + 1$ for all l , and I is γ -admissible with $e(I) \leq r$.

- (2) Let L be the $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebra with underlying $\bar{\mathcal{R}}$ -module $\Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}(\Sigma^{n+k} \mathbb{F}_2)}$, $n \geq 1$ and trivial Lie brackets. Then $\pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}, \text{AR}_{\bullet}(L))$ is the exterior algebra over $\mathbb{F}_2$ on generators $\gamma_I \bar{Q}_J(x_k)$, where $J = (j_1, \dots, j_r)$ satisfies $0 \leq j_l \leq j_{l+1} + 1$ for all $l < r$ and $0 \leq j_r < n$, and I is γ -admissible with $e(I) \leq r$.

Proof (1) In light of [Corollary 3.16](#), it suffices to compute

$$\pi_{*,*}(\text{AR}_{\bullet}(L)) = \pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \mathcal{A}_{\bar{\mathcal{R}}}, \Sigma^k \mathbb{F}_2),$$

where the right hand side is the unstable Tor groups $\text{UnTor}_{*,*}^{\bar{\mathcal{R}}}(\mathbb{F}_2, \Sigma^k \mathbb{F}_2)$. The unstable Tor groups are computed by taking the homotopy groups of the subcomplex of the bar complex computing the Tor groups $\text{Tor}_{*,*}^{\bar{\mathcal{R}}}(\mathbb{F}_2, \Sigma^k \mathbb{F}_2)$ obtained by regarding $\Sigma^k \mathbb{F}_2$ as an unstable trivial module over $\bar{\mathcal{R}}$ and imposing the instability conditions $[\bar{Q}^j | \alpha] = 0$ for $j \leq |\alpha|$; see [\[15, Section 3\]](#).

The quadratic algebra $\bar{\mathcal{R}}$ is a homogeneous Koszul algebra, since the canonical basis

$$\{\bar{Q}^{j_1} \dots \bar{Q}^{j_r}, j_i > 2j_{i+1} \forall i\}$$

of $\bar{\mathcal{R}}$ is a Poincaré–Birkhoff–Witt basis in the sense of Priddy [\[51, Theorem 5.3\]](#). In particular, it follows from Priddy’s machinery [\[51, Theorems 2.5, 3.8\]](#) that the Tor group $\text{Tor}_{s,*}^{\bar{\mathcal{R}}}(\mathbb{F}_2, \mathbb{F}_2)$ has a basis consisting of cycles indexed by $\bar{Q}^{j_1} \dots \bar{Q}^{j_s}$, where $j_i \leq 2j_{i+1}$ for all i .

To compute the unstable Tor groups on a class x_k of internal degree k , we need to impose the instability condition $\bar{Q}^j(x) = 0$ for $j < |x|$, then the basis of $\text{UnTor}_{r,*}^{\bar{\mathcal{R}}}(\mathbb{F}_2, \mathbb{F}_2\{x_k\})$ consists of basis elements of $\text{Tor}_{r,*}^{\bar{\mathcal{R}}}(\mathbb{F}_2, \mathbb{F}_2)$ satisfying $j_i > j_{i-1} - 1 + j_{i-2} - 1 + \dots + j_r - 1 + |x|$ for all $i < r$ and $j_r \geq k$, or equivalently sequences $\bar{Q}_{j_1} \dots \bar{Q}_{j_s}(x_k)$, where $0 \leq j_i \leq j_{i+1} + 1$ for all i .

- (2) Iterating [Proposition 2.10](#) yields a canonical map of $\bar{\mathcal{R}}$ -modules

$$L = \Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}(\Sigma^{n+k} \mathbb{F}_2)} \rightarrow \Omega^\infty \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}(\Sigma^\infty \Sigma^k \mathbb{F}_2)} \cong \Sigma^k \mathbb{F}_2,$$

which gives rise to a surjective map of $\text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}$ -algebras with trivial brackets. The underlying $\mathbb{F}_2$ -module of L has basis $\bar{Q}^J x_k$, where $J = (j_1, \dots, j_r)$ is a basis element of $\bar{\mathcal{R}}$ satisfying $j_r \geq n + k$. Suppose that $\alpha \in \text{AR}_{\bullet}(L)$ is the cycle completion of an element $\bar{Q}^{j_1} | \dots | \bar{Q}^{j_r} | x_k$ with $k \leq j_r < n + k$ and $j_{l+1} - 1 + \dots + j_r - 1 + k \leq j_l \leq 2j_{l+1}$ for $l < r$. Since cycle completion via Behrens’ relations in the sense of [Remark 3.18](#) cannot increase the index of the right most operation, the differentials supported by α are the same as those supported by its image in $\text{AR}_{\bullet}(\mathbb{F}_2\{x_k\})$, so α is a nontrivial cycle. Otherwise, all but the rightmost face maps send α to zero, while the rightmost face map from at least one (distinct) term of α is nonzero, so it is impossible to complete the cycle. Switching to lower-indexing yields the desired answer. $\square$

Combining [Theorem 3.15](#) and [Lemma 3.19](#), we have:

Corollary 3.20 For $\mathfrak{g} = \Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k}\mathbb{F}_2)$ with $1 \leq n \leq \infty$, the E^1 -page $\pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}, \mathfrak{g})$ of the algebraic γ_1 -Bockstein spectral sequence (see [Theorem 3.6](#)) is the (shifted graded) exterior algebra on generators $\gamma_I \bar{Q}_J(x_k)$, where $I = (i_1, \dots, i_m)$ is γ -admissible with $e(I) \leq r$ and $J = (j_1, \dots, j_r)$ satisfies $0 \leq j_l \leq j_{l+1} + 1$ for $l < r$, $0 \leq j_r < n$.

3.3 Quillen homology of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras with trivial brackets

Next we want to identify the differentials in the May-type spectral sequence and the γ_1 -Bockstein spectral sequence when $\mathfrak{g} = \Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k}\mathbb{F}_2)$. There is no canonical map from $\pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$ to the E^1 -page $\pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s, \text{ti}}, \mathfrak{g})$ of the γ_1 -Bockstein spectral sequence; instead we map $\text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g})$ into the bar construction of another variant of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras.

Definition 3.21 Let $\text{Mod}_{\bar{\mathcal{R}}_{>0}} \subset \text{Mod}_{\bar{\mathcal{R}}}$ be the subcategory of allowable $\bar{\mathcal{R}}$ -modules M such that $\bar{Q}_0(x) = 0$ for all $x \in M$. Denote by $\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}_{>0}}}$ the free $\bar{\mathcal{R}}_{>0}$ -module functor, and $\mathcal{A}_{\bar{\mathcal{R}}_{>0}}$ the additive monad associated to the free functor. Let $\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}} = \mathcal{A}_{\bar{\mathcal{R}}_{>0}} \circ \text{Lie}^{s, \text{ti}}$, where the composite monad on the right has distributivity given by $[\bar{Q}^j(-), (-)] = 0$.

By [Proposition 2.8](#), there is an equivalence of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras

$$\text{Lie}_{\bar{\mathcal{R}}}^s(M) = \mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^s(M) / (\bar{Q}_0(x) = [x, x], x \in M).$$

In comparison, there is an equivalence of $\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}$ -algebras

$$\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}(M) = \mathcal{A}_{\bar{\mathcal{R}}_{>0}} \circ \text{Lie}^{s, \text{ti}}(M) = \mathcal{A}_{\bar{\mathcal{R}}} \circ \text{Lie}^s(M) / \langle \bar{Q}_0(x), [x, x], x \in M \rangle,$$

where the quotient is taken with respect to the left $\bar{\mathcal{R}}$ -algebra ideal. Hence the category $\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}$ of $\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}$ -algebras is the subcategory of $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebras L satisfying the condition that $\bar{Q}_0(x) = [x, x] = 0$ for all $x \in L$. The inclusion

$$T_{\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\mathfrak{g}) : \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}} \rightarrow \text{Lie}_{\bar{\mathcal{R}}}^s$$

of subcategory admits a left adjoint $Q_{\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\mathfrak{g})$ that takes the quotient by the $\bar{\mathcal{R}}$ -algebra ideal of the self-brackets. When $\mathfrak{g}$ is a $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra with trivial Lie^s brackets, $Q_{\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\mathfrak{g})$ is given by equipping the $\bar{\mathcal{R}}_{>0}$ -module $Q_{\text{Mod}_{\bar{\mathcal{R}}_{>0}}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\mathfrak{g})$ with trivial $\text{Lie}^{s, \text{ti}}$ brackets.

Lemma 3.22 Let $\mathfrak{g}$ be an $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra. There is a surjective map of simplicial $\mathbb{F}_2$ -modules

$$\varphi : \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g}) \rightarrow \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}, Q_{\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\mathfrak{g})).$$

Proof There is a map of monads $\text{Lie}_{\bar{\mathcal{R}}}^s \rightarrow \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}$ that sends the symbol $\bar{Q}_0$ to 0, and this induces the map of bar constructions in question. $\square$

The homotopy group of $\text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \text{ti}}, L)$ is computed in the same way as for $\text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, L)$ via [Lemmas 2.13](#) and [3.19](#) ($\bar{Q}_0$ operation no longer appears in the generators).

Construction 3.23 For L a $\text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s,\text{ti}}$ -algebra with $\text{Lie}^{s,\text{ti}}$ -bracket $\langle -, - \rangle$, denote by $\text{AR}_{\bullet}^{>0}(L)$ the bar construction $\text{Bar}_{\bullet}(\text{id}, \mathcal{A}_{\bar{\mathcal{R}}_{>0}}, L)$ equipped with the simplicial $\text{Lie}^{s,\text{ti}}$ -algebra structure given levelwise by

$$\langle \alpha_1 | \alpha_2 | \dots | \alpha_n | x, \beta_1 | \beta_2 | \dots | \beta_n | y \rangle = \begin{cases} 1 | \dots | 1 | \langle x, y \rangle & \text{if } \alpha_i = \beta_i = 1, \ 1 \leq i \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3.24 (1) *There is an isomorphism*

$$\begin{aligned} \pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s,\text{ti}}, \Sigma^k \mathbb{F}_2) &\cong \pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}^{s,\text{ti}}, \text{Bar}_{\bullet}(\text{id}, \mathcal{A}_{\bar{\mathcal{R}}_{>0}}, \Sigma^k \mathbb{F}_2)) \\ &\cong \pi_{*,*} \Lambda^{\bullet}(\text{UnTor}_{*,*}^{\bar{\mathcal{R}}_{>0}}(\mathbb{F}_2, \mathbb{F}_2\{x_k\})). \end{aligned}$$

Hence $\pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s,\text{ti}}, \Sigma^k \mathbb{F}_2)$ is the exterior algebra on generators $\gamma_I \bar{Q}_J(x_k)$, where $J = (j_1, \dots, j_r)$ satisfies $1 \leq j_l \leq j_l + 1$ for all l and I is γ -admissible with $e(I) \leq r$.

- (2) *The homotopy group of $\text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s,\text{ti}}, \Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}_{>0}}}(\Sigma^{n+k} \mathbb{F}_2))$ is the exterior algebra on generators $\gamma_I \bar{Q}_J(x_k)$, where $J = (j_1, \dots, j_r)$ satisfies $1 \leq j_r < n$ and $1 \leq j_l \leq j_l + 1$ for $l < r$, while I is γ -admissible with $e(I) \leq r$.*
- (3) *For $L = \Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}_{>0}}}(\Sigma^{n+k} \mathbb{F}_2)$ with $1 \leq n \leq \infty$, the quotient map of monads $\mathcal{A}_{\bar{\mathcal{R}}} \rightarrow \mathcal{A}_{\bar{\mathcal{R}}_{>0}}$ induces a surjective map $\pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^{s,\text{ti}}, L) \rightarrow \pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s,\text{ti}}, L)$ that sends the symbol $\bar{Q}_0$ to 0.*

In order to use the comparison map (see [Lemma 3.22](#))

$$\varphi_*: \pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \mathfrak{g}) \rightarrow \pi_{*,*} \text{Bar}_{\bullet}(\text{id}, \text{Lie}_{\bar{\mathcal{R}}_{>0}}^{s,\text{ti}}, \mathfrak{g})$$

to detect differentials and permanent cycles, we make use of explicit combinatorial formulae of γ_i by Bökstedt and Ottosen. The grading conventions are modified to suit our context.

For $r, i \in \mathbb{N}$ with $1 \leq i \leq r$, let $U(r, i)$ be the set of pairs (A, B) of ordered sequences $a_1 < \dots < a_i$, $b_1 < \dots < b_i$ such that $\{a_1, \dots, a_i\}$ and $\{b_1, \dots, b_i\}$ are complementary subsets of

$$\{r - i, r - i + 1, \dots, r + i - 1\}.$$

Let $V(r, i) \subset U(r, i)$ be the subset with $a_1 = r - i$.

Proposition 3.25 [[13](#), Theorem 1.3, Lemma 3.1] *Suppose that $V_{\bullet}$ is a simplicial $\mathbb{F}_2$ -module with face maps d_j . Let z be a representative of a class $[z] \in \pi_{s,t}(V_{\bullet})$ in the normalized complex $N(V_{\bullet})$. For $2 \leq i \leq s$, define*

$$\gamma_i(z) = \sum_{(A,B) \in V(s,i)} s_{a_i} \cdots s_{a_2} s_{a_1}(z) \otimes s_{b_i} \cdots s_{b_2} s_{b_1}(z) \in \Lambda^2(V_{\bullet}).$$

Then $d_j(\gamma_i(z)) = 0$ for $0 \leq j \leq i + s$, and the induced operation $\gamma_i: \pi_{s,t}(V_{\bullet}) \rightarrow \pi_{s+i+1,2t-1}(\Lambda^2(V_{\bullet}))$ are exactly the Dwyer–Bousfield operations in [Theorem 3.15](#).

Remark 3.26 If in addition $V_\bullet$ is exterior, then the formula above for $i = 1$ induces the operation γ_1 on $\pi_{*,*}(V_\bullet)$. The operation γ_1 is not well-defined when there is some element a in the simplicial commutative algebra $V_\bullet$ such that $a \otimes a \neq 0$. This is because in $N(V_\bullet)$ the differential sends $\gamma_1(a)$ to $a \otimes a$; see [27, Remarks 4.3, 4.4; 13, Remark 3.2].

Hence we obtain natural operations γ_i for $1 \leq i \leq s$ on

$$\pi_{s,*}(\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, \text{AR}_\bullet^{>0}(\Sigma^k \mathbb{F}_2))) \cong \pi_{s,*}(\Lambda^\bullet(\text{Bar}_\bullet(\text{id}, \mathcal{A}_{\bar{\mathcal{R}}_{>0}}, \Sigma^k \mathbb{F}_2))),$$

and similarly on

$$\pi_{s,*}(\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, \text{AR}_\bullet(\Sigma^k \mathbb{F}_2))) \cong \pi_{s,*}(\Lambda^\bullet(\text{Bar}_\bullet(\text{id}, \mathcal{A}_{\bar{\mathcal{R}}}, \Sigma^k \mathbb{F}_2))).$$

Suppose that ξ is a cycle in $\text{AR}_s^{>0}(\Sigma^k \mathbb{F}_2)$. In the total complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, \text{AR}_\bullet^{>0}(\Sigma^k \mathbb{F}_2))$, a representative for the homotopy class $\gamma_i([\xi])$ is

$$\gamma_i(\xi) = \sum_{(A,B) \in V(s,i)} \langle s_{a_i} \cdots s_{a_2} s_{a_1}(\xi), s_{b_1} s_{b_2} \cdots s_{b_i}(\xi) \rangle \in \text{Lie}^{s,\text{ti}} \circ (\mathcal{A}_{\bar{\mathcal{R}}_{>0}})^{\circ(s+i)}(\Sigma^k \mathbb{F}_2).$$

When we iterate the γ_i operations, the formula is harder to write down explicitly.

Notation 3.27 Suppose that $V_\bullet$ is a simplicial $\mathbb{F}_2$ -module as a trivial simplicial $\text{Lie}^{s,\text{ti}}$ -algebra. For distinct classes $[\xi_1], \dots, [\xi_n] \in \pi_{*,*}(V_\bullet)$, denote by $B(\xi_1, \dots, \xi_n)$ the cycle in the normalized complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, V_\bullet)$ that represents the class $[\xi_1] \otimes \cdots \otimes [\xi_n] \in \pi_{*,*}(\Lambda^{n-1}(V_\bullet)) \subset \pi_{*,*}(\text{CE}(V_\bullet)) \cong \pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, V_\bullet)$, which is obtained by cycle completion via the Jacobi identity in the sense of Remark 3.18.

Therefore a homotopy class $[\xi_1] \otimes \cdots \otimes [\xi_l]$ with $l > 1$ in $\pi_{s,*}(\Lambda^\bullet(\text{Bar}_\bullet(\text{id}, \mathcal{A}_{\bar{\mathcal{R}}_{>0}}, \Sigma^k \mathbb{F}_2)))$ is represented by an element $B(\xi_1, \dots, \xi_l)$ in the summand $(\text{Lie}^{s,\text{ti}})^{\circ(l-1)} \circ (\mathcal{A}_{\bar{\mathcal{R}}_{>0}})^{\circ(s-l+1)}(\Sigma^k \mathbb{F}_2)$ of the total complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, \text{AR}_\bullet^{>0}(\Sigma^k \mathbb{F}_2))$. Since a representative for the homotopy class $\gamma_j \gamma_i(\xi)$ in the total complex of $\Lambda^\bullet(\text{Bar}_\bullet(\text{id}, \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}_{>0}}}, \Sigma^k \mathbb{F}_2))$ is given by

$$\gamma_j \gamma_i(\xi) = \sum_{(C,D) \in V(s+i+1,j)} \sum_{(A,B) \in V(s,i)} s_C(s_A(\xi) \otimes s_B(\xi)) \otimes s_D(s_A(\xi) \otimes s_B(\xi)),$$

a representative for $\gamma_j \gamma_i(\xi)$ in the total complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}^{s,\text{ti}}, \text{AR}_\bullet^{>0}(\Sigma^k \mathbb{F}_2))$ is given the sum of over all $(A, B) \in V(s, i)$, $(C, D) \in V(s + i + 1, j)$ of $B(s_C s_A(\xi), s_C s_B(\xi), s_D s_A(\xi), s_D s_B(\xi))$, with the three brackets coming from distinct simplicial filtrations.

Theorem 3.28 *The Quillen homology*

$$\text{HQ}_{*,*}^{\text{Lie}_{\bar{\mathcal{R}}}^s}(\Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2)) \cong \pi_{s,t} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\bar{\mathcal{R}}}^s, \Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2))$$

of the $\text{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\Omega^n \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Mod}_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2)$, $1 \leq n \leq \infty$ is isomorphic as a bigraded vector space to the exterior algebra on generators $\gamma_I \bar{Q}_J(x_k)$, where $I = (i_1, \dots, i_m)$ is γ -admissible with $e(I) \leq r$ and $i_m \geq 2$, whereas $J = (j_1, \dots, j_r)$ satisfies $0 \leq j_l \leq j_{l+1} + 1$ for $l < r$, $0 \leq j_r < n$ and if $j_1 = 0$ then either $r = 1$ or $i_m = 2$.

Recall from [Proposition 2.10](#) that in the case $n = \infty$,

$$\Sigma^{n+k}\mathbb{F}_2 \simeq \lim_{n \rightarrow \infty} \Omega^n \operatorname{Free}_{\operatorname{Mod}\mathbb{F}_2}^{\operatorname{Mod}\bar{\mathcal{R}}}(\Sigma^{n+k}\mathbb{F}_2)$$

is the trivial $\operatorname{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\Sigma^k\mathbb{F}_2$.

Before we proceed to prove the theorem, we provide some intuition about the strategy. Since the input $\operatorname{Lie}_{\bar{\mathcal{R}}}^s$ -algebra $\Omega^n \operatorname{Free}_{\operatorname{Mod}\mathbb{F}_2}^{\operatorname{Mod}\bar{\mathcal{R}}}(\Sigma^{n+k}\mathbb{F}_2)$ has vanishing Lie brackets, [Remark 3.8](#) allows us to consider a single May-type spectral sequence by considering the length filtration on $\Omega^n \operatorname{Free}_{\operatorname{Mod}\mathbb{F}_2}^{\operatorname{Mod}\bar{\mathcal{R}}}(\Sigma^{n+k}\mathbb{F}_2)$ shifted up by two. From the construction of the May-type spectral sequence, we see that there is a higher differential on a class on the E^1 -page $\cong \pi_{*,*}(\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^{s, \operatorname{ti}}, L))$ if and only if its representative cycle, considered as an element in $\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^s, L)$, admits a face map that evaluates a non-self-bracket to a self-bracket. [Remark 3.26](#) and [Corollary 3.20](#) indicate that γ_1 is the only operation that arises in $\pi_{*,*}(\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^{s, \operatorname{ti}}, L)) \cong \Lambda\{\gamma_I \bar{Q}_J(x)\}$ with I γ -admissible precisely because self-brackets are zero in $\operatorname{Lie}_{\bar{\mathcal{R}}}^{s, \operatorname{ti}}$ -algebras and thus generates all the differentials in the May-type spectral sequence. Hence we expect $\pi_{*,*} \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^s, L)$ to be a quotient of $\pi_*(\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^{s, \operatorname{ti}}, L))$ (see [Corollary 3.20](#)) by a suitable ideal generated by $\gamma_1(\alpha)$ for all $\alpha \in \pi_{*,*}(\operatorname{AR}_{\bullet}(L))$, and we use the induced map on homotopy groups of $\varphi: \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^s, L) \rightarrow \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \operatorname{ti}}, L)$ from [Lemma 3.22](#) to help detect the differentials and permanent cycles.

Proof of Theorem 3.28 We focus on the case $L = \Sigma^k\mathbb{F}_2$, since in the cases $n < \infty$ the only difference is an extra condition on the rightmost operation in basis elements, so the same argument applies with no change.

Consider the map $\varphi_*: \pi_{*,*} \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^s, L) \rightarrow \pi_{*,*} \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \operatorname{ti}}, L)$ from [Lemma 3.22](#). Its cokernel consists of all cycles in $\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \operatorname{ti}}, \Sigma^k\mathbb{F}_2)$ whose preimage is the source of a differential to an element that is in the kernel of φ . Since φ is surjective by [Lemma 3.22](#), this is equivalent to finding all classes α that are cycles in $\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \operatorname{ti}}, \Sigma^k\mathbb{F}_2)$ precisely because the differential ∂' in the normalized complex of $\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \operatorname{ti}}, \Sigma^k\mathbb{F}_2)$ sends α to a linear combination of elements that contain self-brackets or $\bar{Q}_0$. In other words, via the inclusion to $\pi_{*,*} \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^{s, \operatorname{ti}}, \Sigma^k\mathbb{F}_2)$ in [Lemma 3.24\(3\)](#), all elements in the cokernel of φ_* support differentials in the May spectral sequence.

We start with the generators of the exterior algebra, see [Lemma 3.19](#). Let $[\alpha] = \bar{Q}_{j_1} \bar{Q}_{j_2} \cdots \bar{Q}_{j_r}(x_k)$ be a basis element of $\pi_{*,*} \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^{s, \operatorname{ti}}, \operatorname{AR}_{\bullet}^{>0}(\Sigma^k\mathbb{F}_2))$, represented by a cycle

$$\alpha = \bar{Q}_{j_1} | \dots | \bar{Q}_{j_r} | x_k + \sum_l \bar{Q}_{j'_1} | \dots | \bar{Q}_{j'_l} | x_k \quad \text{in } \operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \operatorname{ti}}, \Sigma^k\mathbb{F}_2).$$

The terms in the summation comes from cycle completion via Behrens' relations in the sense of [Remark 3.18](#), with the condition that any term containing $\bar{Q}_0$ is 0. It has preimage $\tilde{\alpha}$ the cycle completion of $\bar{Q}_{j_1} | \dots | \bar{Q}_{j_r} | x_k$ in $\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}}^s, \Sigma^k\mathbb{F}_2)$ via Behrens' relations, which is the sum of α and terms $\bar{Q}_{j'_1} | \dots | \bar{Q}_{j'_l} | x_k$ such that at least one of the $\bar{Q}_{j'_l}$, $l > 1$ is equal to $\bar{Q}_0$. By [\[13, Lemma 3.1\]](#), the differential ∂ in the normalized complex of $\operatorname{Bar}_{\bullet}(\operatorname{id}, \operatorname{Lie}_{\bar{\mathcal{R}}_{>0}}^{s, \operatorname{ti}}, \Sigma^k\mathbb{F}_2)$ sends $\gamma_i(\alpha)$, $i \geq 2$ to zero because the terms are either zero or cancel out in pairs due to the simplicial identities of face and degeneracy maps.

Hence its preimage $\gamma_i(\tilde{\alpha})$ is also a cycle in the normalized complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)$ and hence a permanent cycle in the May spectral sequence. Similarly, for any γ -admissible sequence $I = (i_1, \dots, i_m)$ with $i_m \geq 2$, $\gamma_I(\alpha)$ lifts to a cycle $\gamma_I(\tilde{\alpha})$ in $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)$ and hence a permanent cycle in the May spectral sequence. By naturality of the γ_i operations and Lemma 3.24(3), the class $\gamma_I(\alpha)$ with $i_m \geq 2$ and $\alpha \in \pi_{*,*}(\text{AR}_\bullet(\Sigma^k \mathbb{F}_2))$ is also a permanent cycle.

On the other hand, the differential ∂ sends $\gamma_1(\alpha)$ to $\langle \alpha, \alpha \rangle = 0$ in $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}_{>0}}^{s, \text{ti}}, \Sigma^k \mathbb{F}_2)$, whereas its preimage $\gamma_1(\tilde{\alpha}) = [s_0 \tilde{\alpha}, s_1 \tilde{\alpha}]$ maps to $[\tilde{\alpha}, \tilde{\alpha}] = \overline{Q}_0 | \tilde{\alpha}$ under the differential in $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)$, which is in the kernel of φ . In other words, there is a differential in the May-type spectral sequence from $\gamma_1(\alpha) \in \pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \Sigma^k \mathbb{F}_2)$ to $\overline{Q}_0 \alpha$. Similarly, for any γ -admissible sequence $I = (i_1, \dots, i_m)$ with $i_m \geq 2$, $\gamma_I \gamma_1(\alpha)$ is a cycle in $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}_{>0}}^{s, \text{ti}}, \Sigma^k \mathbb{F}_2)$ because of the self-bracket in $\partial \gamma_I \gamma_1(\alpha) = \gamma_I(\partial(\gamma_1(\alpha)))$ if the simplicial degree of α is $r > 1$ and

$$\partial \gamma_I \gamma_1(\alpha) = \partial(\gamma_1(\alpha)) \otimes \gamma_1(\alpha) \otimes \gamma_2 \gamma_1(\alpha) \otimes \cdots \otimes \gamma_{2m-1} \cdots \gamma_2 \gamma_1(\alpha)$$

if $r = 1$; see [34, 3.9(i)]. On the other hand, its preimage $\gamma_I \gamma_1(\tilde{\alpha})$ is mapped by the total differential in $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)$ to the cycle completion $B(\overline{Q}_0 | \tilde{\alpha}, \gamma_1(\tilde{\alpha}), \dots, \gamma_{2m-1} \cdots \gamma_2 \gamma_1(\tilde{\alpha}))$ (see Notation 3.27) if $r = 1$, and to $\gamma_I([\tilde{\alpha}, \tilde{\alpha}])$ when $r > 1$. Note that in $\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \Sigma^k \mathbb{F}_2) \cong \Lambda\{\gamma_I \overline{Q}_J(x_k)\}$ with I γ -admissible and J satisfying certain conditions, we have $[\gamma_I([\tilde{\alpha}, \tilde{\alpha}])] = [\gamma_{I'}(\overline{Q}_0 | \tilde{\alpha})]$ with $I' = (i_1 + 2^{m-1}, \dots, i_m + 1)$. There is a shift in the indexing of the γ operations because by construction the self-brackets appearing in the same bracket term live in distinct filtrations when more γ s are applied, so replacing each self-bracket by a $\overline{Q}_0$ in a cycle will increase the index of the acting γ_i by one. In particular, we note that $i_m + 1 \geq 3$. Hence there is a differential in the May-type spectral sequence from $\gamma_I \gamma_1(\alpha)$ to $\gamma_{I'}(\overline{Q}_0 | \alpha)$, and all the generators $\gamma_I \gamma_1(\alpha)$ of the exterior algebra $\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \text{AR}_\bullet^{>0}(\Sigma^k \mathbb{F}_2))$ are in the cokernel of φ_* . Again by naturality of the γ_i operations and Lemma 3.24(3), the class $\gamma_I \gamma_1(\alpha) \in \pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \Sigma^k \mathbb{F}_2)$ supports a differential to $\gamma_{I'}(\overline{Q}_0 \alpha)$ in the May-type spectral sequence.

In general, suppose $[\alpha]$ is a basis element of

$$\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \Sigma^k \mathbb{F}_2) \cong \pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \text{AR}_\bullet(\Sigma^k \mathbb{F}_2))$$

that is the exterior product of generators $\gamma_{I_1}([\alpha_1]), \dots, \gamma_{I_n}([\alpha_n])$ with each α_i the cycle completion of a basis element $[\alpha_i] \in \pi_{*,*} \text{AR}_\bullet(\Sigma^k \mathbb{F}_2)$. It is represented by a cycle $\alpha = B(\gamma_{I_1}(\alpha_1), \dots, \gamma_{I_n}(\alpha_n))$ in the total complex of $\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \text{AR}_\bullet(\Sigma^k \mathbb{F}_2))$, see Notation 3.27, since $d_j(\gamma_{I_l}(\alpha_l)) = 0$ for all j and l by Proposition 3.25. Then $[\alpha]$ supports a differential in the May-type spectral sequence if and only if at least one of the γ -admissible sequences I_l is of the form $I_l = (i_{l1}, \dots, i_{lm}, 1)$. By Corollary 3.20, the above covers all classes in the $\mathbb{F}_2$ -basis of the E^1 -page of $\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^{s, \text{ti}}, \Sigma^k \mathbb{F}_2)$. $\square$

Remark 3.29 Note that $\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)$ is the cofree coalgebra on $\Sigma^k \mathbb{F}_2$ over the comonad

$$|\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, -)| := \pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Free}_{\text{Lie}_{\overline{\mathcal{R}}}^s}, -)$$

on $\text{Mod}_{\mathbb{F}_2}$. The coalgebra structure map is given by

$$\begin{aligned} |\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)| &\xleftarrow{\simeq} |\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, |\text{Bar}_\bullet(\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\overline{\mathcal{R}}}^s}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)|)| \\ &\rightarrow |\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, |\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)|)|, \end{aligned}$$

where the last map makes use of the augmentation $\text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\overline{\mathcal{R}}}^s} \rightarrow \text{id}$; see [16, Appendix D]. In particular, $\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \Sigma^k \mathbb{F}_2)$ records all natural unary operations on a degree k class in the mod 2 Quillen homology of $\text{Lie}_{\overline{\mathcal{R}}}^s$ -algebras, and Theorem 3.28 gives us a dimension count.

4 Application to the Knudsen spectral sequence

The rest of the paper is devoted to studying the mod p homology of labeled configuration spaces using the computation of Quillen homology of spectral Lie algebras. The coefficients for homology is $\mathbb{F}_2$ unless otherwise specified.

Let M be a manifold of dimension n and X a spectrum. The configuration space of k points in M labeled by X is the spectrum

$$B_k(M; X) = \Sigma_+^\infty \text{Conf}_k(M) \otimes_{\Sigma_k} X^{\otimes k},$$

considered as a weighted spectra of weight k . Here $\text{Conf}_k(M)$ is the space of k -tuples of pairwise distinct points in M . Denote by $s\mathcal{L}$ the monad associated to the free spectral Lie algebra functor $\text{Free}^{s\mathcal{L}}$. The ∞ -category of spectral Lie algebras is cotensored in Spaces , and we write $(-)^{M^+}$ for the cotensor with the one-point compactification of M in this category. In [39], Knudsen established the following equivalence using factorization homology; see [17, Theorem 5.1].

Theorem 4.1 [39, Section 3.4] *Let M be a parallelizable n -manifold and X a spectrum. Consider X as a weighted spectrum of weight one. Then there is an equivalence of weighted spectra*

$$\bigoplus_{k \geq 1} B_k(M; X) \simeq |\text{Bar}_\bullet(\text{id}, s\mathcal{L}, \text{Free}^{s\mathcal{L}}(\Sigma^n X)^{M^+})|.$$

The left hand side is weighted by the index k ; the weight filtration on the right hand side is given by propagating the weight on X via the free spectral Lie operad functor.

Applying the bar spectral sequence (Proposition 2.16) to the bar construction on the right hand side, we obtain the following:

Proposition 4.2 *There is a weighted spectral sequence*

$$(10) \quad E_{s,t}^2 = \text{HQ}_{s,t}^{\text{Lie}_{\overline{\mathcal{R}}}^s} (H_*(\text{Free}^{s\mathcal{L}}(\Sigma^n X)^{M^+})) \Rightarrow \bigoplus_{k \geq 1} H_{s+t}(B_k(M; X)).$$

The $\text{Lie}_{\overline{\mathcal{R}}}^s$ -algebra structure on the $\mathbb{F}_2$ -module

$$H_*(\text{Free}^{s\mathcal{L}}(\Sigma^n X)^{M^+}) \cong \tilde{H}^*(M^+) \otimes H_*(\text{Free}^{s\mathcal{L}}(\Sigma^n X)) \cong \tilde{H}^*(M^+) \otimes \text{Free}_{\text{Mod}\mathbb{F}_2}^{\text{Lie}_{\overline{\mathcal{R}}}^s}(H_*(\Sigma^n X))$$

has an explicit description.

Proposition 4.3 [17, Proposition 5.9] *Let $\mathfrak{g}$ be a spectral Lie algebra. Then there is a spectral Lie algebra structure on the cotensor $\mathfrak{g}^{M^+}$ in the category of spectra. The weight two structural map factors as*

$$\partial_2(\text{Id}) \otimes (\mathbb{D}(M^+) \otimes \mathfrak{g})_{h\Sigma_2}^{\otimes 2} \rightarrow \mathbb{D}(M^+)_{h\Sigma_2}^{\otimes 2} \otimes (\partial_2(\text{Id}) \otimes \mathfrak{g}_{h\Sigma_2}^{\otimes 2}) \xrightarrow{\mathbb{D}(\delta^*) \otimes \xi_*} \mathbb{D}(M^+) \otimes \mathfrak{g},$$

where $\mathbb{D}$ is the Spanier–Whitehead dual and δ the diagonal embedding.

As a result, the shifted Lie bracket on $\tilde{H}^*(M^+) \otimes H_*(\mathfrak{g})$ is given by

$$[y_1 \otimes x_1, y_2 \otimes x_2] := (y_1 \cup y_2) \otimes [x_1, x_2].$$

On the other hand, the Steenrod operations on $H^*(M^+)$ induces a twisted $\overline{\mathcal{R}}$ -module structure in the cotensor.

Proposition 4.4 *The operations $\overline{Q}^j$ act on $\tilde{H}^*(M^+) \otimes H_*(\mathfrak{g})$ by*

$$\overline{Q}^j(y \otimes x) = \sum_i \text{Sq}^{i-j}(y) \otimes \overline{Q}^i(x).$$

Proof Applying the Cartan formula $Q^j(y \otimes x) = \sum_i Q^{j-i}(y) \otimes Q^i(x)$ for the extended Dyer–Lashof operations $Q^j: x \mapsto e_{j-|x|} \otimes x \otimes x$ and the identification $Q^{-i} = \text{Sq}^i$ [46] to the definition of the $\overline{Q}^j$ operations, we have

$$\overline{Q}^j(y \otimes x) = \xi_* \sigma^{-1} \left(\sum_i \text{Sq}^{i-j}(y) \otimes Q^i(x) \right) = \sum_i \text{Sq}^{i-j}(y) \otimes \xi_* \sigma^{-1} Q^i(x) = \sum_i \text{Sq}^{i-j}(y) \otimes \overline{Q}^i(x).$$

Here σ^{-1} is the desuspension isomorphism, and ξ is the second structure map of spectral Lie algebras. $\square$

4.1 The universal case

Now we apply Theorem 3.28 to the case where M is the Euclidean space. While the homology for $B_k(\mathbb{R}^n; X)$ is well-understood [18; 24; 47], we observe interesting patterns of higher differentials in the associated Knudsen spectral sequence. Furthermore, the computation of the E^2 -page in these cases will be useful in deducing the E^2 -page for a general M .

Since $\tilde{H}^*(S^n) = \mathbb{F}_2\{\iota_n\}$ is concentrated in one dimension, the only nonzero Steenrod operation is $\text{Sq}^0 = \text{id}$, so the $\overline{\mathcal{R}}$ -module structure on $\tilde{H}^*(S^n) \otimes H_*(\mathfrak{g})$ is given by

$$\overline{Q}^j(\iota_n \otimes x) = \sigma^{-n} \overline{Q}^j(x) = \overline{Q}^j(\sigma^{-n}x), \quad x \in \mathfrak{g}.$$

In the limiting case $M = \mathbb{R}^\infty = \lim_{n \rightarrow \infty} \mathbb{R}^n$, we have the stabilization

$$\lim_{n \rightarrow \infty} \Omega^n \text{Free}^{s\mathcal{L}}(\Sigma^n X) \simeq X,$$

and the spectral sequence (10) becomes

$$(11) \quad E_{s,t}^2 = \text{HQ}_{s,t}^{\text{Lie}^s_{\mathcal{R}}}(\Sigma^k \mathbb{F}_2) \Rightarrow H_{s+t}(\text{Free}^{\mathbb{E}^\infty}(\mathbb{S}^k)).$$

The E^2 -page is computed in Theorem 3.28. Namely, it is the exterior algebra generated by one class x_k and two types of operations on coalgebras over the comonad $\pi_{*,*} \text{Bar}_\bullet(\text{id}, \text{Lie}^s_{\mathcal{R}}, -)$,

$$\begin{aligned} \bar{Q}^j : E_{h,s,t}^2 &\rightarrow E_{h,s+1,t+j-1}^2, & j \geq t, \\ \gamma_i : E_{h,s,t}^2 &\rightarrow E_{2h+1,s+i,2t-1}^2, & 2 \leq i \leq s, \end{aligned}$$

under a further splitting of the filtration degree into a sum of homological degree h counting the number of brackets and simplicial degree s counting the number of $\bar{Q}^j$.

Comparing with the computation of $H_*(\text{Free}^{\mathbb{E}^\infty}(\mathbb{S}^k))$ [18; 47], which is the E^∞ -page, we can immediately deduce that the E^2 -page is much larger. Using sparsity arguments, we can identify higher differentials in low degrees, which allows us to make the following conjecture.

Conjecture 4.5 *Each page of the spectral sequence*

$$E_{s,t}^2 = \text{HQ}_{s,t}^{\text{Lie}^s_{\mathcal{R}}}(\Sigma^k \mathbb{F}_2) \Rightarrow \pi_{s+t} \text{Bar}_\bullet(\text{id}, s\mathcal{L}, \Sigma^k \mathbb{F}_2) \cong H_{s+t}(\text{Free}^{\mathbb{E}^\infty}(\mathbb{S}^k))$$

is an exterior algebra. The higher differentials on the exterior generators of the E^2 -page are given as follows:

(1) For an exterior generator $\alpha = \bar{Q}_{j_1} \dots \bar{Q}_{j_m}(x_k)$ on the E^2 -page, we have

$$d^{r+1} \gamma_{r+1}(\alpha) = \bar{Q}_r(\alpha)$$

for $r < m$ and $r \leq j_1 + 1$.

(2) For an exterior generator $\beta = \gamma_{n+1} \bar{Q}_{j_1} \dots \bar{Q}_{j_{m-1}}(x_k)$ on the E^2 -page, we have

$$(a) \quad d^{n+1}(\beta) = \bar{Q}_n \bar{Q}_{j_1} \dots \bar{Q}_{j_{m-1}}(x_k),$$

$$(b) \quad d^{n+1} \gamma_{m+n}(\beta) = d^{n+1}(\beta) \otimes \beta,$$

$$(c) \quad \gamma_l d^{n+1}(\beta) = d^{2n+1} \gamma_{n+l-1}(\beta) \text{ for } n+3 \leq l \leq m.$$

These generate all higher differentials under further applications of γ_i operations in accordance with (2)(b) and (2)(c), as well as the exterior product.

Figure 1 is an illustration of the higher differentials in homological Adams grading $(s+t, s)$ for $\beta = \gamma_{n+1} \bar{Q}_{j_1} \dots \bar{Q}_{j_m}(x_k)$ and $\alpha = \bar{Q}_n \bar{Q}_{j_1} \dots \bar{Q}_{j_{m-1}}(x_k)$ with internal degree b . Set $a = 2b + m + 1$. Along the horizontal line $s = m + 1$ we have generators $\bar{Q}_1(\alpha), \dots, \bar{Q}_{n+1}(\alpha)$, each receiving a solid differential via Conjecture 4.5(1). Along the top slope we have, for each l with $n+2 < l \leq m$, a dashed arrow $d_{2n+1}(\gamma_{n+l-1}(\beta)) = \gamma_l(\alpha)$, which correspond to the differentials in Conjecture 4.5(2)(c). Finally we have a dotted arrow hitting the cross term, corresponding to Conjecture 4.5(2)(b).

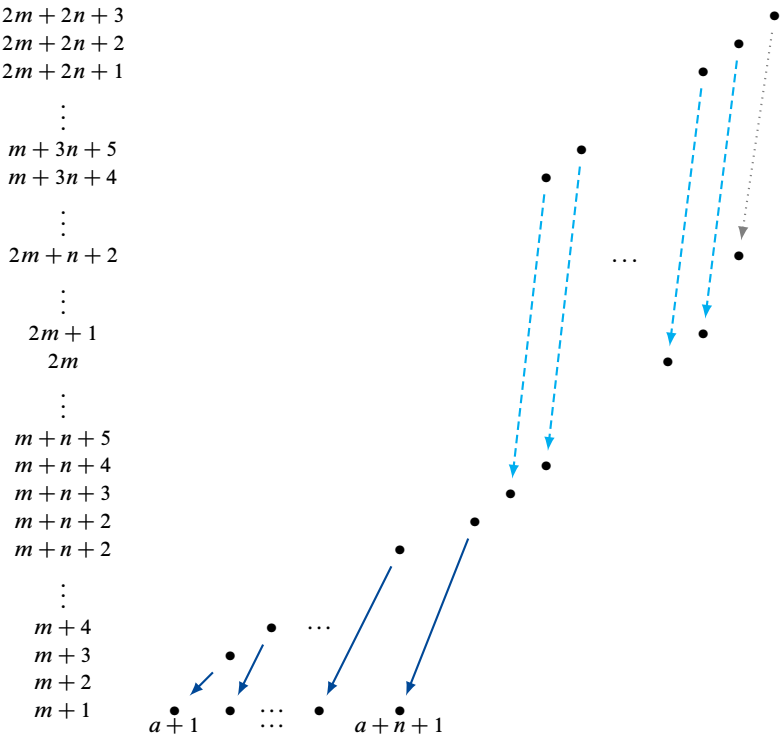


Figure 1: Conjecture 4.5 higher differentials.

Remark 4.6 The pattern in the universal case is similar to the pattern of universal higher differentials in [26, Proposition 2.6; 56], where divided squares kills off Steenrod operations that are not admissible. Here, the Dyer–Lashof operations $\bar{Q}^j$ on the E^∞ -page should be represented by the surviving $\bar{Q}^j$ operations. On the E^2 -page, the admissibility condition for $\bar{Q}^j$ allows for more admissible sequences than the Dyer–Lashof algebra. The γ_i operations eliminate the $\bar{Q}^j$ operations that do not satisfy the admissibility condition for Dyer–Lashof operations via higher differentials.

One major difference is that while Steenrod operations can be defined on the spectral sequence filtrationwise in [26; 56], the operations $\bar{Q}^j$ increase filtration by one. Hence the classical methods of producing operations on spectral sequences by chain-level constructions no longer apply.

In forthcoming work with Robert Burklund and Andrew Senger, we use a suitable deformation of the comonad associated to the bar construction $|\mathrm{Bar}_\bullet(\mathrm{id}, s\mathcal{L}, -)|$ to the ∞ -category of Postnikov-connective filtered $\mathbb{F}_2$ -modules, which allows us to detect the pattern of higher differentials in Conjecture 4.5.

Remark 4.7 The spectral sequence we study here is analogous to the bar spectral sequence

$$E_{s,t}^2 = \pi_s \pi_t \mathrm{Bar}_\bullet(\mathrm{id}, \mathbb{E}_\infty^{\mathrm{nu}} \otimes \mathbb{F}_p, \pi_*(A)) \Rightarrow \pi_{s+t} \mathrm{Bar}_\bullet(\mathrm{id}, \mathbb{E}_\infty^{\mathrm{nu}} \otimes \mathbb{F}_p, A)$$

and its dual. The latter was used to identify operations on homotopy groups of spectral partition Lie algebras and mod p TAQ cohomology operations of nonunital $\mathbb{E}_\infty$ - $\mathbb{F}_p$ -algebras in [57], which subsumes

unpublished work of Kříž, Basterra and Mandell. The E^2 -page of this spectral sequence is the André–Quillen homology of Poly_R -algebras, ie graded $\mathbb{F}_2$ -modules equipped with Dyer–Lashof operations and a polynomial product that satisfying the Cartan formula. In contrast to [Conjecture 4.5](#), this spectral sequence collapses on the E^2 -page. Heuristically, the phenomenon here arises from the nonadditivity of the free Lie^s -algebra functor and the order of the factorization $Q_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^s_{\bar{\mathcal{R}}}} = Q_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^s} \circ Q_{\text{Lie}^s}^{\text{Lie}^s_{\bar{\mathcal{R}}}}$, which results in simplicial homotopy operations. Whereas the Dyer–Lashof operations are additive away from the bottom operations on even degree classes, so the factorization $Q_{\text{Mod}\mathbb{F}_p}^{\text{Poly}_R} = Q_{\text{Mod}\mathbb{F}_p}^{\text{Mod}_{R>0}} \circ Q_{\text{Mod}_{R>0}}^{\text{Poly}_R}$ does not introduce simplicial homotopy operations.

4.2 With coefficients

Next, we take up a slightly more complicated case, where $M = \mathbb{R}^n$ with labels in an arbitrary spectrum X . Then $H_*(\text{Free}^{s\mathcal{L}}(\Sigma^n X)^{M^+}) \cong \Omega^n \text{Free}_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^s_{\bar{\mathcal{R}}}}(\Sigma^n H_*(X))$ and the spectral sequence [\(10\)](#) becomes

$$(12) \quad E_{s,t}^2 = \text{HQ}_{s,t}^{\text{Lie}^s_{\bar{\mathcal{R}}}}(\Omega^n \text{Free}_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^s_{\bar{\mathcal{R}}}}(\Sigma^n H_*(X))) \Rightarrow H_{s+t}(\text{Free}^{\mathbb{E}^n}(X)).$$

When $X = \mathbb{S}^k$, the E^2 -page $\text{HQ}_{s,t}^{\text{Lie}^s_{\bar{\mathcal{R}}}}(\Omega^n \text{Free}_{\text{Mod}\mathbb{F}_2}^{\text{Lie}^s_{\bar{\mathcal{R}}}}(\Sigma^{n+k} \mathbb{F}_2))$ is computed in [Theorem 3.28](#).

Write $H_*(X) \cong \bigoplus_{k,l} \mathbb{F}_2\{x_{k,l}\}$, where $\{x_{k,l}\}_l$ is an $\mathbb{F}_2$ -basis of $H_k(X)$ for each k . Then

$$\begin{aligned} \mathfrak{g} &= H_*(\Omega^n \text{Free}^{s\mathcal{L}}(\Sigma^n H_*(X))) \cong \mathbb{F}_2\{\iota_n\} \otimes H_*(\text{Free}^{s\mathcal{L}}(\Sigma^n H_*(X))) \\ &\cong \mathbb{F}_2\{\iota_n\} \otimes \left(\bigoplus_{w \in W} \mathbb{F}_2\{\bar{Q}^J w, J \in \bar{\mathcal{R}}(d(w))\} \right) \end{aligned}$$

by [\[2, Proposition 7.3\]](#). Here $\bar{\mathcal{R}}(n)$ is the quotient of $\bar{\mathcal{R}}$ by the relations

$$\bar{Q}^{j_1} \dots \bar{Q}^{j_k} = 0 \quad \text{if } j_1 < j_2 + \dots + j_k + n,$$

and W is the set of Lyndon words on the set of letters $\{\sigma^n x_{k,l}\}_{k,l}$, which is a basis for the free $\text{Lie}^{s,\text{ti}}$ -algebra on generators $\{\sigma^n x_{k,l}\}_{k,l}$.

We define the *degree* of a word $w \in W$ to be $d(w) = 1 + \sum_{k,l} m_{k,l}(w)(n+k-1)$, where $m_{k,l}(w)$ counts the number of times the letter $\sigma^n x_{k,l}$ appears in w . Set

$$\mathfrak{g}_w = \mathbb{F}_2\{\iota_n\} \otimes \mathbb{F}_2\{\bar{Q}^J w, J \in \bar{\mathcal{R}}(n+|w|)\}.$$

Then $\mathfrak{g} \simeq \bigoplus_{w \in W} \mathfrak{g}_w$ with trivial brackets. Note that this splitting is induced by an equivalence of $s\mathcal{L}$ -algebras in $\mathbb{F}_2$ -module spectra

$$\begin{aligned} (\text{Free}^{s\mathcal{L}}(\Sigma^n X))^{(\mathbb{R}^n)^+} \otimes \mathbb{F}_2 &\simeq \mathbb{D}(S^n) \otimes \text{Free}^{s\mathcal{L}}(\Sigma^n X \otimes \mathbb{F}_2) \simeq \mathbb{D}(S^n) \otimes \text{Free}^{s\mathcal{L}}\left(\bigvee_{x_{k,l}} \Sigma^{n+k} \mathbb{F}_2\right) \\ &\simeq \bigvee_{w \in W} (\text{Free}^{s\mathcal{L}}(\Sigma^{d(w)} \mathbb{F}_2))^{(\mathbb{R}^n)^+}, \end{aligned}$$

where the last step makes use of [Corollary 5.13](#) in [\[3\]](#). The equivalence above would only be that of $\mathbb{F}_2$ -module spectra if we did not kill the brackets by cotensoring with $(\mathbb{R}^n)^+$. Therefore we deduce:

Proposition 4.8 The spectral sequence $E_{s,t}^2 = \mathrm{HQ}_{s,t}^{\mathrm{Lie}^s_{\bar{\mathcal{R}}}}(\Omega^n \mathrm{Free}_{\mathrm{Mod} \mathbb{F}_2}^{\mathrm{Lie}^s_{\bar{\mathcal{R}}}}(\Sigma^n H_*(X))) \Rightarrow H_{s+t}(\mathrm{Free}^{\mathbb{E}_n}(X))$ splits as

$$E_{s,t}^2 \cong \bigoplus_{w \in W} \mathrm{HQ}_{s,t}^{\mathrm{Lie}^s_{\bar{\mathcal{R}}}}(\mathfrak{g}_w) \Rightarrow \bigoplus_{w \in W} \pi_{s+t} \mathrm{Bar}_\bullet(\mathrm{id}, s\mathcal{L}, \Omega^n \mathrm{Free}^{s\mathcal{L}}(\Sigma^n \Sigma^{d(w)-n} \mathbb{F}_2)).$$

Remark 4.9 The canonical map of spectral Lie algebras

$$\Omega^n \mathrm{Free}^{s\mathcal{L}}(\Sigma^n \mathbb{S}^k) \rightarrow \Omega^\infty \mathrm{Free}^{s\mathcal{L}}(\Sigma^\infty \mathbb{S}^k)$$

via stabilization induces an embedding of the E^2 -pages

$$\mathrm{HQ}^{\mathrm{Lie}^s_{\bar{\mathcal{R}}}}(\Omega^n \mathrm{Free}_{\mathrm{Mod} \mathbb{F}_2}^{\mathrm{Mod} \bar{\mathcal{R}}}(\Sigma^{n+k} \mathbb{F}_2)) \rightarrow \mathrm{HQ}^{\mathrm{Lie}^s_{\bar{\mathcal{R}}}}(\Sigma^k \mathbb{F}_2)$$

by [Proposition 2.10](#) and [Theorem 3.28](#). We expect that the higher differentials in the target ([Conjecture 4.5](#)) pull back to higher differentials in the source. Indeed, combinatorially this will yield the computation of the free $\mathbb{E}_n$ -algebra on a single generator. If $H_*(X)$ has multiple generators, then the splitting of the spectral sequence above via Lyndon words corresponds precisely the Browder bracket on the free $\mathbb{E}_n$ -algebra on those generators; see [\[24, III\]](#).

5 Upper bounds and low weight computations

For a general parallelizable manifold M of dimension n , the $\mathrm{Lie}^s_{\bar{\mathcal{R}}}$ -algebra

$$\mathfrak{g} = \tilde{H}^*(M^+) \otimes \mathrm{Free}_{\mathrm{Mod} \mathbb{F}_2}^{\mathrm{Lie}^s_{\bar{\mathcal{R}}}}(\Sigma^n H_*(X))$$

has nontrivial Lie^s -brackets and the precise image of the comparison map φ_* in [Lemma 3.22](#) becomes much harder to pin down. Nonetheless, [Theorem 3.6](#) and [Corollary 3.10](#) allow us to obtain a formula for an upper bound of $\pi_{*,*} \mathrm{Bar}_\bullet(\mathrm{id}, \mathrm{Lie}^s_{\bar{\mathcal{R}}}, \mathfrak{g})$ by

$$\pi_{*,*} \mathrm{Bar}_\bullet(\mathrm{id}, \mathrm{Lie}^{s,\mathrm{ti}}_{\bar{\mathcal{R}}}, \tilde{\mathfrak{g}}) \cong \pi_{*,*}(\mathrm{CE}(\mathrm{AR}_\bullet(\tilde{\mathfrak{g}})))$$

that is an equivalence in weight less than four. Here $\tilde{\mathfrak{g}} = \tilde{H}^*(M^+) \otimes \mathrm{Free}^{\mathrm{Lie}^{s,\mathrm{ti}}_{\bar{\mathcal{R}}}}(H_*(X))$ is the associated $\mathrm{Lie}^{s,\mathrm{ti}}_{\bar{\mathcal{R}}}$ -algebra, where $\tilde{H}^*(M^+)$ is equipped with the $\mathrm{Lie}^{s,\mathrm{ti}}$ -bracket coming from the associated $\mathrm{Lie}^{s,\mathrm{ti}}$ -algebra of the Lie^s -algebra $H^*(M^+)$ with its usual cup product; see [Construction 3.4](#). In particular, it follows from [Corollary 3.7](#) that in weight less than four, the two homotopy groups are isomorphic.

5.1 General upper bounds

We will see that $\pi_{*,*}(\mathrm{CE}(\mathrm{AR}_\bullet(\tilde{\mathfrak{g}})))$ admits a description in terms of the $\mathrm{Lie}^{s,\mathrm{ti}}$ -algebra homology of $\tilde{\mathfrak{g}}$. The key observation is that for $\tilde{\mathfrak{g}} = \tilde{H}^*(M^+) \otimes \mathrm{Free}^{\mathrm{Lie}^{s,\mathrm{ti}}_{\bar{\mathcal{R}}}}(H_*(X))$, $\mathrm{AR}_\bullet(\tilde{\mathfrak{g}})$ has trivial $\mathrm{Lie}^{s,\mathrm{ti}}$ -structure away from simplicial degree 0 and its degeneracies, see [Construction 3.9](#), and the $\mathrm{Lie}^{s,\mathrm{ti}}$ -bracket on $\tilde{\mathfrak{g}}$ vanishes on elements that involve $\bar{Q}^i$ operations.

Definition 5.1 For a $\text{Lie}^{s,\text{ti}}$ -algebra $\mathfrak{g}$, we say that its $\text{Lie}^{s,\text{ti}}$ -structure is *supported entirely* by a sub- $\text{Lie}^{s,\text{ti}}$ -algebra $\mathfrak{g}'$ if the $\text{Lie}^{s,\text{ti}}$ -algebra $\mathfrak{g}$ is isomorphic to the product $\text{Lie}^{s,\text{ti}}$ -algebra $N \oplus \mathfrak{g}'$, where the $\text{Lie}^{s,\text{ti}}$ bracket vanishes on the complement $N \subset \mathfrak{g}$.

Lemma 5.2 Let $\tilde{\mathfrak{g}} = L \otimes \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^{s,\text{ti}}_{\bar{\mathcal{R}}}}(V)$ be a $\text{Lie}^{s,\text{ti}}_{\bar{\mathcal{R}}}$ -algebra, where L is a nonunital graded commutative algebra over $\mathbb{F}_2$ and the $\text{Lie}^{s,\text{ti}}$ -structure on $\tilde{\mathfrak{g}}$ is the usual one on the tensor product. Then

$$\pi_{*,*}(\text{CE}(\text{AR}_{\bullet}(\tilde{\mathfrak{g}}))) \cong \Lambda\{\gamma_I(\alpha), \alpha \in A\} \otimes H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}}),$$

where $\alpha \in A$ is an element of an $\mathbb{F}_2$ -basis for $\pi_{\geq 1,*}(\text{AR}_{\bullet}(\tilde{\mathfrak{g}}))$ with simplicial degree $s(\alpha)$, and I is γ -admissible with $e(I) \leq s(\alpha)$.

Proof Since brackets of operations are zero, the $\text{Lie}^{s,\text{ti}}$ -algebra $\tilde{\mathfrak{g}}$ is supported entirely by the sub- $\text{Lie}^{s,\text{ti}}$ -algebra $\mathfrak{g}'_0 = L \otimes \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}^{s,\text{ti}}}(V)$. Furthermore, for all $m \geq 1$, the $\text{Lie}^{s,\text{ti}}$ -algebra $\text{AR}_m(\tilde{\mathfrak{g}})$ is supported entirely by the degeneracies coming from $\mathfrak{g}'_0$ by [Construction 3.9](#). Hence each simplicial level $\text{AR}_m(\tilde{\mathfrak{g}})$ is isomorphic to the product $\text{Lie}^{s,\text{ti}}$ -algebra $T_m \oplus \mathfrak{g}'_m$, where $\mathfrak{g}'_m$ is the sub- $\text{Lie}^{s,\text{ti}}_{\mathbb{F}_2}$ -algebra consisting of degeneracies of $\mathfrak{g}'_0$ and T_m a trivial $\text{Lie}^{s,\text{ti}}$ -algebra. Since the splittings respect the simplicial $\text{Lie}^{s,\text{ti}}$ -algebra structure of $\text{AR}_{\bullet}(\tilde{\mathfrak{g}})$, we deduce that $\text{AR}_{\bullet}(\tilde{\mathfrak{g}}) \cong T_{\bullet} \oplus \mathfrak{g}'_{\bullet}$ as simplicial $\text{Lie}^{s,\text{ti}}_{\mathbb{F}_2}$ -algebras. This induces a splitting of chain complexes

$$\text{CE}(\text{AR}_{\bullet}(\tilde{\mathfrak{g}})) \cong \text{CE}(T_{\bullet}) \otimes \text{CE}(\mathfrak{g}'_{\bullet}),$$

where $T_{\bullet}$ is a trivial simplicial $\text{Lie}^{s,\text{ti}}$ -algebra and $\mathfrak{g}'_{\bullet}$ the constant simplicial object on $\mathfrak{g}'_0$. The lemma then follows from [Theorem 3.15](#), noting that $H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}}) \cong H_{*,*}^{\text{Lie}^{s,\text{ti}}}(T_0) \otimes H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\mathfrak{g}'_0)$. $\square$

It remains to compute $\pi_{*,*}(\text{AR}_{\bullet}(\tilde{\mathfrak{g}}))$ for $\tilde{\mathfrak{g}} = \tilde{H}^*(M^+) \otimes \text{Free}_{\bar{\mathcal{R}}}^{\text{Lie}^{s,\text{ti}}}(H_*(X))$. Since $\mathfrak{g}$ and $\tilde{\mathfrak{g}}$ are isomorphic as $\bar{\mathcal{R}}$ -modules (see [Remark 3.2](#)), we will not distinguish the two. Recall from [Proposition 4.4](#) that the $\bar{\mathcal{R}}$ -module structure on $\mathfrak{g}$ is twisted by the Steenrod operations in the sense that

$$\bar{Q}^j(y \otimes \alpha) = \sum_{0 \leq s \leq n} \text{Sq}^{j+s}(y) \otimes \bar{Q}^s(\alpha).$$

Notation 5.3 Let $H \cup \{z\}$ be an $\mathbb{F}_2$ -basis of the cohomology ring $H^*(M^+)$, where z corresponds to the added point in the one-point compactification and H is a basis for $\tilde{H}^*(M^+)$. For $y \in H$, denote by $|y|$ the cohomological degree of y .

Let $\tilde{B} = \{x_a\}_a$ be a totally ordered basis for $V = H_*(X)$ and $B = \{\sigma^n x_a\}_a$ with the induced ordering, where n is the dimension of M . Denote by W the set of basic products on the set B . Then

$$\mathfrak{g} = \tilde{H}^*(M^+) \otimes H_*(\text{Free}^{s,\mathcal{L}}(\Sigma^n X)) \cong \bigoplus_{w \in W, y \in H} \mathbb{F}_2\{y\} \otimes \mathbb{F}_2\{\bar{Q}^J w, J \in \bar{\mathcal{R}}(|w|)\}.$$

Proposition 5.4 The bigraded homotopy group

$$\pi_{*,*}(\text{AR}_{\bullet}(\tilde{\mathfrak{g}})) = \pi_{*,*}(\text{AR}_{\bullet}(\mathfrak{g}))$$

is isomorphic to $\pi_{*,*}(\text{AR}_{\bullet}(\mathfrak{g}_{\text{triv}}))$, where the untwisted $\bar{\mathcal{R}}$ -module $\mathfrak{g}_{\text{triv}}$ has the same underlying $\mathbb{F}_2$ -module as $\mathfrak{g}$ and the $\bar{\mathcal{R}}$ -module structure is given by $\bar{Q}^j(y \otimes x) = y \otimes \bar{Q}^j(x)$ for all j .

Proof We make use of a spectral sequence to filter away the twisting by the action of the Steenrod operations. We abuse notation here and denote again by $\mathrm{AR}_\bullet(\mathfrak{g})$ the associated chain complex of $\mathrm{AR}_\bullet(\mathfrak{g})$. Filter $\mathfrak{g}$ in terms of decreasing cohomological degree of $\tilde{H}^*(M^+)$, so we have

$$F_{-p}(\mathfrak{g}) = \tilde{H}^{\geq p}(M^+) \otimes \mathrm{Free}_{\mathrm{Mod}\mathbb{F}_2}^{\mathrm{Lie}_{\bar{\mathcal{R}}}^s}(V) \cong \bigoplus_{w \in W, y \in H, |y| \geq p} \mathbb{F}_2\{y \otimes \bar{Q}^J(w), J \in \bar{\mathcal{R}}(|w|)\}$$

with associated graded pieces given by

$$G_{-p}(\mathfrak{g}) = F_{-p}(\mathfrak{g})/F_{-p-1}(\mathfrak{g}) \cong \bigoplus_{w \in W, y \in H, |y|=p} \mathbb{F}_2\{y \otimes \bar{Q}^J(w), J \in \bar{\mathcal{R}}(|w|)\}.$$

Since action by Steenrod operations does not decrease cohomological degree, the induced filtration

$$F_{-p}(\mathrm{AR}_\bullet(\mathfrak{g})) := \mathrm{AR}_\bullet(F_{-p}(\mathfrak{g}))$$

makes $\mathrm{AR}_\bullet(\mathfrak{g})$ a filtered chain complex. The associated graded pieces are

$$G_{-p}(\mathrm{AR}_\bullet(\mathfrak{g})) = \mathrm{AR}_\bullet(G_{-p}(\mathfrak{g})) = \bigoplus_{w \in W, y \in H, |y|=p} \mathrm{AR}_\bullet(\mathbb{F}_2\{y \otimes \bar{Q}^J(w), J \in \bar{\mathcal{R}}(|w|)\})$$

and the induced differential preserves direct summands.

Using the case $M = \mathbb{R}^n$ in [Proposition 4.8](#), we deduce that

$$\begin{aligned} E_{-p,q}^1 &= H_{-p+q}(G_p(\mathrm{AR}_\bullet(\mathfrak{g}))) \cong \bigoplus_{w \in W, y \in H, |y|=p} \pi_* (\mathrm{AR}_\bullet(\mathbb{F}_2\{y \otimes \bar{Q}^J(w), J \in \bar{\mathcal{R}}(|w|)\})) \\ &\cong \bigoplus_{w \in W, y \in H, |y|=p} \mathbb{F}_2\{\bar{Q}^{j_1} \dots \bar{Q}^{j_m}(y \otimes w), (j_1, \dots, j_m) \in \bar{\mathcal{R}}(p, |w|)\}, \end{aligned}$$

where $\bar{\mathcal{R}}(p, |w|)$ is the set of sequences $(j_1, \dots, j_m)$ such that:

- (1) $j_l \leq 2j_{l+1}$ for $1 \leq l < m$ and $|w| - p \leq j_m < |w|$.
- (2) If $m \geq 2$ then $j_l \geq j_{l+1} + \dots + j_m + |w| - (m - l)$ for

$$2 \leq l \leq m-1 \quad \text{and} \quad j_1 > j_2 + \dots + j_m + |w| - (m-1).$$

We claim that every class on the E^1 -page survives to a class on the E^∞ -page by induction on $\tilde{H}^*(M^+)$ along decreasing cohomological degree.

For $y \in \tilde{H}^n(M^+) \in F_{-n}(\mathfrak{g})$ a top cohomology class, there is no nonzero action by a Steenrod operation on y other than Sq^0 , so the differential on β in $\mathrm{AR}_\bullet(\mathfrak{g})$ is the same as the differential in $G_{-n}(\mathrm{AR}_\bullet(\mathfrak{g}))$, ie β survives to a nonzero class on the E^∞ -page.

Suppose that in $F_{-p-1}(\mathrm{AR}_\bullet(\mathfrak{g})) = \mathrm{AR}_\bullet(F_{-p-1}(\mathfrak{g}))$, any basis element $\beta' = \bar{Q}^{j'_1} \dots \bar{Q}^{j'_m}(y' \otimes w')$ of the E^1 -page is a permanent cycle and they span all permanent cycles in $F_{-p-1}(\mathrm{AR}_\bullet(\mathfrak{g}))$. Let $[\beta] = \bar{Q}^{j_1} \dots \bar{Q}^{j_m}(y \otimes w)$ be a basis element on the E^1 -page, with $y \in \tilde{H}^p(M^+)$. A cycle representing this class in $\mathrm{AR}_\bullet(G_{-p}(\mathfrak{g}))$ is a finite sum

$$\beta = \bar{Q}^{j_1} | \dots | \bar{Q}^{j_m} | (y \otimes w) + \sum_l \bar{Q}^{l_1} | \dots | \bar{Q}^{l_m} | (y \otimes w)$$

obtained by cycle completion via Behrens' relations in the sense of [Remark 3.18](#). Note that $l_m \leq j_m < |w|$ for all l . Let d_m be the rightmost face map. Then in $\mathrm{AR}_\bullet(\mathfrak{g})$

$$\begin{aligned} \partial\beta &= \partial\left(\bar{Q}^{j_1}|\dots|\bar{Q}^{j_m}|(y \otimes w) + \sum_l \bar{Q}^{l_1}|\dots|\bar{Q}^{l_m}|(y \otimes w)\right) \\ &= 0 + d_m\left(\bar{Q}^{j_1}|\dots|\bar{Q}^{j_m}|(y \otimes w) + \sum_l \bar{Q}^{l_1}|\dots|\bar{Q}^{l_m}|(y \otimes w)\right) \\ &= \sum_{s \geq 0} \bar{Q}^{j_1}|\dots|\bar{Q}^{j_{m-1}}|\mathrm{Sq}^s(y) \otimes \bar{Q}^{j_m+s}(w) + \sum_l \sum_{s \geq 0} \bar{Q}^{l_1}|\dots|\bar{Q}^{l_{m-1}}|\mathrm{Sq}^s(y) \otimes \bar{Q}^{l_m+s}(w). \end{aligned}$$

Note that the sum of these

$$\theta_l = \bar{Q}^{l_1}|\dots|\bar{Q}^{l_{m-1}}|\mathrm{Sq}^s(y) \otimes \bar{Q}^{l_m+s}(w) \quad \text{or} \quad \bar{Q}^{j_1}|\dots|\bar{Q}^{j_{m-1}}|\mathrm{Sq}^s(y) \otimes \bar{Q}^{j_m+s}(w)$$

over $s \geq 0$ is a boundary in $\mathrm{AR}_\bullet(\mathfrak{g})$: If $l_m + s < |w|$, then $\theta_l = 0$. If $l_m + s \geq |w|$ or $j_m + s \geq |w|$, then $s \geq 1$, since $l_m \leq j_m < |w|$, so $\theta_l \in F_{-p-1}(\mathrm{AR}_\bullet(\mathfrak{g}))$. By the inductive hypothesis, the sum of such θ_l is not a nonzero cycle on the E^∞ -page and thus the boundary of a finite sum of classes in $F_{-p-1}(\mathrm{AR}_\bullet(\mathfrak{g}))$ of the form $\bar{Q}^{j'_1}|\dots|\bar{Q}^{j'_m}|(y' \otimes w')$ with $|y'| \geq p + s > p$. Denote by ξ this finite sum, so $\partial(\beta + \xi) = 0$ in $\mathrm{AR}_\bullet(\mathfrak{g})$. Note that ξ is not a boundary because it is maximally nondegenerate and $\xi \neq \beta$ since β is not in $F_{-p-1}(\mathrm{AR}_\bullet(\mathfrak{g}))$. Hence $\beta + \xi$ is a cycle in $\mathrm{AR}_\bullet(\mathfrak{g})$ corresponding to the basis element $\beta = \bar{Q}^{j_1}|\dots|\bar{Q}^{j_m}|(y \otimes w)$ on the E^1 -page. Therefore the dimension of the E^1 -page is at most that of the E^∞ -page, so no differential can happen in the spectral sequence. $\square$

Combing [Lemma 5.2](#), [Proposition 5.4](#) and [Corollary 3.7](#), we deduce the following general upper bound and low weight computation of the E^2 -page of the Knudsen spectral sequence.

Theorem 5.5 *Let M be a parallelizable manifold of dimension n and X any spectrum. Let $\mathfrak{g}$ denote the $\mathrm{Lie}_{\mathcal{R}}^s$ -algebra $\tilde{H}^*(M^+) \otimes \mathrm{Free}_{\mathrm{Mod}_{\mathbb{F}_2}}^{\mathrm{Lie}_{\mathcal{R}}^s}(\Sigma^n H_*(X))$ with $\mathbb{F}_2$ -basis B , and $\tilde{\mathfrak{g}}$ the associated $\mathrm{Lie}_{\mathcal{R}}^{s, \mathrm{ti}}$ -algebra. An upper bound for the E^2 -page of the weighted spectral sequence*

$$(13) \quad E_{s,t}^2 = \mathrm{HQ}_{s,t}^{\mathrm{Lie}_{\mathcal{R}}^s}(\mathfrak{g}) \Rightarrow \bigoplus_{k \geq 1} H_{s+t}(B_k(M; X))$$

is given by

$$\pi_{*,*}(\mathrm{CE}(\mathrm{AR}_\bullet(\tilde{\mathfrak{g}}))) \cong \Lambda\{\gamma_I \bar{Q}_J(y \otimes w), y \otimes w \in H \otimes B\} \otimes H_{*,*}^{\mathrm{Lie}_{\mathcal{R}}^{s, \mathrm{ti}}}(\tilde{\mathfrak{g}}),$$

where $\gamma_I \bar{Q}_J(y \otimes w)$ satisfies the conditions that

- (1) $J = (j_1, \dots, j_m)$ with $m \geq 1$, $0 \leq j_l \leq j_{l+1} + 1$ for $1 \leq l < m$, and $0 \leq j_m < |y|$,
- (2) I is γ -admissible with $e(I) \leq m$.

Furthermore, in weight less than four equality is achieved.

5.2 Low weight computations

Theorem 5.5 allows us to deduce the degeneration of the spectral sequence at weight two and three using sparsity arguments. Denote by $\text{wt}_k(M)$ the weight k part of a weighted (bi)graded $\mathbb{F}_2$ -module M and set $E^r(k) = \text{wt}_k(E^r)$.

Corollary 5.6 Let $\mathfrak{g}, \tilde{\mathfrak{g}}$ be the same as in [Theorem 5.5](#) and B, H bases given in [Notation 5.3](#). The weight two part of the spectral sequence (13)

$$E_{s,t}^2(2) = \text{wt}_2(\text{HQ}_{s,t}^{\text{Lie}^s_{\mathcal{R}}}(\mathfrak{g})) \Rightarrow H_{s+t}(B_2(M; X))$$

collapses on the E^2 -page, and hence

$$E^\infty(2) \cong E^2(2) \cong \text{wt}_2(H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}})) \oplus \bigoplus_{x \in B, y \in H} \{\bar{Q}_j(y \otimes x), 0 \leq j < |y|\}.$$

Proof Since classes in the tensor factor

$$A = \Lambda\{\gamma_I(\bar{Q}_J(y \otimes w)), y \otimes w \in H \otimes B\}$$

of [Theorem 5.5](#) have weight at least two, classes of weight two lie in exactly one of the two tensor components A and $H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}})$. The weight two classes in A are of the form $\bar{Q}_j(y \otimes w)$, where w has weight one, ie w is an element of the $\mathbb{F}_2$ -basis B of $V = H_*(X)$; see [Notation 5.3](#). The weight two classes in $H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}})$ are of the form $y \otimes \langle x_a, x_b \rangle$ and $(y \otimes x_a) \otimes (y' \otimes x_b)$. Hence the weight two part of the spectral sequence has E^2 -page concentrated in simplicial degrees 0, 1 and thus cannot admit higher differentials. $\square$

In particular, this demonstrates that for a parallelizable M , the $\mathbb{F}_2$ -module $H_*(B_2(M; X))$ depends on and only on the cohomology ring $H^*(M^+)$ when $H_*(X)$ has at least two generators.

Remark 5.7 This is in contrast to the case where $X = \mathbb{S}^r$ has only one generator in its homology: Bödiger, Cohen and Taylor [\[12\]](#) showed that for any n -manifold M ,

$$\bigoplus_{k \geq 1} H_*(B_k(M; \mathbb{S}^r)) \cong \bigotimes_{i=0}^n H_*(\Omega^{n-i} S^{n+r})^{\otimes \dim H_i(M)}$$

depends only on $H^*(M)$ as an $\mathbb{F}_2$ -module.

There is a clear bijection from the weight 2 part of their decomposition to the basis above: Let x_k denote the generator of the free $\mathbb{E}_n$ -algebra $H_*(\Omega^n \Sigma^n S^k)$. For y a basis element of $\tilde{H}^i(M^+) \cong H_i(M)$, the bijection sends $\bar{Q}_j(y \otimes x_{n+r})$ to the tensor with $Q_j(x_{r+i})$ in the tensor factor $H_*(\Omega^{n-i} S^{n+r})$ corresponding to y and 1 in all other tensor factors. The $\text{Lie}^{s,\text{ti}}$ -algebra $\text{Lie}^{s,\text{ti}} \mathfrak{g}$ is trivial, so $\text{wt}_2(H^{\text{Lie}^{s,\text{ti}}}(\text{Lie}^{s,\text{ti}} \mathfrak{g})) \cong \{yy'\}$ where y, y' ranges over distinct basis of $\tilde{H}^i(M^+)$ and the bijection sends yy' to the tensor with factors y, y' and 1 in all other slots.

On the other hand, the homology of $\text{Conf}_2(M)$, the space of ordered configurations of two points in M , also depends only on the cup product structure of $H^*(M)$ as discussed in [\[50, Section 1.1\]](#).

Corollary 5.8 *If in addition M is a closed manifold, then the weight three part of the spectral sequence (13) collapses on the E^2 -page, and a basis for $H_*(B_3(M; X))$ is given by*

$$E^\infty(3) \cong E^2(3) \cong \bigoplus_{x, x' \in B, y, y' \in H} \mathbb{F}_2\{(\bar{Q}_j(y \otimes x)) \otimes (y' \otimes x'), 0 \leq j < |y|\} \oplus \text{wt}_3(H_{*,*}^{\text{Lie}^{s, \text{ti}}}(\tilde{\mathfrak{g}})),$$

Proof Let d denote the generator for $\tilde{H}^0(M^+) \cong H^0(M)$. Then any element that is a sum of $y \otimes \langle \langle x_1, x_2 \rangle, x_3 \rangle \in H \otimes B$ is killed by a sum of $(y \otimes \langle x_1, x_2 \rangle) \otimes (d \otimes x_3)$. Since classes in A have weights positive powers of two, weight three classes on the E^2 -page either live in $\text{wt}_3(H_{*,*}^{\text{Lie}^{s, \text{ti}}}(\tilde{\mathfrak{g}}))$ with simplicial degree one or two, or have the form

$$(\bar{Q}^j(y \otimes x)) \otimes (y' \otimes x') \in \text{wt}_2(A) \otimes \text{wt}_1(H_{*,*}^{\text{Lie}^{s, \text{ti}}}(\tilde{\mathfrak{g}}))$$

with simplicial degree two. Hence $E^2(3)$ is concentrated in simplicial degree 1 and 2, so there cannot be any higher differentials. $\square$

At the weight four part we can no longer deduce that the spectral sequence (13) collapses on the E^2 -page using sparsity arguments. An upper bound for the bigraded $\mathbb{F}_2$ -module $E^2(4)$ is given by the weight four part of $A \otimes H^{\text{Lie}^{s, \text{ti}}}(\tilde{\mathfrak{g}})$, which consists of

- (1) $\bar{Q}_i(y \otimes \langle x, x' \rangle)$ in simplicial degree one,
- (2) $\bar{Q}_i \bar{Q}_j(y \otimes x)$ and $\bar{Q}_i(y \otimes x) \otimes (y' \otimes \langle x_1, x_2 \rangle)$ in simplicial degree two,
- (3) $\bar{Q}_i(y \otimes x) \otimes \bar{Q}_j(y' \otimes x')$ and $\bar{Q}_i(y \otimes x) \otimes (y_1 \otimes x_1) \otimes (y_2 \otimes x_2)$ in simplicial degree three,
- (4) the weight four part of $H^{\text{Lie}^{s, \text{ti}}}(\tilde{\mathfrak{g}})$.

There could well be a d^2 -differential from degree considerations.

We close this section by a few example computations: the closed torus, the punctured genus g surfaces with $g \geq 1$ and the (punctured) real projective space $\mathbb{RP}^3$.

5.3 Example computations: closed torus and punctured genus g surfaces

Let $\Sigma_{g,1}$ be a once-punctured surface of genus $g \geq 1$ and Σ_1 the closed torus. Let $\tilde{B} = \{x_i\}_i$ be a totally ordered basis for $H_*(X)$ and $B = \{\sigma^2 x_i\}_i$ with the induced ordering. Then

$$\tilde{H}^*(\Sigma_{g,1}^+) = \begin{cases} \mathbb{F}_2\{a_i \oplus b_i, i = 1, \dots, g\} & \text{if } * = 1, \\ \mathbb{F}_2\{c\} & \text{if } * = 2, \\ 0 & \text{otherwise,} \end{cases}$$

with nonzero cup products $a_i \cup b_i = c$ for all i and no nontrivial Steenrod operations.

For the closed surface Σ_1 , we have

$$\tilde{H}^*(\Sigma_1^+) \cong H^*(\Sigma_1) = \begin{cases} \mathbb{F}_2\{d\} & \text{if } * = 0, \\ \mathbb{F}_2\{a, b\} & \text{if } * = 1, \\ \mathbb{F}_2\{c\} & \text{if } * = 2, \\ 0 & \text{otherwise,} \end{cases}$$

with nonzero cup products $a \cup b = c$ and $d \cup y = y$ for all $y \in H^*(\Sigma_1)$.

5.3.1 Weight two For $M = \Sigma_{g,1}$, the weight two classes supporting nonzero CE differentials are $\delta(a_i \otimes x_1, b_i \otimes x_2) = c \otimes \langle x_1, x_2 \rangle$ for $x_1 \neq x_2 \in B$, since these are the only nonzero cup products not involving the unit. Denote by H^1 the set of generators $\{a_i, b_i, i = 1, \dots, g\}$ for $\tilde{H}^1(\Sigma_{g,1}^+)$. Impose a total ordering on $H^1 \cup \{c, d\}$. By [Corollary 5.6](#), a basis for $H_*(B_2(\Sigma_{g,1}; X))$ is given by

$$\begin{aligned} E^\infty(2) &= E^2(2) \\ &\cong \bigoplus_{x \in B} \mathbb{F}_2 \{ \bar{Q}_0(y \otimes x), y \in H^1; \bar{Q}_0(c \otimes x), \bar{Q}_1(c \otimes x) \} \\ &\quad \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2 \{ y \otimes \langle x_1, x_2 \rangle, (y \otimes x_1) \otimes (y \otimes x_2), y \in H^1; (c \otimes x_1) \otimes (c \otimes x_2) \} \\ &\quad \oplus \bigoplus_{x_1, x_2 \in B} \mathbb{F}_2 \{ (y \otimes x_1) \otimes (c \otimes x_2), y \in H^1 \} \oplus \bigoplus_{x \in B} \mathbb{F}_2 \{ (y \otimes x) \otimes (y' \otimes x), y < y' \in H^1 \cup \{c\} \} \\ &\quad \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2 \{ (y \otimes x_1) \otimes (y' \otimes x_2) + (a_1 \otimes x_1) \otimes (b_1 \otimes x_2), y \neq y' \in H^1, (y, y') \neq (a_i, b_i) \}. \end{aligned}$$

For $M = \Sigma_1$, the weight two classes supporting CE differentials are

$$\delta((a \otimes x_2) \otimes (b \otimes x_2)) = c \otimes \langle x_1, x_2 \rangle \quad \text{and} \quad \delta((d \otimes x_1) \otimes (y \otimes x_2)) = y \otimes \langle x_1, x_2 \rangle$$

for $x_1 \neq x_2 \in B$ and $y \in \tilde{H}^*(\Sigma_1^+)$. By [Corollary 5.6](#), a basis for $H_*(B_2(\Sigma_1; X))$ is given by

$$\begin{aligned} E^\infty(2) &= E^2(2) \\ &\cong \bigoplus_{x \in B} \mathbb{F}_2 \{ \bar{Q}_0(y \otimes x), y \in H^1; \bar{Q}_0(c \otimes x), \bar{Q}_1(c \otimes x) \} \\ &\quad \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2 \{ (y \otimes x_1) \otimes (y \otimes x_2), y \in H^1; (z \otimes x_1) \otimes (z \otimes x_2) \} \\ &\quad \oplus \bigoplus_{x_1 \neq x_2 \in B} \mathbb{F}_2 \{ (y \otimes x_1) \otimes (z \otimes x_2), y \in H^1 \} \\ &\quad \oplus \bigoplus_{x \in B} \mathbb{F}_2 \{ (y \otimes x) \otimes (y' \otimes x), \{y < y'\} \in \{a, b, c, d\} \} \\ &\quad \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2 \{ (y \otimes x_1) \otimes (y' \otimes x_2), y, y' \in H^1, \{y, y'\} \neq \{a, b\} \} \\ &\quad \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2 \{ (y \otimes x_1) \otimes (y' \otimes x_2) + (d \otimes x_1) \otimes (c \otimes x_2), \{y, y'\} = \{a, b\} \text{ or } (y, y') = (c, d) \}. \end{aligned}$$

Example 5.9 When $X = \mathbb{S}^k$ with $k \geq 1$, we have $B = \{x = \sigma_2 \iota_k\}$, so $H_*(B_2(\Sigma_1, S^k))$ has $\mathbb{F}_2$ -basis

$$\{ \bar{Q}_0(a \otimes x), \bar{Q}_0(b \otimes x), \bar{Q}_0(c \otimes x), \bar{Q}_1(c \otimes x); (y \otimes x) \otimes (y' \otimes x), \{y < y'\} \subset \{a, b, c, d\} \}.$$

The weight two part of the Bökigheimer–Cohen–Taylor decomposition [\[12\]](#),

$$\begin{aligned} (14) \quad \bigoplus_{k \geq 1} H_*(B_k(\Sigma_1; \mathbb{S}^k)) &\cong \bigotimes_{i=0}^n H_*(\Omega^{2-i} S^{2+k})^{\otimes \dim H_i(M)} \\ &\cong H_*(\Omega^2 \Sigma^2 S^k) \otimes H_*(\Omega \Sigma S^{1+k})^{\otimes 2} \otimes H_*(S^{2+k}), \end{aligned}$$

is an $\mathbb{F}_2$ -module on generators

$$Q_0(x_k) \otimes 1 \otimes 1 \otimes 1, \quad Q_1(x_k) \otimes 1 \otimes 1 \otimes 1, \quad 1 \otimes \bar{Q}_0(x_{k+1}) \otimes 1 \otimes 1, \quad 1 \otimes 1 \otimes \bar{Q}_0(x_{k+1}) \otimes 1,$$

as well as 6 other elements, where we let two of the four tensor factors be 1 and the other two be the weight 1 generators. There is a one-to-one correspondence by sending $y \otimes x$ to $x_{k+2-|y|}$ and $\bar{Q}_i(y \otimes x)$ to $Q_i(x_{k+2-|y|})$ for $y = a, b, c, d$.

5.3.2 Weight three Classes in $A = \Lambda(\gamma_I(\bar{Q}^{j_1} | \dots | \bar{Q}^{j_m} | (y \otimes w)), m \geq 1)$ have weights positive powers of 2. Hence weight three classes in $E^2(3)$ either live in $\text{wt}_3(H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}}))$ or has the form

$$(\bar{Q}_j(y \otimes x)) \otimes (y' \otimes x') \in A \otimes H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}}), \quad x, x' \in B.$$

Let H be the set of generators for $\tilde{H}^*(\Sigma_{g,1}^+) \cong \tilde{H}^*(\Sigma_g)$ and H^1 the set of generators for $\tilde{H}^1(\Sigma_{g,1}^+)$. Recall that $\tilde{\mathfrak{g}} = \tilde{H}^*(\Sigma_g) \otimes \text{Free}_{\text{Mod } \mathbb{F}_2}^{\text{Lie}^{s,\text{ti}}_{\tilde{\mathfrak{R}}}(\Sigma^n H_*(X))}(\Sigma^n H_*(X))$. Then we have

$$\begin{aligned} E^2(3) \cong & \bigoplus_{x_1, x_2 \in B} \mathbb{F}_2\{(\bar{Q}_0(y \otimes x_1)) \otimes (y' \otimes x_2), y \in H^1, y' \in H\} \\ & \oplus \bigoplus_{x_1, x_2 \in B} \mathbb{F}_2\{(\bar{Q}_0(c \otimes x_1)) \otimes (y \otimes x_2), (\bar{Q}_1(c \otimes x)) \otimes (y \otimes x_2), y \in H\} \oplus \text{wt}_3(H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}})). \end{aligned}$$

A complete list of an $\mathbb{F}_2$ -basis of $\text{wt}_3(H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}}))$ can be written down in a straight forward way.

The E^2 -page is concentrated in simplicial degree 0, 1, 2. We need to investigate all classes in $E_{2,*}^2(3)$ to see if they support nontrivial d^2 -differentials to $E_{0,*+1}^2(3)$. Note that all classes in $E_{0,*}^2(3)$ are of the form $y \otimes \langle \langle x_1, x_2 \rangle, x_3 \rangle$ for $y \in H^1$. Since $E^2(3)$ is natural in $H_*(V)$, we can assume $x_1, x_2, x_3 \in B$ have internal degree k respectively. There are two cases:

- (1) The class $(\bar{Q}_j(y_1 \otimes x_1)) \otimes (y_2 \otimes x_2) \in E_{2,*}^2(3)$ has internal degree at most $3k - 5$ for all $y_1, y_2 \in H$, while the class $y \otimes \langle \langle x_1, x_2 \rangle, x_1 \rangle$ has internal degree $3k - 3$ for all $y \in H^1$. Hence they do not support d^2 -differentials.
- (2) The other type of classes in filtration 2 are of the form $(y_1 \otimes x_1) \otimes (y_2 \otimes x_2) \otimes (y_3 \otimes x_3)$ with internal degrees at most $3k - 5$, while the class $y \otimes \langle \langle x_1, x_2 \rangle, x_3 \rangle$ has internal degree $3k - 3$. Hence these classes do not support d^2 -differentials either.

Therefore the weight three part of the spectral sequence collapses at the E^2 -page, and we obtain a basis for $H_*(B_3(\Sigma_{g,1}; X))$.

For the closed surface Σ_1 , $\tilde{\mathfrak{g}} = H^*(\Sigma_1) \otimes \text{Free}_{\text{Mod } \mathbb{F}_2}^{\text{Lie}^{s,\text{ti}}_{\tilde{\mathfrak{R}}}(\Sigma^n H_*(X))}(\Sigma^n H_*(X))$ and [Corollary 5.8](#) says that

$$\begin{aligned} E^\infty(3) = E^2(3) \cong & \bigoplus_{x_1, x_2 \in B} \mathbb{F}_2\{(\bar{Q}_0(y \otimes x_1)) \otimes (y' \otimes x_2), y \in H^1, y' \in H \cup \{d\}\} \\ & \oplus \bigoplus_{x_1, x_2 \in B} \mathbb{F}_2\{(\bar{Q}_0(c \otimes x_1)) \otimes (y \otimes x_2), (\bar{Q}_1(c \otimes x)) \otimes (y \otimes x_2), y \in H \cup \{d\}\} \\ & \oplus \text{wt}_3(H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}})). \end{aligned}$$

We do not list the $\mathbb{F}_2$ -basis of $\text{wt}_3(H_{*,*}^{\text{Lie}^{s,\text{ti}}}(\tilde{\mathfrak{g}}))$ for simplicity.

Example 5.10 As in the weight two case, our basis for $H_*(B_3(\Sigma_1, \mathbb{S}^k))$, $k \geq 1$ is in bijection with the weight 3 part of Equation (14) by sending $y \otimes x$ to $x_{k+2-|y|}$ and $\bar{Q}_i(y \otimes x)$ to $Q_i(x_{k+2-|y|})$ for $y = a, b, c, d$.

5.4 Example computations: (punctured) real projective space

The simplest examples of parallelizable manifolds admitting nontrivial Steenrod actions other than Sq^0 are the real projective space $\mathbb{RP}^3$ and the once-punctured real projective space $\mathbb{RP}^3$.

Let y be a generator for $H^1(\mathbb{RP}^3)$. Then

$$\tilde{H}^*((\mathbb{RP}^3)^+) \cong H^*(\mathbb{RP}^3) = \mathbb{F}_2[y]/(y^4), \tilde{H}^*((\mathbb{RP}^3)^+) = \tilde{H}^*(\mathbb{RP}^3) = \mathbb{F}_2\{y, y^2, y^3\}$$

with the obvious cup products and one nontrivial Steenrod operation $\text{Sq}^1(y) = y^2$.

5.4.1 Weight two We deduce $H_*(B_2(\mathbb{RP}^3; X))$ and $H_*(B_2(\mathbb{RP}^3; X))$ from Corollary 5.6. For $M = \mathbb{RP}^3$, there is only one nontrivial cup product $y \cup y^2 = y^3$, so

$$\begin{aligned} E^\infty(2) = E^2(2) = & \bigoplus_{x \in B, a=1,2,3} \mathbb{F}_2\{\bar{Q}_j(y^a \otimes x), 0 \leq j < a\} \oplus \bigoplus_{x \in B} \mathbb{F}_2\{(y^a \otimes x) \otimes (y^b \otimes x), 1 \leq a < b \leq 3\} \\ & \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2\{y \otimes \langle x_1, x_2 \rangle; (y^a \otimes x_1) \otimes (y^a \otimes x_2), a = 2, 3\} \\ & \oplus \bigoplus_{x_1 \neq x_2 \in B} \mathbb{F}_2\{(y^a \otimes x_1) \otimes (y^3 \otimes x_2), a = 1, 2\} \\ & \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2\{(y^1 \otimes x_1) \otimes (y^2 \otimes x_2) + (y^2 \otimes x_1) \otimes (y^1 \otimes x_2)\}. \end{aligned}$$

For $M = \mathbb{RP}^3$, the nonzero cup products are $y \cup y = y^2$, $y \cup y^2 = y^3$ and $1 \cup y^a = y^a$ for $0 \leq a \leq 3$, so

$$\begin{aligned} E^\infty(2) = E^2(2) = & \bigoplus_{x \in B, a=1,2,3} \mathbb{F}_2\{\bar{Q}_j(y^a \otimes x), 0 \leq j < a\} \oplus \bigoplus_{x \in B} \mathbb{F}_2\{(y^a \otimes x) \otimes (y^b \otimes x), 0 \leq a < b \leq 3\} \\ & \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2\{(y^a \otimes x_1) \otimes (y^a \otimes x_2), a = 2, 3\} \\ & \oplus \bigoplus_{x_1 \neq x_2 \in B} \mathbb{F}_2\{(y^a \otimes x_1) \otimes (y^3 \otimes x_2), a = 1, 2\} \\ & \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2\{(y^a \otimes x_1) \otimes (y^b \otimes x_2) + (y^3 \otimes x_1) \otimes (1 \otimes x_2), (a, b) \neq (3, 0)\} \\ & \oplus \bigoplus_{x_1 < x_2 \in B} \mathbb{F}_2\{(y^a \otimes x_1) \otimes (1 \otimes x_2) + (1 \otimes x_1) \otimes (y^a \otimes x_2), a = 1, 2, 3\}. \end{aligned}$$

Example 5.11 When $X = \mathbb{S}^k$ with $k \geq 1$, we have $B = \{x = \sigma_2 \iota_k\}$, so $H_*(B_2(\mathbb{RP}^3, S^k))$ has $\mathbb{F}_2$ -basis

$$\{\bar{Q}_j(y^a \otimes x), 0 \leq j < a, a = 1, 2, 3; (y^a \otimes x) \otimes (y^b \otimes x), 0 \leq a < b \leq 3\}.$$

A bijection with weight 3 part of the Bödighheimer–Cohen–Taylor decomposition [12],

$$\begin{aligned} \bigoplus_{k \geq 1} H_*(B_k(\mathbb{RP}^3; S^k)) &\cong \bigotimes_{i=0}^n H_*(\Omega^{3-i} S^{3+k})^{\otimes \dim H_i(M)} \\ &\cong H_*(\Omega^3 \Sigma^3 S^k) \otimes H_*(\Omega^2 \Sigma^2 S^{k+1}) \otimes H_*(\Omega \Sigma S^{k+2}) \otimes H_*(S^{k+3}) \\ &\cong \text{Free}^{E^3}(\mathbb{F}_2\{x_k\}) \otimes \text{Free}^{E^2}(\mathbb{F}_2\{x_{k+1}\}) \otimes \text{Free}^{E^1}(\mathbb{F}_2\{x_{k+2}\}) \otimes \mathbb{F}_2\{x_k\}, \end{aligned}$$

is given by sending $y^a \otimes x$ to x_{k+3-a} and $\bar{Q}_i(y^a \otimes x)$ to $Q_i(x_{k+3-a})$ for $0 \leq a \leq 3$.

5.4.2 Weight three For the closed manifold $\mathbb{RP}^3$ and $\tilde{\mathfrak{g}} = H^*(\mathbb{RP}^3) \otimes \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\mathbb{F}_2}^{s, \text{ti}}}(\Sigma^n H_*(X))$, it follows from Corollary 5.8 that

$$E^\infty(3) = E^2(3) \cong \text{wt}_3(H_{*,*}^{\text{Lie}_{*,*}^{s, \text{ti}}}(\tilde{\mathfrak{g}})) \oplus \bigoplus_{\substack{x_1, x_2 \in B \\ 1 \leq a \leq 3, 0 \leq b \leq 3}} \mathbb{F}_2\{(\bar{Q}_j(y^a \otimes x_1)) \otimes (y^b \otimes x_2), 0 \leq j < a\}.$$

For the punctured real projective space $\mathbb{RP}^3$ and $\tilde{\mathfrak{g}} = \tilde{H}^*(\mathbb{RP}^3) \otimes \text{Free}_{\text{Mod}_{\mathbb{F}_2}}^{\text{Lie}_{\mathbb{F}_2}^{s, \text{ti}}}(\Sigma^n H_*(X))$, weight three classes in $E^2(3)$ either live in $\text{wt}_3(H_{*,*}^{\text{Lie}_{*,*}^{s, \text{ti}}}(\tilde{\mathfrak{g}}))$ or has the form

$$(\bar{Q}_j(y^a \otimes x)) \otimes (y^b \otimes x') \in A \otimes H_{*,*}^{\text{Lie}_{*,*}^{s, \text{ti}}}(\tilde{\mathfrak{g}})$$

with $x, x' \in B$ and $1 \leq a, b \leq 3$. Therefore

$$E^2(3) = \text{wt}_3(H_{*,*}^{\text{Lie}_{*,*}^{s, \text{ti}}}(\tilde{\mathfrak{g}})) \oplus \bigoplus_{x_1, x_2 \in B, 1 \leq a, b \leq 3} \mathbb{F}_2\{(\bar{Q}_j(y^a \otimes x_1)) \otimes (y^b \otimes x_2), 0 \leq j < a\}.$$

A complete list of an $\mathbb{F}_2$ -basis for $\text{wt}_3(H_{*,*}^{\text{Lie}_{*,*}^{s, \text{ti}}}(\tilde{\mathfrak{g}}))$ is given by

- (1) $y \otimes \langle x_1, x_2 \rangle, x_3$ for $x_1, x_2, x_3 \in B$, $x_1 < x_2$, $x_1 < x_3$ in simplicial degree 0;
- (2) $(y^3 \otimes \langle x_1, x_2 \rangle) \otimes (y^b \otimes x_3) + (y^3 \otimes \langle x_1, x_3 \rangle) \otimes (y^b \otimes x_2) + (y^3 \otimes \langle x_2, x_3 \rangle) \otimes (y^b \otimes x_1)$ for $b = 1, 2$ and $(y \otimes \langle x_1, x_2 \rangle) \otimes (y^2 \otimes x_3) + (y \otimes \langle x_1, x_3 \rangle) \otimes (y^2 \otimes x_2) + (y \otimes \langle x_2, x_3 \rangle) \otimes (y^2 \otimes x_1)$ for distinct $x_i \in B$ in simplicial degree 1;
- (3) $(y^a \otimes x_1) \otimes (y^b \otimes x_2) \otimes (y^c \otimes x_3)$ for $\{1, 2\}, \{1, 1\} \not\subseteq \{a, b, c\}$ and $x_i \in B$;

$$\sum_{\{i, j, k\}=\{1, 2, 3\}, i < j} (y \otimes x_i) \otimes (y \otimes x_j) \otimes (y^2 \otimes x_k), \quad \sum_{\{i, j, k\}=\{1, 2, 3\}, j < k} (y \otimes x_i) \otimes (y^2 \otimes x_j) \otimes (y^2 \otimes x_k)$$

for distinct $x_1, x_2, x_3 \in B$ in simplicial degree 2.

Again the E^2 -page is concentrated in simplicial degrees 0, 1, 2, and we use sparsity to rule out higher differentials. Suppose that x_1, x_2, x_3 have internal degree k . We examine the two cases that could potentially support a d^2 -differential.

- (1) The class $(\bar{Q}_j(y^a \otimes x_1)) \otimes (y^b \otimes x_2) \in E_{2,*}^2(3)$ has internal degree at most $3k - 5$ for all $1 \leq a, b \leq 3$, while the class $y \otimes \langle x_1, x_2 \rangle, x_1$ has internal degree $3k - 3$. Hence they do not support d^2 -differentials.

- (2) The other type of classes in simplicial degree 2 are of the form $(y^a \otimes x_1) \otimes (y^b \otimes x_2) \otimes (y^c \otimes x_3)$ with internal degrees at most $3k - 5$, while the class $y \otimes \langle \langle x_1, x_2 \rangle, x_3 \rangle$ has internal degree $3k - 3$. Hence these classes do not support d^2 -differentials either.

Therefore the weight three part of the spectral sequence collapses on the E^2 -page, and we obtain a basis for $H_*(B_3(\mathbb{R}\mathbb{P}^3; X))$.

6 Odd primary homology

In this last section, we apply the same methods to study the mod p homology of $B_k(M; X)$ for $p > 2$ via the Knudsen spectral sequence with $\mathbb{F}_p$ coefficient.

6.1 Odd primary Knudsen spectral sequence

We start by recalling partial progress in understanding the unary operations on the mod p homology of spectral Lie algebras by Kjaer [37]. He constructed weight p Dyer–Lashof-type operations in analogy to Behrens’ construction of $\bar{Q}^j$, which was further clarified by the work of Konovalov.

Proposition 6.1 [37, Definition 3.2; 40, Definition 2.5.17] *Let L be a spectral Lie algebra. Then $H_*(L; \mathbb{F}_p)$ admits unary operations*

$$\overline{\beta^\epsilon Q^j}: H_*(L; \mathbb{F}_p) \rightarrow H_{*+2(p-1)i-\epsilon-1}(L; \mathbb{F}_p), \quad \epsilon \in \{0, 1\}, \quad j \in \mathbb{Z}.$$

On a class $x \in H_*(L; \mathbb{F}_p)$ such that if $|x|$ is even then $2j \neq |x|$, the class $\overline{\beta^\epsilon Q^j}(x)$ is given by $\xi_*(\sigma^{-1} \beta^\epsilon Q^j(x))$, where $\beta^\epsilon Q^j$ is a mod p Dyer–Lashof operation, σ^{-1} the desuspension isomorphism, and $\xi: \partial_p(\text{Id}) \otimes_{h\Sigma_p} L^{\otimes p} \rightarrow L$ the p^{th} structure map of the spectral Lie algebra L . When $|x| = 2l$, define $\overline{\beta Q^l}(x)$ via the isomorphism $H_*(\partial_p(\text{id}) \otimes_{h\Sigma_p} (S^{2l})^{\otimes p}) \cong H_*(\Sigma^{-1}(\partial_p(\text{id}) \otimes_{h\Sigma_p} (S^{2l-1})^{\otimes p}))$.

It follows from the instability condition of Dyer–Lashof operations that the allowability condition for the operations $\overline{\beta^\epsilon}$ are given by $\overline{\beta^\epsilon Q^j}(x) = 0$ if $j < \frac{1}{2}|x|$. Analogous to the case $p = 2$, brackets of unary operations always vanish.

Proposition 6.2 [37, Proposition 3.7] *For L a spectral Lie algebra, $[\overline{\beta^\epsilon Q^j}(x), y] = 0$ for any ϵ, j and $x, y \in H_*(L; \mathbb{F}_p)$.*

The relations among the unary operations were obtained by Konovalov.

Proposition 6.3 [40, Theorem 8.2.14] *Let $\bar{\mathcal{R}}$ be the free algebra over $\mathbb{F}_p$ on generators $\overline{\beta^\epsilon Q^j}$, $\epsilon \in \{0, 1\}$, subject to the relations*

$$\begin{aligned} \overline{\beta^\epsilon Q^j} \cdot \overline{\beta Q^i} &= (-1)^{\epsilon+1} \sum_{m=pi}^{i+j-1} \binom{p(m-i) - (p-1)j + \epsilon - 1}{m-pi} \overline{\beta Q^m} \cdot \overline{\beta^\epsilon Q^{j+i-m}} \\ &\quad + (1-\epsilon) \sum_{m=pi+1}^{i+j-1} \binom{p(m-i) - (p-1)j}{m-pi} \overline{Q^m} \cdot \overline{\beta Q^{j+i-m}} \end{aligned}$$

for $j < pi$, and

$$\overline{\beta^\epsilon Q^j} \cdot \overline{Q^i} = \sum_{m=pi+1}^{i+j-1} \binom{p(m-i)-(p-1)j-1}{m-pi-1} \overline{\beta^\epsilon Q^m} \cdot \overline{Q^{j+i-m}}$$

for $j \leq pi$. Then the mod p homology of a spectral Lie algebra is an allowable module over $\overline{\mathcal{R}}$.

Denote by $\mathcal{A}_{\overline{\mathcal{R}}}$ the free allowable $\overline{\mathcal{R}}$ -module monad. Let $\text{Lie}_{\overline{\mathcal{R}}}^s : \text{Mod}_{\mathbb{F}_p} \rightarrow \text{Mod}_{\mathbb{F}_p}$ be the composite monad $\mathcal{A}_{\overline{\mathcal{R}}} \circ \text{Lie}_{\mathbb{F}_p}^s$ subject to the commuting relations [Proposition 6.2](#) when $p > 3$, and the monad given by $\text{Lie}_{\overline{\mathcal{R}}}^s(M) = \mathcal{A}_{\overline{\mathcal{R}}} \circ \text{Lie}_{\mathbb{F}_3}^s(M) / \langle \overline{\beta^\epsilon Q^{|x|/2}}(x) = [[x, x], x] \rangle$, where we take the quotient by the $\overline{\mathcal{R}}$ -module ideal ranging over $x \in M$ in even degree. For $M \in \text{Mod}_{\mathbb{F}_p}$, let A be an $\mathbb{F}_p$ -basis for the free shifted Lie algebra $\text{Free}_{\text{Mod}_{\mathbb{F}_p}}^{\text{Lie}_{\mathbb{F}_p}^s}(M)$. The graded $\mathbb{F}_p$ -module $\text{Lie}_{\overline{\mathcal{R}}}^s(M)$ has basis

$$\{\overline{\beta_1^{\epsilon_1} Q^{j_1}} \dots \overline{\beta_k^{\epsilon_k} Q^{j_k}} | x, x \in A, j_k \geq \frac{1}{2}|x|, j_i \geq pj_{i+1} - \epsilon_{i+1} \forall i\}.$$

Theorem 6.4 [[37](#), Theorem 5.2; [40](#), Theorem 8.2.17] *For X a spectrum. there is an isomorphism of $\text{Lie}_{\overline{\mathcal{R}}}^s$ -algebras*

$$\text{Lie}_{\overline{\mathcal{R}}}^s(H_*(X; \mathbb{F}_p)) \rightarrow H_*(\text{Free}^{s\mathcal{L}}(X); \mathbb{F}_p).$$

Remark 6.5 For $p = 3$, Kjaer claimed in [[37](#), Corollary 4.7] that the triple bracket on an even degree homology class ι_{2l} of a spectral Lie algebra is zero by showing that

$$[[\iota_{2l}, \iota_{2l}], \iota_{2l}] \in H_*(\partial_3(\text{id}) \otimes_{h\Sigma_3} (S^{2l})^{\otimes 3})$$

vanishes. The claim is incorrect in light of [Proposition 6.8](#) below, and was independently observed by Nikolai Konovalov. Specifically, Kjaer argued that in the long exact sequence

$$\dots \rightarrow H_{6l-2}(\Sigma^{-2}(S^{2l})_{h\Sigma_3}^{\otimes 3}) \rightarrow H_{6l-2}(\partial_3(\text{id}) \otimes_{h\Sigma_3} (S^{2l})^{\otimes 3}) \rightarrow H_{6l-2}((\Sigma^{-1}(S^{2l})_{h\Sigma_3}^{\otimes 3}) \rightarrow \dots,$$

the middle group is generated as an $\mathbb{F}_3$ -module by the bottom operation $\overline{\beta Q^l} \iota_{2l}$, which is mapped isomorphically onto $\sigma^{-1} \beta Q^l \iota_{2l}$ by definition of the bottom operation in [[loc. cit.](#), Definition 3.2]. However, $\sigma^{-1} \beta Q^l \iota_{2l} \in H_{6l-2}(\Sigma^{-1}(S^{2l})_{h\Sigma_3}^{\otimes 3}) = 0$. In fact, one can see that the confusion was caused by incorrect placement of parentheses. Since the left term is one-dimensional on $[[\iota_{2l}, \iota_{2l}], \iota_{2l}]$, we see that $[[\iota_{2l}, \iota_{2l}], \iota_{2l}] = \mu_l \overline{\beta Q^l} \iota_{2l}$, where $\mu_l = \pm 1$. This also motivates the modification of the definition of the bottom operation on an even class in [Proposition 6.1](#).

Now we turn to the odd primary Knudsen spectral sequence

$$(15) \quad E_{s,t}^2(k) = \pi_{s,t}(\text{Bar}_\bullet(\text{id}, \text{Lie}_{\overline{\mathcal{R}}}^s, \mathfrak{g}) \otimes \mathbb{F}_p)(k) \Rightarrow H_{s+t}(B_k(M; X); \mathbb{F}_p),$$

where

$$\mathfrak{g} = H_*(\text{Free}^{s\mathcal{L}}(\Sigma^n X)^{M^+}; \mathbb{F}_p) \cong \tilde{H}^*(M^+; \mathbb{F}_p) \otimes \text{Lie}_{\overline{\mathcal{R}}}^s(\Sigma^n H_*(X; \mathbb{F}_p)).$$

Furthermore, $\mathfrak{g}$ has a $\text{Lie}_{\mathbb{F}_p}^s$ -structure given by [Proposition 4.3](#), ie

$$[y_1 \otimes x_1, y_2 \otimes x_2] := (-1)^{|x_1||y_2|} (y_1 \cup y_2) \otimes [x_1, x_2].$$

We proceed to compute the E^2 -page of the spectral sequence (15) in low weights in terms of $\text{Lie}_{\mathbb{F}_p}^s$ -algebra homology.

Definition 6.6 [22; 45] For a shifted Lie algebra L over $\mathbb{F}_p$, let L_{even} and L_{odd} denote the elements in L with even and odd degrees, respectively. The *Chevalley–Eilenberg* complex of L is the chain complex

$$\text{CE}(L) = (\Gamma^\bullet(L_{\text{even}}) \otimes \Lambda^\bullet(L_{\text{odd}}), \partial),$$

where $\Gamma^\bullet$ and $\Lambda^\bullet$ are respectively the graded, shifted divided power and exterior algebra functor over $\mathbb{F}_p$, and the differential ∂ on a general element

$$\gamma_{k_1}(x_1)\gamma_{k_2}(x_2)\cdots\gamma_{k_m}(x_m)\langle y_1, y_2, \dots, y_n \rangle \in \Gamma^\bullet(L_{\text{even}}) \otimes \Lambda^\bullet(L_{\text{odd}})$$

is given by

$$\begin{aligned} & \sum_{1 \leq i < j \leq m} \gamma_{k_1}(x_1) \cdots \gamma_{k_{i-1}}(x_i) \cdots \gamma_{k_j-1}(x_j) \cdots \gamma_{k_m}(x_m) \langle [x_i, x_j], y_1, \dots, y_n \rangle \\ & + \sum_{1 \leq i < j \leq n} (-1)^{i+j-1} \gamma_{k_1}(x_1) \cdots \gamma_{k_m}(x_m) \langle [y_i, y_j], y_1, \dots, \widehat{y_i}, \dots, \widehat{y_j}, \dots, y_n \rangle \\ & + \frac{1}{2} \sum_{i=1}^m \gamma_{k_1}(x_1) \cdots \gamma_{k_{i-2}}(x_i) \cdots \gamma_{k_m}(x_m) \langle [x_i, x_i], y_1, \dots, y_n \rangle \\ & + \sum_{i=1}^m \sum_{j=1}^n (-1)^{j-1} \gamma_1([x_i, y_j]) \gamma_{k_1}(x_1) \cdots \gamma_{k_{i-1}}(x_i) \cdots \gamma_{k_m}(x_m) \langle y_1, \dots, \widehat{y_j}, \dots, y_n \rangle. \end{aligned}$$

Proposition 6.7 Let M^n be a parallelizable manifold and X any spectrum.

(1) For $k < p$, the weight k part of the spectral sequence

$$E_{s,t}^2(k) = \pi_s \pi_t (\text{Bar}_\bullet(\text{id}, s\mathcal{L}, \text{Free}^{s\mathcal{L}}(\Sigma^n X)^{M^+}) \otimes \mathbb{F}_p)(k) \Rightarrow H_{s+t}(B_k(M; X); \mathbb{F}_p)$$

has E^2 -page given by $\text{wt}_k(H_{*,*}(\text{CE}(\mathfrak{g})))$, where $\mathfrak{g} = \tilde{H}^*(M^+; \mathbb{F}_p) \otimes \text{Lie}_{\mathbb{F}_p}^s(\Sigma^n H_*(X; \mathbb{F}_p))$.

(2) For $p \geq 5$, the weight p part of the spectral sequence has E^2 -page given by

$$E_{*,*}^2(k) \cong \text{wt}_p(H_{*,*}(\text{CE}(\mathfrak{g}))) \oplus \bigoplus_{y \in H, x \in B} \mathbb{F}_p \{ \overline{\beta^\epsilon Q^j} | y \otimes x, \frac{1}{2}(|x| - |y|) \leq j < \frac{1}{2}|x| \},$$

where H is an $\mathbb{F}_p$ -basis of $\tilde{H}^*(M^+; \mathbb{F}_p)$ and B an $\mathbb{F}_p$ -basis of $H_*(X; \mathbb{F}_p)$.

Proof For $k < p$, all elements in the weight k part of the E^2 -page of the spectral sequence do not contain unary operations $\overline{\beta^\epsilon Q^j}$. When $k = p$, nondegenerate elements of weight p on the E^2 -page are either of the form $\overline{\beta^\epsilon Q^j} | y \otimes x \in \text{Lie}_{\mathbb{F}_p}^s(\mathfrak{g})$, $\overline{\beta^\epsilon Q^j}(y \otimes x) \in \mathfrak{g}$, or a bracket of weight p . When $p \geq 5$, the unary operation $\beta^\epsilon Q^j$ cannot be an iteration of brackets on a single element, since $[[x, x], x] = 0$ for any x by the Jacobi identity. Hence there is no d_1 -differential from a weight p bracket to $\overline{\beta^\epsilon Q^j} | y \otimes x$ or $y \otimes \overline{\beta^\epsilon Q^j}(x)$. The same argument in Proposition 5.4 implies that the twisting of the action of $\overline{\beta^\epsilon Q^j}$ by Steenrod operations can be ignored when computing a basis for the E^2 -page. $\square$

The condition $p \geq 5$ in part (2) is necessary in light of the following computation for Euclidean spaces.

Proposition 6.8 For $p \geq 5$, the only higher differential in the weight p part of the spectral sequence (15) for $M = \mathbb{R}^n$, $2 \leq n \leq \infty$, which converges to $H_*(B_p(\mathbb{R}^n; \mathbb{S}^{2l}); \mathbb{F}_p)$, is a d_{p-2} -differential

$$\gamma_p(x) \mapsto \overline{\beta^\epsilon Q^l} | y_n \otimes \sigma^n(x).$$

When $p = 3$, the above spectral sequence has a d^1 -differential $\gamma_3(x) \mapsto \overline{\beta^\epsilon Q^l} | x$.

Heuristically, this is because the bottom nonvanishing mod p Dyer–Lashof operation on a class x of degree $2l$ in the mod p homology of an $\mathbb{E}_n$ -algebra is given by $\overline{Q^l}(x) = x^{\otimes p}$, so $\gamma_p(x)$ is redundant.

Proof Consider the spectral sequence (15) when $M = \mathbb{R}^n$ and $X = \mathbb{S}^{2l}$ with $n > 2$, so

$$\mathfrak{g} = \mathbb{F}_p\{y_n\} \otimes \text{Lie}_{\mathcal{R}}^s(\mathbb{F}_p\{\sigma^n(x_{2l})\})$$

with y in internal degree $-n$ and x_{2l} in degree $2l$. Set $x = y_n \otimes \sigma^n(x_{2l})$. Then the weight p part of the E^2 -page has basis

$$\{\overline{\beta^\epsilon Q^j} | x, l \leq j < \frac{1}{2}(2l + n); \gamma_p(x)\}.$$

Comparing with the weight p part of the E^∞ -page, which is the weight p part of the mod p homology of the free $\mathbb{E}_n$ -algebra on the $\mathbb{S}^{2l}$, we see that there are two classes that do not survive to the E^∞ -page, ie $\gamma_p(x)$ in bidegree $(p-1, 2pl - (p-1))$ and $\overline{\beta^\epsilon Q^l} | x$ in bidegree $(1, 2pl - 2)$ (see [24, III]). Hence there has to be a d_{p-2} -differential from $\gamma_p(x)$ to $\overline{\beta^\epsilon Q^l} | x$.

When $p = 3$, $\gamma_3(x)$ is represented by the element $[[x, x], x] \in \text{Lie}_{\mathbb{F}_p}^s \circ \text{Lie}_{\mathbb{F}_p}^s(\mathfrak{g}) \subset \text{Lie}_{\mathcal{R}}^s \circ \text{Lie}_{\mathcal{R}}^s(\mathfrak{g})$. It is mapped by the differential to $[[x, x], x] \in \text{Lie}_{\mathbb{F}_p}^s(\mathfrak{g})$, which by Remark 6.5 is indeed $\overline{\beta^\epsilon Q^l} | x$. $\square$

As an immediate corollary to Proposition 6.7, we see that the weight two part of the spectral sequence (15) collapses on the E^2 -page, since the E^2 -page is concentrated in simplicial degree 0 and 1. When $p > 3$, weight three elements on the E^2 -page are in simplicial degree 1 or 2 since $[[x, x], x] = 0$ by the Jacobi identity. Hence the weight three part of the spectral sequence (15) also collapses on the E^2 -page.

Corollary 6.9 Let M^n be a parallelizable manifold and X any spectrum. Let $\mathfrak{g}$ be the $\text{Lie}_{\mathbb{F}_p}^s$ -algebra $\tilde{H}^*(M^+; \mathbb{F}_p) \otimes \text{Lie}_{\mathbb{F}_p}^s(\Sigma^n H_*(X; \mathbb{F}_p))$.

(1) For all i , there is an isomorphism of $\mathbb{F}_p$ -modules

$$H_i(B_2(M; X); \mathbb{F}_p) \cong \bigoplus_{s+t=i} \text{wt}_2(H_{s,t}(\text{CE}(\mathfrak{g}))).$$

(2) If $p \geq 5$, then for all i

$$H_i(B_3(M; X); \mathbb{F}_p) \cong \bigoplus_{s+t=i} \text{wt}_3(H_{s,t}(\text{CE}(\mathfrak{g}))).$$

Remark 6.10 For M a connected n -manifold, Bödigheimer, Cohen and Taylor [12] showed that

$$\bigoplus_{k \geq 1} H_*(B_k(M; S^r); \mathbb{F}_p) \cong \bigotimes_{i=0}^n H_*(\Omega^{n-i} S^{n+r}; \mathbb{F}_p)^{\otimes \dim H_i(M; \mathbb{F}_p)}$$

for $r + n$ odd and $r \geq 0$. Their proof does not work in the case where $r + n$ is even due to the existence of nontrivial self-brackets in $H_*(\Omega^m \Sigma^m S^l; \mathbb{F}_p)$ when l is even. Roughly speaking, their inductive proof relies on the canonical map

$$H_*(\Omega^m \Sigma^m S^l; \mathbb{F}_p) \rightarrow H_*(\Omega^\infty \Sigma^\infty S^l; \mathbb{F}_p)$$

being an injection, which is only true when l is odd. [Corollary 6.9](#) shows that when l is even, the mod p homology of $B_k(M; S^r)$, $k = 2, 3$ depends on the cup product structure on $H^*(M^+; \mathbb{F}_p)$: if $a \cup b = c$ in $\tilde{H}^*(M^+; \mathbb{F}_p)$, then the d_1 -differential sends $(a \otimes x) \otimes (b \otimes x)$ to

$$c \otimes [x, x] \in \mathfrak{g} = \tilde{H}^*(M^+; \mathbb{F}_p) \otimes \mathrm{Lie}_{\mathcal{R}}^s(\mathbb{F}_p\{x\}),$$

which is not zero since x has internal degree l .

At higher weights, there generally will be higher differentials in the odd primary Knudsen spectral sequence (15). In recent work with Matthew Chen [21], we make use of [Proposition 6.7](#) and the Drummond-Cole–Knudsen computation [25] of the rational homology of the unordered configurations space $B_k(M)$, where $M = \Sigma_1$ or $\Sigma_{g,1}$ to identify the differentials in the Knudsen spectral sequence for $B_k(\Sigma_g; \mathbb{S})$. As a result, we show that the integral homology of $B_k(\Sigma_1)$ is p -torsion-free for $k \leq p$. The same argument works for the punctured surface $\Sigma_{g,1}$ with $g \geq 0$, thereby providing a more elementary proof for [17, Theorem 1.10].

References

- [1] **J Adem**, *The iteration of the Steenrod squares in algebraic topology*, Proc. Nat. Acad. Sci. U.S.A. 38 (1952) 720–726 [MR](#) [Zbl](#)
- [2] **O Antolín Camarena**, *The mod 2 homology of free spectral Lie algebras*, Trans. Amer. Math. Soc. 373 (2020) 6301–6319 [MR](#) [Zbl](#)
- [3] **G Z Arone**, **D L B Brantner**, *The action of Young subgroups on the partition complex*, Publ. Math. Inst. Hautes Études Sci. 133 (2021) 47–156 [MR](#) [Zbl](#)
- [4] **G Arone**, **M Kankaanrinta**, *The homology of certain subgroups of the symmetric group with coefficients in $\mathrm{Lie}(n)$* , J. Pure Appl. Algebra 127 (1998) 1–14 [MR](#) [Zbl](#)
- [5] **G Arone**, **M Mahowald**, *The Goodwillie tower of the identity functor and the unstable periodic homotopy of spheres*, Invent. Math. 135 (1999) 743–788 [MR](#) [Zbl](#)
- [6] **D Ayala**, **J Francis**, *Factorization homology of topological manifolds*, J. Topol. 8 (2015) 1045–1084 [MR](#) [Zbl](#)
- [7] **M Basterra**, *André–Quillen cohomology of commutative S -algebras*, J. Pure Appl. Algebra 144 (1999) 111–143 [MR](#) [Zbl](#)

- [8] **J Beck**, *Distributive laws*, from “Seminar on Triples and Categorical Homology Theory”, Lecture Notes in Math. 80, Springer (1969) 119–140 [MR](#) [Zbl](#)
- [9] **M Behrens**, *The Goodwillie tower and the EHP sequence*, Mem. Amer. Math. Soc. 1026, Amer. Math. Soc., Providence, RI (2012) [MR](#) [Zbl](#)
- [10] **C-F Bökigheimer, F R Cohen**, *Rational cohomology of configuration spaces of surfaces*, from “Algebraic topology and transformation groups”, Lecture Notes in Math. 1361, Springer (1988) 7–13 [MR](#) [Zbl](#)
- [11] **C-F Bökigheimer, F R Cohen, R J Milgram**, *Truncated symmetric products and configuration spaces*, Math. Z. 214 (1993) 179–216 [MR](#) [Zbl](#)
- [12] **C-F Bökigheimer, F Cohen, L Taylor**, *On the homology of configuration spaces*, Topology 28 (1989) 111–123 [MR](#)
- [13] **M Bökstedt, I Ottosen**, *An alternative approach to homotopy operations*, Trans. Amer. Math. Soc. 358 (2006) 3985–3995 [MR](#) [Zbl](#)
- [14] **A K Bousfield**, *Operations on derived functors of non-additive functors*, preprint (1967) Available at <https://math.mit.edu/~hrm/manuscripts/bousfield-operations.pdf>
- [15] **A K Bousfield, E B Curtis**, *A spectral sequence for the homotopy of nice spaces*, Trans. Amer. Math. Soc. 151 (1970) 457–479 [MR](#) [Zbl](#)
- [16] **D L B Brantner**, *The Lubin–Tate theory of spectral Lie algebras*, PhD thesis, Harvard University (2017) Available at <https://www.proquest.com/docview/2449294095>
- [17] **D L B Brantner, J Hahn, B Knudsen**, *The Lubin–Tate theory of configuration spaces, I*, J. Topol. (online publication October 2024) [Zbl](#)
- [18] **R R Bruner, J P May, J E McClure, M Steinberger**, *H_∞ ring spectra and their applications*, Lecture Notes in Math. 1176, Springer (1986) [MR](#) [Zbl](#)
- [19] **H Cartan**, *Sur les groupes d’Eilenberg–Mac Lane, II*, Proc. Nat. Acad. Sci. U.S.A. 40 (1954) 704–707 [MR](#) [Zbl](#)
- [20] **H Cartan**, *Puissances divisées*, from “Séminaire Henri Cartan, 1954/1955”, Secrét. Math., Paris (1955) Exposé 7
- [21] **M Chen, A Y Zhang**, *mod p homology of unordered configuration spaces of p points in parallelizable surfaces*, Proc. Amer. Math. Soc. 152 (2024) 2239–2248 [MR](#) [Zbl](#)
- [22] **C Chevalley, S Eilenberg**, *Cohomology theory of Lie groups and Lie algebras*, Trans. Amer. Math. Soc. 63 (1948) 85–124 [MR](#) [Zbl](#)
- [23] **M Ching**, *Bar constructions for topological operads and the Goodwillie derivatives of the identity*, Geom. Topol. 9 (2005) 833–933 [MR](#) [Zbl](#)
- [24] **F R Cohen, T J Lada, J P May**, *The homology of iterated loop spaces*, Lecture Notes in Math. 533, Springer (1976) [MR](#) [Zbl](#)
- [25] **G C Drummond-Cole, B Knudsen**, *Betti numbers of configuration spaces of surfaces*, J. Lond. Math. Soc. 96 (2017) 367–393 [MR](#) [Zbl](#)
- [26] **W G Dwyer**, *Higher divided squares in second-quadrant spectral sequences*, Trans. Amer. Math. Soc. 260 (1980) 437–447 [MR](#) [Zbl](#)
- [27] **W G Dwyer**, *Homotopy operations for simplicial commutative algebras*, Trans. Amer. Math. Soc. 260 (1980) 421–435 [MR](#) [Zbl](#)

- [28] **E Dyer, R K Lashof**, *Homology of iterated loop spaces*, Amer. J. Math. 84 (1962) 35–88 [MR](#) [Zbl](#)
- [29] **Y Félix, J-C Thomas**, *Rational Betti numbers of configuration spaces*, Topology Appl. 102 (2000) 139–149 [MR](#) [Zbl](#)
- [30] **B Fresse**, *Koszul duality of operads and homology of partition posets*, from “Homotopy theory: relations with algebraic geometry, group cohomology, and algebraic K -theory”, Contemp. Math. 346, Amer. Math. Soc., Providence, RI (2004) 115–215 [MR](#) [Zbl](#)
- [31] **V Ginzburg, M Kapranov**, *Koszul duality for operads*, Duke Math. J. 76 (1994) 203–272 [MR](#) [Zbl](#)
- [32] **P G Goerss, J F Jardine**, *Simplicial homotopy theory*, Progr. Math. 174, Birkhäuser, Basel (1999) [MR](#) [Zbl](#)
- [33] **M Hall, Jr.**, *A basis for free Lie rings and higher commutators in free groups*, Proc. Amer. Math. Soc. 1 (1950) 575–581 [MR](#) [Zbl](#)
- [34] **R Haugseng, H Miller**, *On a spectral sequence for the cohomology of infinite loop spaces*, Algebr. Geom. Topol. 16 (2016) 2911–2947 [MR](#) [Zbl](#)
- [35] **B Johnson**, *The derivatives of homotopy theory*, Trans. Amer. Math. Soc. 347 (1995) 1295–1321 [MR](#) [Zbl](#)
- [36] **N Johnson, J Noel**, *Lifting homotopy T -algebra maps to strict maps*, Adv. Math. 264 (2014) 593–645 [MR](#) [Zbl](#)
- [37] **J J Kjaer**, *On the odd primary homology of free algebras over the spectral Lie operad*, J. Homotopy Relat. Struct. 13 (2018) 581–597 [MR](#) [Zbl](#)
- [38] **B Knudsen**, *Betti numbers and stability for configuration spaces via factorization homology*, Algebr. Geom. Topol. 17 (2017) 3137–3187 [MR](#) [Zbl](#)
- [39] **B Knudsen**, *Higher enveloping algebras*, Geom. Topol. 22 (2018) 4013–4066 [MR](#) [Zbl](#)
- [40] **N Konovalov**, *Algebraic Goodwillie spectral sequence*, preprint (2023) [arXiv 2303.06240](#)
- [41] **I Kříž**, *Towers of $\mathbb{E}_\infty$ ring spectra with an application to BP* , preprint (1993)
- [42] **I Kříž**, *On the rational homotopy type of configuration spaces*, Ann. of Math. 139 (1994) 227–237 [MR](#) [Zbl](#)
- [43] **T Kudo, S Araki**, *Topology of H_n -spaces and H -squaring operations*, Mem. Fac. Sci. Kyūsyū Univ. A 10 (1956) 85–120 [MR](#) [Zbl](#)
- [44] **N J Kuhn**, *The homology of the James–Hopf maps*, Illinois J. Math. 27 (1983) 315–333 [MR](#) [Zbl](#)
- [45] **J P May**, *The cohomology of restricted Lie algebras and of Hopf algebras*, J. Algebra 3 (1966) 123–146 [MR](#) [Zbl](#)
- [46] **J P May**, *A general algebraic approach to Steenrod operations*, from “The Steenrod algebra and its applications”, Lecture Notes in Math. 168, Springer (1970) 153–231 [MR](#) [Zbl](#)
- [47] **J P May**, *The geometry of iterated loop spaces*, Lecture Notes in Math. 271, Springer (1972) [MR](#) [Zbl](#)
- [48] **D McDuff**, *Configuration spaces of positive and negative particles*, Topology 14 (1975) 91–107 [MR](#) [Zbl](#)
- [49] **R J Milgram, P Löffler**, *The structure of deleted symmetric products*, from “Braids”, Contemp. Math. 78, Amer. Math. Soc., Providence, RI (1988) 415–424 [MR](#) [Zbl](#)
- [50] **D Petersen**, *Cohomology of generalized configuration spaces*, Compos. Math. 156 (2020) 251–298 [MR](#) [Zbl](#)
- [51] **S B Priddy**, *Koszul resolutions*, Trans. Amer. Math. Soc. 152 (1970) 39–60 [MR](#) [Zbl](#)

- [52] **S B Priddy**, *Transfer, symmetric groups, and stable homotopy theory*, from “Algebraic K -theory, I: Higher K -theories”, Lecture Notes in Math. 341, Springer (1973) 244–255 [MR](#) [Zbl](#)
- [53] **P Salvatore**, *Configuration operads, minimal models and rational curves*, PhD thesis, University of Oxford (1998)
- [54] **G Segal**, *Configuration-spaces and iterated loop-spaces*, Invent. Math. 21 (1973) 213–221 [MR](#) [Zbl](#)
- [55] **B Totaro**, *Configuration spaces of algebraic varieties*, Topology 35 (1996) 1057–1067 [MR](#) [Zbl](#)
- [56] **J M Turner**, *Operations and spectral sequences, I*, Trans. Amer. Math. Soc. 350 (1998) 3815–3835 [MR](#) [Zbl](#)
- [57] **A Y Zhang**, *Operations on spectral partition Lie algebras and TAQ cohomology*, preprint (2022) [arXiv 2203.15771](#) To appear in Geom. Topol.

Department of Mathematical Sciences, University of Copenhagen
Copenhagen, Denmark

adelazhang224@gmail.com

Received: 23 December 2021 Revised: 1 February 2024

Coarse Alexander duality for pairs and applications

G CHRISTOPHER HRUSKA

EMILY STARK

HÙNG CÔNG TRẦN

For a group G (of type F) acting properly on a coarse Poincaré duality space X , Kapovich and Kleiner introduced a coarse version of Alexander duality between G and its complement in X . More precisely, the cohomology of G with group ring coefficients is dual to a certain Čech homology group of the family of increasing neighborhoods of a G -orbit in X . This duality applies more generally to coarse embeddings of certain contractible simplicial complexes into coarse $\text{PD}(n)$ spaces. In this paper we introduce a relative version of this Čech homology that satisfies the Eilenberg–Steenrod exactness axiom, and we prove a relative version of coarse Alexander duality.

As an application we provide a detailed proof of the following result, first stated by Kapovich and Kleiner. Given a 2-complex formed by gluing k halfplanes along their boundary lines and a coarse embedding into a contractible 3-manifold, the complement consists of k deep components that are arranged cyclically in a pattern called a *Jordan cycle*. We use the Jordan cycle as an invariant in proving the existence of a 3-manifold group that is virtually Kleinian but not itself Kleinian.

20F65, 55M05, 55N05

1 Introduction

The coarse Alexander duality theorem of Kapovich and Kleiner [17] is a powerful homological tool for the study of coarse embeddings via the asymptotic shapes of their complements, generalizing earlier work of Farb and Schwartz [7] and Block and Weinberger [3]. In certain circumstances, a coarse embedding of simplicial complexes $f: K \rightarrow X$ induces a duality between the coarse cohomology of K and a certain Čech homology group associated to the “end” of the complement $X \setminus f(K)$, which, roughly speaking, measures the homology of cycles and boundaries in X that are “far from $f(K)$ ”.

Coarse cohomology was introduced by Roe, who shows in [23, Chapter 5] that for a uniformly acyclic simplicial complex K , the coarse cohomology of K with integral coefficients coincides with the simplicial cohomology with compact supports $H_c^*(K)$. Furthermore, in the presence of a proper, cocompact action of a group G on K , it also coincides with the group cohomology $H^*(G; \mathbb{Z}G)$ of G with coefficients in the group ring $\mathbb{Z}G$. (See also Brown [4, Section VIII.7].)

The *end* of X is the inverse sequence of spaces $\epsilon X = \{X \setminus N_i\}$, where N_i is the closed simplicial ball in X of radius i centered at a fixed basepoint in X and the bonding maps are inclusions. The Čech homology of

the end of X is the inverse limit of the integral homology groups $H_*(\overline{X \setminus N_i})$. In a sense, Čech homology of the end detects the shape of a space near infinity (see Geoghegan [8]; Guilbault [11]; Geoghegan and Swenson [10]). The increasing family of balls may be viewed as increasing neighborhoods of a point in X . If, instead of balls, one chooses $\{N_i\}$ to be a sequence of increasing tubular neighborhoods of a subcomplex $Y \subset X$, one obtains the Y -end of X and the Čech homology of the Y -end, denoted $\check{H}_k^D(Y^c)$. In this notation Y^c denotes the complement $X \setminus Y$. Informally, this Čech homology group may also be called *deep homology of the complement* and its elements are informally called *deep homology classes*.

A coarse PD(n) space, or coarse Poincaré duality space of formal dimension n , is a type of space introduced by Kapovich and Kleiner [17] that has many large scale homological features in common with contractible n -manifolds (see Section 4.1). For example, coarse PD(n) spaces include the universal covers of closed aspherical n -dimensional PL manifolds or any acyclic simplicial complex that admits a free cocompact action by a Poincaré duality group of dimension n .

Suppose X is a coarse PD(n) space, and Z is a uniformly acyclic simplicial complex. Kapovich and Kleiner [17] show that a coarse embedding $f: Z \rightarrow X$ induces a duality isomorphism between the Čech reduced homology $\check{H}_k^D(fZ^c)$ of the $f(Z)$ -end of X and the compactly supported cohomology of Z . This coarse Alexander duality generalizes the classical Alexander duality for embeddings of compacta in the n -sphere to a coarse setting suitable for the study of coarse embeddings between spaces or groups.

In order to facilitate computations in a homology theory, it is useful to introduce relative homology groups and study the corresponding homology long exact sequences. In this article, we introduce a relative version of deep homology of the complement. A simplicial pair (Z_0, Z_1) is a *coarse simplicial pair* if Z_0 and Z_1 are each endowed with a length metric in which each edge has length one, and the inclusion $Z_1 \rightarrow Z_0$ is a coarse embedding. Given a coarse simplicial pair (Z_0, Z_1) , a coarse embedding $f: Z_0 \rightarrow X$ induces a coarse embedding $f: Z_1 \rightarrow X$. In this context, we introduce the Čech relative homology of the $(f(Z_0), f(Z_1))$ -end, denoted $\check{H}_k^D(fZ_1^c, fZ_0^c)$.

A well-known difficulty with the notion of Čech homology is that it typically does not satisfy the Eilenberg–Steenrod exactness axiom, which requires that the homology sequence for any pair be exact (Eilenberg and Steenrod [6]; Milnor [21]). The main source of this difficulty is that Čech homology theory is based on inverse limits, but the limit of an inverse system of exact sequences need not be exact. For this reason, the standard exact sequence techniques of classical algebraic topology are severely limited when working with Čech homology. However, we prove in the setting of coarse embeddings into coarse PD(n) spaces, the Čech homology groups as above do satisfy the exactness axiom.

Theorem 1.1 (deep homology is exact) *Let (Z_0, Z_1) be a coarse simplicial pair of uniformly acyclic complexes, each with bounded geometry. Consider a coarse PD(n) space X and a coarse embedding $f: Z_0 \rightarrow X$. Then the following Čech homology sequence is exact:*

$$\cdots \rightarrow \check{H}_{n-k}^D(fZ_1^c, fZ_0^c) \xrightarrow{\partial} \check{H}_{n-k-1}^D(fZ_0^c) \rightarrow \check{H}_{n-k-1}^D(fZ_1^c) \rightarrow \cdots$$

We refer to [Section 3](#) for precise definitions of uniformly acyclic, bounded geometry, and coarse embeddings.

In particular, as explained by Milnor [\[21\]](#), exactness implies that these Čech homology groups coincide with the corresponding Steenrod homology groups (equivalently, the filtered locally finite homology groups) of the $f(K)$ -end as defined in [\[9, Section 14.2\]](#) by Geoghegan. A key technical step in the proof of [Theorem 1.1](#) is provided by a result of Dydak [\[5\]](#) analogous to the five lemma, but in the setting of exact sequences of inverse sequences (see [Lemma 2.7](#)).

Generalizing Kapovich–Kleiner [\[17\]](#), we show there is a duality isomorphism between compactly supported relative cohomology and relative Čech homology of the complement, which induces an isomorphism of long exact sequences in the following sense.

Theorem 1.2 (coarse Alexander duality of pairs) *Let (Z_0, Z_1) be a coarse simplicial pair of uniformly acyclic complexes, each with bounded geometry. Consider a coarse PD(n) space X and a coarse embedding $f: Z_0 \rightarrow X$. There exists a commutative diagram*

$$\begin{array}{ccccccc} \cdots & \longrightarrow & \check{H}_{n-k}^D(fZ_1^c, fZ_0^c) & \xrightarrow{\partial} & \check{H}_{n-k-1}^D(fZ_0^c) & \longrightarrow & \check{H}_{n-k-1}^D(fZ_1^c) \longrightarrow \cdots \\ & & \cong \downarrow A & & \cong \downarrow A & & \cong \downarrow A \\ \cdots & \xrightarrow{\delta} & H_c^k(Z_0, Z_1) & \longrightarrow & H_c^k(Z_0) & \longrightarrow & H_c^k(Z_1) \xrightarrow{\delta} \cdots \end{array}$$

in which the vertical maps are coarse Alexander duality isomorphisms, and the rows are homology and cohomology long exact sequences.

1.1 Jordan separation and Jordan cycles

The classical Alexander duality theorem proves for any embedding $S^{n-2} \rightarrow S^n$, the complement $S^n \setminus S^{n-2}$ always has the same reduced homology as S^1 , an infinite cyclic group concentrated in dimension one. A generator of this cyclic homology group is called a *Jordan cycle*, by analogy with the codimension-one case of Jordan–Brouwer separation. For a related example, consider the join Θ_k^{n-1} of S^{n-2} with a discrete set of k points. According to Alexander duality, any subspace of S^n homeomorphic to Θ_k^{n-1} separates the n -sphere into k components arranged cyclically in a pattern corresponding to the Jordan cycle. To make precise the similarity between the Jordan cycle and the adjacency of components of $S^n \setminus \Theta_k^{n-1}$, one may use the long exact sequence of a pair as a fundamental computational tool.

The main objects of study in coarse Alexander duality are complexes coarsely embedded in a coarse PD($n+1$) space X . To keep the analogy with the classical setting, we imagine that X is compactified by a fictitious sphere S^n at infinity (a more realistic fantasy might involve a homology n -manifold that is a homology n -sphere as in Bestvina’s article [\[1\]](#)) and that coarse embeddings in X correspond to topological embeddings in S^n .

We demonstrate the use of relative coarse Alexander duality by formalizing the notion of a Jordan cycle in the coarse setting, a notion that is implicit in work of Kapovich and Kleiner [17]. In [17] and in our previous work [14], this notion is used to prove that certain groups are not 3-manifold groups.

By analogy with the embeddings $\Theta_k^{n-1} \rightarrow S^n$, we consider coarse embeddings into a coarse $\text{PD}(n+1)$ space of a union of k half-spaces $\mathbb{R}_{\geq 0}^n$ glued together along their boundaries to form a space W_k . For $n = 2$, these spaces arise naturally in Cayley 2-complexes for Baumslag–Solitar groups and certain other amalgams of surface groups; see [14; 17] and Section 8. The coarse Alexander duality theorem immediately implies that a coarse embedding of W_k into a coarse $\text{PD}(n+1)$ space X coarsely separates it into k deep components corresponding to the set of relative ends $\mathcal{E}(X, f(W_k))$.

We use 1-dimensional relative deep homology groups and the long exact sequence of a pair to define when two deep components of $\mathcal{E}(X, f(W_k))$ are adjacent. More specifically, a pair of deep components is *adjacent* if the image of the connecting homomorphism $\partial: \check{H}_1^D((fW_{k-1}^i)^c, fW_k^c) \rightarrow \check{H}_0^D(fW_k^c)$ for some i is supported on the pair of components, where $W_{k-1}^i \subset W_k$ is obtained by removing the i^{th} halfspace from W_k . These notions define an adjacency graph Γ_k , with vertices corresponding to deep components, and an edge if two components are adjacent.

The coarse Jordan cycle, a 1-dimensional deep homology class, is then used to describe the adjacency relations and the structure of the graph Γ_k . Let $W_0 \subset W_k$ denote the image of the boundaries of the halfspaces in W_k after gluing. Coarse Alexander duality implies that the 1-dimensional homology $\check{H}_1^D(fW_0^c)$ is infinite cyclic. We call a generator of this homology group a *Jordan cycle*. Using relative homology, we recover a proof of the following result that extends a theorem of Kapovich and Kleiner [17, Lemma 7.11].

Theorem 1.3 (*Theorem 7.18.*) *The adjacency graph Γ_k is a circuit. If W_{k-1}^i is obtained from W_k by removing the i^{th} halfspace, the inclusion $W_{k-1}^i \rightarrow W_k$ induces a natural map of cycles $\Gamma_k \rightarrow \Gamma_{k-1}^i$ given by collapsing a single edge to a vertex.*

Although the extra details in the above statement are implicit in the proof of [17, Lemma 7.11], we make this conclusion more explicit using the general methods of this paper. These extra details are required to complete the proof of Theorem 5.6 in [14]. We also use these extra details in an essential way in the proof of Theorem 1.4, discussed below. We suspect that these general methods may prove useful more broadly for studying the relationship between coarse embeddings of groups and coarse embeddings of their various subgroups.

1.2 Applications to 3-manifolds

A *Kleinian group* is a discrete group of isometries of $\mathbb{H}^3$. Torsion-free groups that are virtually Kleinian but that are not the fundamental group of any 3-manifold are constructed by Kapovich and Kleiner in [16]

and by us in [14]. Section 8 is an extended study of one example that arises in [14] for which a more subtle virtual geometric property arises. We show that, among the family of fundamental groups of 3-manifolds, those that admit Kleinian representations are not closed under commensurability.

The fundamental group of the 2-complex formed by gluing a closed orientable hyperbolic surface to a torus along an essential simple closed curve in each surface is a geometrically finite Kleinian group (see Proposition 8.6). In contrast, we show that a certain amalgam of a closed hyperbolic surface group with a Klein bottle group over a $\mathbb{Z}$ subgroup does not admit a Kleinian representation, even though it is a 3-manifold group and it is virtually Kleinian.

Theorem 1.4 *Let X be the space obtained from a closed orientable surface of negative Euler characteristic and a Klein bottle by gluing a nonseparating simple closed curve on the surface to an orientation reversing simple closed curve on the Klein bottle. Then $G = \pi_1(X)$ satisfies the following properties.*

- (1) *G is the fundamental group of a 3-manifold.*
- (2) *G has an index-2 subgroup that is isomorphic to a Kleinian group.*
- (3) *G is not isomorphic to a Kleinian group.*

Statement (1) of Theorem 1.4 is proved by a direct construction in [14]. Statement (2) is a simple application of Thurston's hyperbolization theorem (see Proposition 8.6). The proof of statement (3) uses Jordan cycles as a tool. We briefly sketch this proof here. In order to convert this sketch into a proof, we use the naturality properties of Jordan cycles that are established in Theorem 1.3. We refer the reader to Section 8 for more details. The universal cover of X contains the union of a Euclidean plane and a hyperbolic plane that intersect in a line. The stabilizer H of this line acts on the corresponding adjacency graph given by Theorem 1.3. Since this graph is a circuit of length four, its symmetry group is dihedral. As the induced action of H on the cycle contains a transposition of the Euclidean halfplanes, we conclude that the Euclidean and hyperbolic planes must intersect transversely (in a coarse sense). On the other hand, in any Kleinian representation, the Euclidean plane must coarsely map onto a horosphere, giving a contradiction. We conclude that the virtually Kleinian groups given by Theorem 1.4 are 3-manifold groups, but any such 3-manifold must have a nontrivial geometric decomposition.

Among the family of irreducible 3-manifolds (possibly with boundary) with zero Euler characteristic, the property of admitting a geometric structure (in the sense of Thurston) is a quasi-isometry invariant of the fundamental group by work of Kapovich and Leeb [19] and the geometrization theorem. In contrast, among compact, irreducible 3-manifolds whose boundary has negative Euler characteristic, the property of admitting a geometric structure is not a quasi-isometry invariant of the fundamental group. Indeed Theorem 1.4 illustrates the existence of two compact 3-manifolds with the same fundamental group such that one is geometric and the other is not. On the other hand, we note that the canonical geometric decomposition along tori and annuli as defined, for example, in [22] by Neumann and Swarup, is nontrivial for these manifolds, even though one admits a hyperbolic structure.

1.3 Potential directions for future study

The *action dimension* of a discrete group G is the minimum dimension of a contractible manifold M that admits a proper action of the group G (see Bestvina, Kapovich and Kleiner [2]). If G is torsion free, then the aspherical manifold M/G is a finite-dimensional model for the classifying space BG , so the action dimension is greater than or equal to the geometric dimension, ie the minimal dimension of a $K(G, 1)$ CW complex. In [14], the authors use coarse Alexander duality to show that many amalgams of surface groups over cyclic groups have action dimension four, generalizing examples due to Kapovich and Kleiner [17].

Broadly speaking, the concepts of coarse Alexander duality and its relative version may be useful in the study of actions of groups on contractible n -manifolds M , in terms of restricting the dimension n (as in [14; 17]) or restricting possible geometric structures on M (as in Theorem 1.4). Obstructions to coarse embedding (in particular, the embedding induced by a proper action) are often related to the shape of the complement far from the orbit of the group. Using coarse Alexander duality, one can detect the shape of the complement via the relative Čech homology of the complement. One may be able to use this information as a tool to prove that a group G coarsely embeds into a contractible n -manifold M but cannot act properly on M .

Coarse Alexander duality and its relative version could potentially be used to study the geometric relationship between a finitely generated group G and its subgroups. More precisely, we highlight the following question for which coarse Alexander duality may be useful.

Problem 1.5 Classify subgroups of G up to equivalence under the action of the quasi-isometry group $QI(G)$. More precisely, given two subgroups H and K of a finitely generated group G , determine whether there exists a quasi-isometry of G that maps H to K within a bounded distance.

Study of the above problem may illuminate the general study of the quasi-isometry group of G in the presence of “pattern rigidity” results. In the literature, various quasi-isometry invariants of a pair (G, H) with $H \leq G$ have been studied, such as subgroup distortion and its higher dimensional variants, relative divergence, and many others, but they have limited use. Using coarse Alexander duality, one may be able to use the Čech homology of the complement of $H \leq G$ to better understand the large-scale geometry of G .

Suppose a Poincaré duality group G of dimension n acts freely, cocompactly, and simplicially on a contractible coarse $PD(n)$ space X . The following is a special case of a conjecture due to Bestvina on $\mathbb{Z}$ -set compactifications: does G admit a $\mathbb{Z}$ -structure boundary in the sense of [1]? If so, then a theorem of Bestvina [1, Theorem 2.8] states that (in many cases) such a boundary must be a homology manifold that is also a homology sphere. The existence of such a boundary at infinity would provide a more concrete interpretation of the ends of filtered spaces (and of pairs of filtrations) studied in this paper as direct analogues of classical Alexander duality of subsets (and pairs of subsets) of homology spheres.

1.4 Organization of the paper

Section 2 contains a review of inverse systems and stability properties. In Section 3, we review background on coarse geometry and applications to the study of chain maps. Section 4 introduces Čech homology of complements in ordinary and relative forms and discusses the relation with the notion of ends of pairs of spaces. In Section 5, we review the construction of coarse Alexander duality due to Kapovich and Kleiner. We establish terminology and notation that are useful for introducing the relative version in Section 6. Theorems 1.1 and 1.2 are immediate consequences of Theorem 6.2. Section 7 contains the study of Jordan cycles and the proof of Theorem 1.3. An application to 3-dimensional manifolds is discussed in Section 8, which includes the proof of Theorem 1.4.

Acknowledgements

The authors are thankful for helpful discussions with Arka Banerjee, Max Forester, Ross Geoghegan, Craig Guilbault, Joseph Maher, Alexander Margolis, Boris Okun, Jing Tao, and Genevieve Walsh. We are grateful for the feedback of the referee, whose comments helped to improve the exposition of the paper. Hruska is partially supported by grants #318815 and #714338 from the Simons Foundation. Stark was supported by the Azrieli Foundation; was partially supported at the Technion by a Zuckerman Fellowship; was supported by NSF RTG grant #1840190; and was supported by NSF Grant DMS-2204339. Trần was partially supported by an AMS-Simons Travel Grant.

2 Stability of inverse sequences

In this section, we review some definitions and useful properties of inverse systems and inverse sequences. For detailed background, we refer the reader to either Mardešić and Segal [20] or Geoghegan [9]. In this paper we mainly treat inverse systems using the category of pro-groups. In fact, our main results deal only with inverse systems of abelian groups, so they could be expressed in the category of pro-(Abelian groups). In the language of category theory, we note that the inclusion functor from the pro-Abelian category to the pro-group category is a fully faithful embedding. We work with pro-groups in order to keep statements simpler and more general whenever possible.

We note that in [17] Kapovich and Kleiner work in a related category, which they call *Approx*. As explained in [17, Section 4], their main theorems dealing with inverse sequences in the *Approx* category can all be expressed in the language of pro-(Abelian groups). Throughout this paper we apply this translation without comment whenever citing their results.

Definition 2.1 (inverse system) A *directed set* is a partially ordered set $\mathcal{A}$ such that for each $\alpha_0, \alpha_1 \in \mathcal{A}$ there exists $\alpha \in \mathcal{A}$ with $\alpha_0 \leq \alpha$ and $\alpha_1 \leq \alpha$.

An *inverse system* $(G_\alpha, p_\alpha^{\alpha'})$ consists of a family of groups $(G_\alpha)_{\alpha \in \mathcal{A}}$ indexed by a directed set $\mathcal{A}$ and a collection of homomorphisms $p_\alpha^{\alpha'}: G_{\alpha'} \rightarrow G_\alpha$ for $\alpha \leq \alpha'$ in $\mathcal{A}$ satisfying the following rules:

- $p_{\alpha}^{\alpha'} p_{\alpha'}^{\alpha''} = p_{\alpha}^{\alpha''}$ whenever $\alpha \leq \alpha' \leq \alpha''$, and
- p_{α}^{α} is the identity homomorphism for each α .

The groups G_{α} are the *terms*, the maps $p_{\alpha}^{\alpha'}$ are the *bonding maps*, and $\mathcal{A}$ is the *index set* of the inverse system. When the bonding maps and index set are clear from context, the inverse system $(G_{\alpha}, p_{\alpha}^{\alpha'})$ may be denoted by $(G_{\alpha})_{\alpha \in \mathcal{A}}$ or $(G_{\alpha})_{\alpha}$ or (G_{α}) .

An inverse system indexed by the natural numbers is an *inverse sequence*. An inverse system indexed by a singleton is a *rudimentary system* and is denoted by (G) , where G is its only term.

Definition 2.2 (morphisms) A *morphism of inverse systems*

$$f: (G_{\alpha})_{\alpha \in \mathcal{A}} \rightarrow (H_{\beta})_{\beta \in \mathcal{B}}$$

consists of an increasing function $\phi: \mathcal{B} \rightarrow \mathcal{A}$ called the *index function* and for each $\beta \in \mathcal{B}$ a homomorphism $f_{\beta}: G_{\phi(\beta)} \rightarrow H_{\beta}$ such that whenever $\beta \leq \beta'$ the following diagram commutes:

$$\begin{array}{ccc} G_{\phi(\beta)} & \longleftarrow & G_{\phi(\beta')} \\ f_{\beta} \downarrow & & \downarrow f_{\beta'} \\ H_{\beta} & \longleftarrow & H_{\beta'} \end{array}$$

For simplicity, we suppress the labels of the bonding maps between terms of the same inverse system. The notion of composition of morphisms is defined in the obvious way.

A *level morphism* between inverse systems indexed by the same set $\mathcal{A}$ is a morphism whose index function $\phi: \mathcal{A} \rightarrow \mathcal{A}$ is the identity function. The *identity morphism* $(G_{\alpha}) \rightarrow (G_{\alpha})$ is the level morphism consisting of the family of identity maps $G_{\alpha} \rightarrow G_{\alpha}$ for each $\alpha \in \mathcal{A}$.

Two morphisms f and f' of inverse systems $(G_{\alpha})_{\alpha \in \mathcal{A}} \rightarrow (H_{\beta})_{\beta \in \mathcal{B}}$ with index functions ϕ and ϕ' are *equivalent* if for each $\beta \in \mathcal{B}$ there exists $\alpha \in \mathcal{A}$ with $\alpha \geq \phi(\beta)$ and $\alpha \geq \phi'(\beta)$ such that the following diagram commutes:

$$\begin{array}{ccc} G_{\phi(\beta)} & \longleftarrow & G_{\alpha} \\ f_{\beta} \downarrow & & \downarrow \\ H_{\beta} & \xleftarrow{f'_{\beta}} & G_{\phi'(\beta)} \end{array}$$

Definition 2.3 (pro-groups) The category of *pro-groups* is defined as follows: the objects are inverse systems of groups, and the morphisms $[f]$ are equivalence classes of morphisms f between inverse systems.

Definition 2.4 A morphism of inverse systems $f: (G_{\alpha}) \rightarrow (H_{\beta})$ is a *pro-isomorphism* if it represents an isomorphism in the category of pro-groups, in other words, if there exists a morphism of inverse systems $g: (H_{\beta}) \rightarrow (G_{\alpha})$ such that fg and gf are each equivalent to identity morphisms. An inverse system is *stable* if it is pro-isomorphic to a rudimentary inverse system.

If $\mathcal{A}'$ is a cofinal subset of the directed set $\mathcal{A}$, then the restriction of an inverse system $(G_\alpha)_{\alpha \in \mathcal{A}}$ to the subsystem indexed by $\mathcal{A}'$ is a pro-isomorphism (see [20], Section I.1.1, Theorem 1).

The *inverse limit* of an inverse system of groups (G_α) is the group

$$\varprojlim (G_\alpha) = \left\{ (g_\alpha) \in \prod_{\alpha \in \mathcal{A}} G_\alpha \mid p_\alpha^{\alpha'}(g_{\alpha'}) = g_\alpha \text{ for all } \alpha \leq \alpha' \right\}.$$

The following well-known result is a basic tool for the study of inverse limits. (See, for example, [20, Section I.5.1].)

Proposition 2.5 *A morphism of inverse systems of groups induces a unique homomorphism of their inverse limits via the universal mapping property of inverse limits. Equivalent morphisms induce the same homomorphism. The assignment $(G_\alpha) \mapsto \varprojlim (G_\alpha)$ is a covariant functor. A pro-isomorphism of inverse systems induces an isomorphism of inverse limits.*

Definition 2.6 *An exact sequence of inverse systems with a common index set $\mathcal{A}$ consists of a sequence of level morphisms*

$$\cdots \rightarrow (A_\alpha) \rightarrow (B_\alpha) \rightarrow (C_\alpha) \rightarrow \cdots$$

such that for each fixed $\alpha \in \mathcal{A}$ the sequence of terms

$$\cdots \rightarrow A_\alpha \rightarrow B_\alpha \rightarrow C_\alpha \rightarrow \cdots$$

is exact.

Dydak [5] proved the following analogue of the five lemma for an exact sequence of inverse sequences.

Lemma 2.7 (five lemma for stability) *Given an exact sequence of inverse sequences (indexed by natural numbers),*

$$(A_m) \rightarrow (B_m) \rightarrow (C_m) \rightarrow (D_m) \rightarrow (E_m),$$

if (A_m) , (B_m) , (D_m) , and (E_m) are stable, then (C_m) is stable.

Question 2.8 Does the conclusion of Lemma 2.7 hold in general for inverse systems over arbitrary index sets?

For the purposes of this article, the conclusion of Lemma 2.7 is sufficient. Indeed most inverse systems arising in geometric topology are pro-isomorphic to inverse sequences. However the authors do not know of any obstruction to a five lemma for stability in the general category of pro-groups. We note that the methods of Dydak's proof do not immediately extend to the general case.

Exactness of inverse systems is not always preserved when passing to the inverse limit [6; 21]. However inverse limits are better behaved if all inverse systems involved are stable, as illustrated by the following result.

Lemma 2.9 (exactness of stable inverse limits) *Given a long exact sequence of stable inverse systems of groups*

$$\cdots \rightarrow (A_\alpha) \rightarrow (B_\alpha) \rightarrow (C_\alpha) \rightarrow \cdots$$

the induced sequence of inverse limits

$$\cdots \rightarrow \varprojlim (A_\alpha) \rightarrow \varprojlim (B_\alpha) \rightarrow \varprojlim (C_\alpha) \rightarrow \cdots$$

is also exact.

Proof Consider the induced commutative diagram in the category of pro-groups

$$\begin{array}{ccccccc} \cdots & \longrightarrow & \varprojlim (A_\alpha) & \longrightarrow & \varprojlim (B_\alpha) & \longrightarrow & \varprojlim (C_\alpha) \longrightarrow \cdots \\ & & \downarrow & & \downarrow & & \downarrow \\ \cdots & \longrightarrow & (A_\alpha) & \longrightarrow & (B_\alpha) & \longrightarrow & (C_\alpha) \longrightarrow \cdots \end{array}$$

in which the groups in the first row are considered to be rudimentary pro-groups, and the vertical maps are canonical projections. Since the inverse systems in the second row are stable, the vertical projections are isomorphisms of pro-groups [20, Section I.5.2, Theorem 2]. Therefore the first row of rudimentary pro-groups is an exact sequence. (See [20, Section II.2.3] for more information about exactness in the nonabelian category of pro-groups.) An exact sequence of rudimentary pro-groups is also exact when considered as a sequence of groups, completing the proof. $\square$

In the setting of inverse sequences (ie systems indexed by the natural numbers) exactness is also preserved under the more general hypothesis that the first derived limit $\lim^1$ vanishes for every term in the given exact sequence [5; 9; 21]. We will not need this more general result here.

3 Coarse geometry

In this section we review concepts in coarse geometry. We refer the reader to [23] for more details and references.

Definition 3.1 A map $f: Y \rightarrow X$ between metric spaces is a *coarse embedding* if there exist unbounded increasing functions $\rho_-, \rho_+: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ such that

$$\rho_-(d(y_1, y_2)) \leq d(f(y_1), f(y_2)) \leq \rho_+(d(y_1, y_2)) \quad \text{for all } y_1, y_2 \in Y.$$

The functions ρ_- and ρ_+ are the *lower control* and *upper control* functions. We note that if Y is a length space, then one can always choose the upper control to be an affine function; ie every coarse embedding of a length space is coarse Lipschitz (see [23, Lemma 1.10]).

Definition 3.2 A *metric simplicial complex* is a connected, locally finite simplicial complex in which each simplex is isometric to the standard Euclidean simplex with edges of length one. A metric simplicial complex X has *bounded geometry* if the vertex links of X have a uniformly bounded number of simplices.

If $K \subset X$ is a subcomplex and R is a positive integer, then $N_R(K)$ is the combinatorial R -tubular neighborhood of K obtained by taking the R -fold iterated closed star of K .

Example 3.3 Suppose a finitely generated group G acts properly by simplicial automorphisms on a bounded geometry metric simplicial complex X . Endow G with the word metric for any finite generating set. Then the orbit map $G \rightarrow X$ given by $g \mapsto g(x_0)$ is a coarse embedding for any choice of basepoint $x_0 \in X$.

Throughout this article all chain complexes and homology groups are considered in the simplicial category and all coefficients are integral. A chain complex $\mathcal{C} = \{C_i, \partial_i\}$ may be considered as a single abelian group equipped with a grading $\mathcal{C} = \bigoplus C_i$. The notations $C_\circ(X) = \{C_i(X), \partial_i\}$ and $C_c^\circ(X) = \{C_c^i(X), \delta^i\}$ denote, respectively, the chain complex of simplicial chains and the cochain complex of compactly supported simplicial cochains on X , each with integer coefficients.

Definition 3.4 If X is a simplicial complex and A is any subset of $C_\circ(X)$, then the *support* of A is the smallest subcomplex K of X so that $A \subset C_\circ(K)$. A homomorphism $h: C_i(X) \rightarrow C_c^j(X)$ has *displacement* at most D for some positive integer D if for every simplex σ of X , the cochain $h(\sigma)$ is supported in $N_D(\sigma)$. A homomorphism $C_c^j(X) \rightarrow C_i(X)$ has *displacement* at most D if for every simplex σ of X , the image of the composition $C_c^j(\bar{\sigma}) \rightarrow C_c^j(X) \rightarrow C_i(X)$ has support lying in $N_D(\sigma)$, where the map $C_c^j(\bar{\sigma}) \rightarrow C_c^j(X)$ is given by extending a cochain on the subcomplex $\bar{\sigma}$ to a cochain on X by setting it to evaluate to zero on simplices of X outside of $\bar{\sigma}$.

Definition 3.5 An n -dimensional metric simplicial complex X is *uniformly k -acyclic* for some $k \leq n$ if for every $R_1 > 0$ there exists $R_2 > 0$ so that for each subcomplex $K \subset X$ of diameter at most R_1 the inclusion $K \rightarrow N_{R_2}(K)$ induces zero on $\tilde{H}_k$. The complex X is *uniformly acyclic* if X is uniformly k -acyclic for all $k \leq n$.

By the Mayer–Vietoris theorem, if X is a union of two subcomplexes A and B such that A , B , and $A \cap B$ are uniformly acyclic, then so is X .

Definition 3.6 Let Y and X be finite dimensional metric simplicial complexes. A chain map

$$h: C_\circ(Y) \rightarrow C_\circ(X)$$

or a chain homotopy $h: C_\circ(Y) \rightarrow C_{\circ+1}(X)$ is *D -controlled* for a constant $D < \infty$ if for each simplex σ of Y , the support of $h(C_\circ(\sigma))$ has diameter at most D . A chain map h is a *coarse embedding* if h is D -controlled for some D and there exists a proper function $\phi: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that for each subcomplex $K \subset Y$ of diameter at least r , the support of $h(C_\circ(K))$ has diameter at least $\phi(r)$. Similar notions for cochain maps and cochain homotopies are defined analogously.

Remark 3.7 Suppose Z' is a subcomplex of Z such that the inclusion $Z' \hookrightarrow Z$ is a coarse embedding. It follows immediately from the definitions that the induced chain map of simplicial chain complexes $C_\circ(Z') \rightarrow C_\circ(Z)$ is a coarse embedding.

Coarse embeddings of spaces do not need to be continuous. However, if the spaces are finite dimensional uniformly acyclic metric simplicial complexes, then a coarse embedding induces a chain map that is also a coarse embedding in the following sense.

Definition 3.8 Given a coarse embedding of metric simplicial complexes $f: Z \rightarrow X$, a chain map $f_{\#}: C_{\circ}(Z) \rightarrow C_{\circ}(X)$ of simplicial chain complexes M -approximates f for some number M if for each simplex σ^i of Z , the support of the chain $f_{\#}(\sigma^i)$ is within Hausdorff distance M of the image $f(\sigma^i)$. We say that $f_{\#}$ approximates f if it M -approximates for some M . Note that an approximating chain map $f_{\#}$ of a coarse embedding is again a coarse embedding.

Lemma 3.9 Assume Z and X are metric simplicial complexes such that X is uniformly acyclic and Z is finite dimensional. For each coarse embedding $f: Z \rightarrow X$ there exists a chain map $f_{\#}: C_{\circ}(Z) \rightarrow C_{\circ}(X)$ approximating f .

Although $f_{\#}$ is not uniquely determined by f , it is well-defined in a weaker sense: Let $f: Z \rightarrow X$ be a coarse embedding, and let $f_{\#}$ and $f'_{\#}$ be two chain maps that M' -approximate f . There exists a constant $D = D(Z, X, M', f)$ such that $f_{\#}$ and $f'_{\#}$ are chain homotopic by a D -controlled chain homotopy.

Proof We construct $f_{\#}$ on $C_{\circ}(Z) = \bigoplus_{i \geq 0} C_i(Z)$ by induction on the dimension i . In dimension 0, define $f_{\#}$ by arbitrarily choosing for each 0-simplex σ^0 of Z a 0-simplex of X within a distance 1 of the point $f(\sigma^0)$. In dimension 1, define $f_{\#}$ using the upper control of f (the boundary of a 1-simplex has image in X with a uniformly bounded diameter) and the uniform acyclicity of X (a pair of points separated by a uniformly bounded distance is the boundary of a 1-chain with uniformly bounded diameter). In dimension $k > 1$, define $f_{\#}$ inductively, since the boundary of a k -simplex is a uniformly bounded diameter $(k-1)$ -cycle in X by hypothesis, so it is the boundary of a uniformly bounded diameter k -chain by uniformly acyclic. The uniform bounds obtained in this inductive argument do depend on the dimension k , but by hypothesis Z is finite dimensional, so there exists an upper bound independent of dimension.

To see that the chain map obtained is an approximation in the desired sense, first note that each simplex σ^k of Z has diameter 1, so its image $f(\sigma^k)$ has uniformly bounded diameter in X . In particular, the image $f(\sigma^k)$ lies in a uniformly bounded neighborhood of the image of the 0-skeleton of σ^k , which lies uniformly close to the support of the chain $f_{\#}(\sigma^k)$.

The construction of the controlled chain homotopy involves a similar induction on dimension. The base of the induction relies on the fact that any two approximating chain maps are uniformly close when restricted to the 0-skeleton because they are both uniformly close to f . The induction follows as before using that X is uniformly acyclic and Z is finite dimensional. $\square$

The previous lemma implies that coarse embeddings of metric simplicial complexes (as above) induce well-defined maps on homology and on cohomology with compact supports as explained in the following corollary.

Corollary 3.10 *Let $f: Z \rightarrow X$ be a coarse embedding of metric simplicial complexes, where Z is finite dimensional and X is uniformly acyclic. Any two chain maps approximating f induce the same homomorphism in homology and also in cohomology with compact supports (with integer coefficients).*

Proof The claim regarding homology is immediate since two chain homotopic maps of chain complexes induce the same map on homology.

Simplicial cohomology with compact supports makes sense only when the spaces involved are locally finite. Indeed local finiteness implies that the dual of the boundary map takes compactly supported cochains to compactly supported cochains. Because an approximating chain map $f_{\#}: C_{\circ}(Z) \rightarrow C_{\circ}(X)$ is a coarse embedding, its dual $f^{\#}$ also takes compactly supported cochains to compactly supported cochains. Since the dual of a chain map is a cochain map, each approximating chain map $f_{\#}$ induces a homomorphism on cohomology with compact supports. Since the dual of a chain homotopy is a cochain homotopy, any two approximating chain maps $f_{\#}$ and $g_{\#}$ induce the same maps on cohomology with compact supports. $\square$

4 Čech homology of filtered ends

The study of ends of a space X largely involves properties of the inverse system of subspaces $X - K$ indexed by the family of bounded subsets $K \subset X$. More generally, if K is an unbounded subspace, one can study the analogous filtration of X by metric neighborhoods of K . For each subcomplex K of a metric simplicial complex X , the family of simplicial neighborhoods $N_R(K)$ of K is a filtration of X in the sense of Geoghegan [9, Chapter 14]. This inverse sequence is unchanged up to pro-isomorphism if one replaces K with a subset K' at finite Hausdorff distance from K . In order to see this equivalence it is more convenient to work with a larger inverse system, as described below. We note that the following definition may be considered as a special case of the general notion of homology of filtered ends of spaces discussed by Geoghegan.

Definition 4.1 (Čech homology of filtered ends) Let X be a metric simplicial complex. For each subset K of X , a *neighborhood* L of K is a subcomplex of X such that L is contained in a metric tubular neighborhood $N_R(K)$ of K for some $R < \infty$. The family $\mathcal{N}(K)$ of all neighborhoods of K is directed by inclusion: $L \leq L'$ if $L \subseteq L'$. The i^{th} *homology pro-group of the K -end* is a pro-group associated to the end of the filtration $\mathcal{N}(K)$ and is defined by the inverse system

$$\text{pro-}H_i^D(K^c) = (H_i(\overline{X - L}))$$

indexed by $L \in \mathcal{N}(K)$. Taking inverse limits we get the i^{th} Čech homology of the K -end,

$$\check{H}_i^D(K^c) = \varprojlim_L (H_i(\overline{X - L})).$$

Notice that by definition $\text{pro-}H_i^D(K_1^c) = \text{pro-}H_i^D(K_2^c)$ for any subsets K_1 and K_2 at a finite Hausdorff distance from each other, since $\mathcal{N}(K_1) = \mathcal{N}(K_2)$. In the special case that K is a bounded subset, the i^{th}

homology pro-group of the K -end is also known as the i^{th} homology pro-group at infinity of X and does not depend on the choice of bounded set K . Its inverse limit is the i^{th} Čech homology group at infinity of X .

If K is the empty set, then by definition $\mathcal{N}(K)$ is also empty. In this degenerate case, we note that $\text{pro-}\check{H}_i^D(\emptyset^c)$ reduces to $H_i(X)$.

Under certain conditions, a map of pairs $(A, B) \rightarrow (X, Y)$ induces a corresponding map $\check{H}_i^D(B^c) \rightarrow \check{H}_i^D(Y^c)$ on Čech homology of filtered ends. Rather than developing the most general case, we focus on the following simple setup. Consider a coarse embedding of metric simplicial complexes $f: A \rightarrow X$, and let B be a subcomplex of A with image $f(B)$. Roughly speaking, any such f induces a homomorphism from Čech homology of the B -end of A to the Čech homology of the $f(B)$ -end of X , provided that we first choose a chain map that approximates f .

Proposition 4.2 *Let $f: A \rightarrow X$ be a coarse embedding of metric simplicial complexes, where A is finite dimensional and X is uniformly acyclic. Let B be any subcomplex of A . Then there exists a unique homomorphism*

$$\check{H}_i^D(B^c) \xrightarrow{\check{f}} \check{H}_i^D((fB)^c)$$

induced by f in the following sense. Every chain map

$$C_\circ(A) \xrightarrow{f_\#} C_\circ(X)$$

approximating f induces (by restriction to subspaces) a morphism between the inverse systems defining the Čech homology groups. This morphism induces a homomorphism $\check{f}$ of inverse limits that does not depend on any of the choices made in its construction.

Proof We first show that any approximating chain map $f_\#$ induces a map on deep homology via restrictions. Suppose $f_\#$ is a chain map approximating f , and let K be the support of $f_\#(C_\circ(B))$. By construction, the Hausdorff distance between K and $f(B)$ is finite, for each simplex σ^k of A the support of the chain $f_\#(\sigma^k)$ is within a uniformly bounded distance of the image $f(\sigma^k)$, and f is a coarse embedding. Therefore for each R there is a constant M such that whenever $L \subseteq N_R(f(B))$ we have

$$(4.3) \quad f_\#(C_\circ(\overline{A - N_M(B)})) \subseteq C_\circ(\overline{X - N_R(f(B))}) \subseteq C_\circ(\overline{X - L}).$$

In particular $f_\#$ induces, via restriction, a homomorphism

$$H_i(\overline{A - N_M(B)}) \xrightarrow{f_L} H_i(\overline{X - L}).$$

For concreteness, for each L we choose $R = R(L)$ and $M = M(L)$ to be the minimal natural numbers satisfying (4.3), so that the mapping $L \mapsto M(L)$ is nondecreasing. Since at the level of chain maps the family of such restrictions commutes with inclusions, the maps f_L define a morphism of inverse systems, which induces a corresponding homomorphism $\check{f}: \check{H}_i^D(B^c) \rightarrow \check{H}_i^D(K^c)$.

Now suppose $f_{\#}$ and $f'_{\#}$ are two chain maps that approximate f . We will show that the induced maps $\check{f}$ and $\check{f}'$ constructed as above are equal. Recall that by [Lemma 3.9](#) the chain maps $f_{\#}$ and $f'_{\#}$ are chain homotopic by a D -controlled chain homotopy F for some D depending on $f_{\#}$ and $f'_{\#}$. Since $f_{\#}$ is a coarse embedding, the chain homotopy F is also a coarse embedding. Suppose $L \in \mathcal{N}(f(B))$, and choose M' sufficiently large that the chain homotopy F restricts to a map

$$C_o(\overline{A - N_{M'}(B)}) \rightarrow C_{o+1}(\overline{X - L}).$$

Then the two homology maps

$$H_i(\overline{A - N_{M'}(B)}) \rightarrow H_i(\overline{X - L})$$

induced by restriction of $f_{\#}$ and $f'_{\#}$ are equal. It follows that the induced maps $\check{H}_i^D(B^c) \rightarrow \check{H}_i^D(K^c)$ are equal as well. $\square$

Remark 4.4 Throughout this section we have discussed the ordinary Čech homology of filtered ends. The analogous reduced Čech homology groups are defined similarly. We omit the details, which are essentially the same as in the ordinary case.

4.1 Relative homology of filtered ends

This section discusses a relative version of Čech homology of filtered ends. In general Čech homology theories do not satisfy the Eilenberg–Steenrod exactness axiom. However in certain well-behaved situations exactness still holds. One such natural situation is when X is a coarse Poincaré duality space, as discussed in the following definition.

Definition 4.5 A coarse PD(n) space X is a bounded geometry, uniformly acyclic, metric simplicial complex equipped with cochain maps

$$C_{n-o}(X) \xrightarrow{\bar{P}} C_c^o(X) \xrightarrow{P} C_{n-o}(X)$$

such that for some integer constant E_0 we have

- (1) P and $\bar{P}$ have displacement at most E_0 ,
- (2) $\bar{P} \circ P$ and $P \circ \bar{P}$ are cochain homotopic to the identity by E_0 -controlled cochain homotopies.

Definition 4.6 (relative Čech homology) Let X be a metric simplicial complex and let $K_1 \subset K_0$ be subsets of X . A *neighborhood* of the pair (K_0, K_1) is a pair of subcomplexes (L_0, L_1) of X with $L_1 \subseteq L_0$ such that there exists R with $L_i \subseteq N_R(K_i)$. As in the absolute case, the family $\mathcal{N}(K_0, K_1)$ of all neighborhoods of (K_0, K_1) is directed by inclusion, and we define the i^{th} homology pro-group of the (K_0, K_1) -end,

$$\text{pro-}H_i^D(K_1^c, K_0^c) = (H_i(\overline{X - L_1}, \overline{X - L_0}))$$

indexed by $(L_0, L_1) \in \mathcal{N}(K_0, K_1)$. Taking inverse limits we get the corresponding i^{th} Čech homology group of the (K_0, K_1) -end,

$$\check{H}_i^D(K_1^c, K_0^c) = \varprojlim (H_i(\overline{X - L_1}, \overline{X - L_0})).$$

Suppose X is held fixed. An inclusion $(C, D) \rightarrow (A, B)$ of pairs of subsets induces a morphism of homology pro-groups:

$$\text{pro-}H_i^D(B^c, A^c) \rightarrow \text{pro-}H_i^D(D^c, C^c).$$

Thus the relative homology pro-group of filtered ends may be considered as a contravariant functor on the pairs of subsets of X .

When X is a coarse PD(n) space, one may identify the homology pro-group associated to the pair $(K, \emptyset)$ with the reduced homology pro-group associated to K as defined in [Definition 4.1](#). Indeed a neighborhood of $(K, \emptyset)$ has the form $(L, \emptyset)$ for some neighborhood L of K . Since X is acyclic, the connecting homomorphism ∂ of the reduced homology long exact sequence of the pair $(X, X - L)$ is an isomorphism in every dimension:

$$H_i(X, \overline{X - L}) \xrightarrow{\partial} \tilde{H}_{i-1}(\overline{X - L}).$$

4.2 Stable deep components

We describe what the 0^{th} Čech homology of the K -end represents in terms of a natural basis that we use later.

Definition 4.7 Let X be a metric simplicial complex, and let K be a subset of X . The set of *ends of the pair* (X, K) or *relative ends* is the set

$$\mathcal{E}(X, K) = \varprojlim \pi_0(\overline{X - L}).$$

indexed by the directed system of all subcomplexes $L \in \mathcal{N}(K)$.

A component $U \in \pi_0(\overline{X - L})$ is *deep* if it is in the image of the natural projection $\mathcal{E}(X, K) \rightarrow \pi_0(\overline{X - L})$, or equivalently if U is not contained in a bounded neighborhood of K . A deep component of $\overline{X - L}$ is *stable* if it intersects exactly one deep component of $\overline{X - L'}$ for each $L' \in \mathcal{N}(K)$ with $L \subseteq L'$.

Note that if U is a stable deep component of $\overline{X - L}$, then there exists a unique end of the pair (X, K) mapping to U under the natural projection.

The relation between stable deep components and the 0^{th} Čech homology of the K -end is summarized in the following result.

Proposition 4.8 [9, Section 14.3] Suppose the inverse system $\{H_0(\overline{X-L})\}_{L \in \mathcal{N}(K)}$ is stable and its inverse limit is free abelian of finite rank. Then the set of relative ends $\mathcal{E}(X, K)$ is finite and freely generates $\check{H}_0^D(K^c)$. Furthermore there exists $R_0 < \infty$ such that whenever $L \in \mathcal{N}(K)$ satisfies $N_{R_0}(K) \subseteq L$ all deep components $U \in \pi_0(\overline{X-L})$ are stable, and the natural projection

$$\mathcal{E}(X, K) \rightarrow \pi_0(\overline{X-L})$$

is a bijection onto the set of deep components of $\overline{X-L}$.

In addition, for each $L \in \mathcal{N}(K)$ there exists $L' \in \mathcal{N}(K)$ such that any component $U \in \pi_0(\overline{X-L})$ that is not deep is contained in L' .

5 Coarse Alexander duality

In this section, we review the coarse Alexander duality isomorphism theorem of Kapovich and Kleiner [17].

Suppose X is a coarse PD(n) space and Z is a uniformly acyclic metric simplicial complex with bounded geometry. Let $f: Z \rightarrow X$ be a coarse embedding. For each k , coarse Alexander duality is a pro-group morphism

$$\text{pro-}\tilde{H}_{n-k-1}^D(fZ^c) \xrightarrow{A} H_c^k(Z),$$

defined as follows.

Although by definition the homology pro-group is indexed by the family of subcomplexes $\mathcal{N}(fZ)$, it suffices to restrict our attention to the cofinal sequence of subcomplexes $\{N_R(K) \mid R = 1, 2, 3, \dots\}$, where K is any fixed subcomplex at a finite Hausdorff distance from the set fZ . Applying this restriction, we may consider the homology pro-group to be an inverse sequence indexed by the natural number R .

Fix a subcomplex $K \in \mathcal{N}(fZ)$ at a finite Hausdorff distance from fZ , and let

$$Y^R = \overline{X - N_R(K)}.$$

Because the cochain map $\bar{P}$ given by Definition 4.5 has bounded displacement, it restricts to a cochain map

$$C_{n-\circ}(Y^{R+E}) \xrightarrow{\bar{P}} C_c^\circ(X, N_R(K)),$$

where $E = E_0 + 1$ and E_0 is the constant from Definition 4.5. Indeed $\bar{P}$ maps each chain in Y^{R+E} to a cochain supported in $N_{E_0}(Y^{R+E}) = N_{E_0}(Y^{R+1+E_0})$, which lies in $\overline{X - N_{R+1}(K)}$. Such a cochain evaluates to zero on every simplex of $N_R(K)$, so it lies in the given relative cochain complex. Therefore $\bar{P}$ induces a cochain map, also denoted by $\bar{P}$, between the two quotient cochain complexes in the following commutative diagram of cochain maps:

$$\begin{array}{ccccccc} 0 & \longrightarrow & C_{n-\circ}(Y^{R+E}) & \longrightarrow & C_{n-\circ}(X) & \longrightarrow & C_{n-\circ}(X, Y^{R+E}) \longrightarrow 0 \\ & & \downarrow \bar{P} & & \downarrow \bar{P} & & \downarrow \bar{P} \\ 0 & \longrightarrow & C_c^\circ(X, N_R(K)) & \longrightarrow & C_c^\circ(X) & \longrightarrow & C_c^\circ(N_R(K)) \longrightarrow 0 \end{array}$$

Choose an approximating chain map $f_{\#}: C_{\circ}(Z) \rightarrow C_{\circ}(X)$. We now add the assumption that K has been chosen large enough so as to contain the support of $f_{\#}(C_{\circ}(Z))$. This assumption is sufficient for defining a morphism of pro-groups, since the collection of sets $\{N_R(K)\}$ is cofinal in $\mathcal{N}(fZ)$. Thus we may replace K with its R_0 -neighborhood for a large enough value of R_0 . In particular, the chain map $f_{\#}$ restricts to a chain map

$$C_{\circ}(Z) \xrightarrow{f_{\#}} C_{\circ}(N_R(K)).$$

As discussed in [Corollary 3.10](#), a chain map that is a coarse embedding induces a dual cochain map of compactly supported cochain complexes. In the present setting, we get an induced cochain map

$$C_c^{\circ}(N_R(K)) \xrightarrow{f^{\#}} C_c^{\circ}(Z).$$

The connecting homomorphism ∂ of the reduced homology long exact sequence of the pair (X, Y^{R+E}) is an isomorphism in every dimension because X is acyclic. Consider, for each fixed k , the map $A^{R+E}: \tilde{H}_{n-k-1}(Y^{R+E}) \rightarrow H_c^k(Z)$ defined by the composition of homomorphisms

$$\tilde{H}_{n-k-1}(Y^{R+E}) \xrightarrow{\partial^{-1}} H_{n-k}(X, Y^{R+E}) \xrightarrow{\bar{P}} H_c^k(N_R(K)) \xrightarrow{f^*} H_c^k(Z),$$

in which ∂^{-1} denotes the inverse of the connecting homomorphism ∂ , and $\bar{P}$ and f^* are the induced maps on homology/cohomology discussed above. When $R < R'$ consider the following diagram in which the vertical maps are induced by inclusion:

$$\begin{array}{ccccccc} \tilde{H}_{n-k-1}(Y^{R'+E}) & \xrightarrow{\partial^{-1}} & H_{n-k}(X, Y^{R'+E}) & \xrightarrow{\bar{P}} & H_c^k(N_{R'}(K)) & \xrightarrow{f^*} & H_c^k(Z) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \text{id} \\ \tilde{H}_{n-k-1}(Y^{R+E}) & \xrightarrow{\partial^{-1}} & H_{n-k}(X, Y^{R+E}) & \xrightarrow{\bar{P}} & H_c^k(N_R(K)) & \xrightarrow{f^*} & H_c^k(Z) \end{array}$$

Observe that the diagram above commutes. Indeed, the leftmost square commutes by naturality of the connecting homomorphism ∂ . The middle square commutes at the level of cochain complexes, since the two horizontal maps labeled $\bar{P}$ are induced by the same map $\bar{P}: C_{n-\circ}(X) \rightarrow C_c^{\circ}(X)$. Similarly the right-hand square commutes at the level of cochains by the definition of $f^{\#}$.

Recall that a morphism of inverse systems with rudimentary target is determined by a single homomorphism of groups. In particular, for each $R > E$ the map A^R determines a morphism

$$(\tilde{H}_{n-k-1}(Y^R))_R \xrightarrow{A^R} (H_c^k(Z))$$

of inverse systems. The argument in the previous paragraph implies that the morphisms A^R and $A^{R'}$ are equivalent for all $R, R' > E$. The equivalence class $A = [A^R]$ is a morphism of pro-groups, known as *coarse Alexander duality*.

Kapovich and Kleiner have shown the following coarse Alexander duality isomorphism theorem.

Theorem 5.1 [17, Theorem 7.7] Suppose Z is a uniformly acyclic metric simplicial complex with bounded geometry, and suppose X is a coarse PD(n) space. Let $f: Z \rightarrow X$ be a coarse embedding. Then the coarse Alexander duality morphism

$$\text{pro-}\tilde{H}_{n-k-1}^D(fZ^c)_R \xrightarrow{A} H_c^k(Z)$$

is a pro-isomorphism. In particular, the reduced homology pro-groups of the (fZ) -end of X are stable, and A induces an isomorphism

$$\check{H}_{n-k-1}^D(fZ^c) \xrightarrow{A} H_c^k(Z)$$

called the coarse Alexander duality isomorphism.

6 Coarse Alexander duality for pairs

The ordinary coarse Alexander duality isomorphism theorem deals with a single coarse embedding $Z \rightarrow X$. In this section, we establish coarse Alexander duality for pairs. More specifically we introduce a duality theorem for a coarse embedding of a pair of complexes $Z_0 \subseteq Z_1$ in a coarse PD(n) space X . In a general context, Čech homology theories do not satisfy the Eilenberg–Steenrod axioms since they do not satisfy the exactness axiom for pairs. However in the setting of coarse PD(n) spaces, coarse Alexander duality implies that the inverse system defining the homology pro-group of the filtered end is stable. Thus in this setting we establish the exactness of the long exact sequence of a pair. The main objective of this section is to explain the details of this exactness statement.

More precisely, the coarse Alexander duality isomorphism for pairs is proved using the long exact sequence of a pair in the setting of Čech homology of relative ends, which we show is exact and satisfies a naturality property with respect to the coarse Alexander duality isomorphism.

Suppose X is a coarse PD(n) space, and $Z_1 \subset Z_0$ is a pair of uniformly acyclic complexes with bounded geometry such that the inclusion $Z_1 \hookrightarrow Z_0$ is a coarse embedding. If $f: Z_0 \rightarrow X$ is a coarse embedding, then it restricts to a coarse embedding $f: Z_1 \rightarrow X$. In this context we introduce a relative version of Čech homology associated to the complement of a pair of subcomplexes, defined as follows. The coarse embedding f induces an approximating chain map $f_\#: C_\circ(Z_0) \rightarrow C_\circ(X)$ by Lemma 3.9. We let K_i denote the subcomplex of X that is the support of $f_\#(C_\circ(Z_i))$ for each i . The relative Čech homology of the (Z_0, Z_1) -end is given by

$$\check{H}_i^D(fZ_1^c, fZ_0^c) = \varprojlim_R H_i(Y_1^R, Y_0^R)$$

where $Y_i^R = \overline{X - N_R(K_i)}$.

As in the definition of coarse Alexander duality, the cochain maps

$$\bar{P}: C_{n-\circ}(X) \rightarrow C_c^\circ(X) \quad \text{and} \quad f^\#: C_c^\circ(X) \rightarrow C_c^\circ(Z_0)$$

induce cochain maps that make the following diagram commute:

$$\begin{array}{ccccccc}
 0 & \longrightarrow & C_{n-\circ}(Y_1^{R+E}, Y_0^{R+E}) & \longrightarrow & C_{n-\circ}(X, Y_0^{R+E}) & \longrightarrow & C_{n-\circ}(X, Y_1^{R+E}) \longrightarrow 0 \\
 & & \downarrow \bar{P} & & \downarrow \bar{P} & & \downarrow \bar{P} \\
 0 & \longrightarrow & C_c^\circ(N_R(K_0), N_R(K_1)) & \longrightarrow & C_c^\circ(N_R(K_0)) & \longrightarrow & C_c^\circ(N_R(K_1)) \longrightarrow 0 \\
 & & \downarrow f^\# & & \downarrow f^\# & & \downarrow f^\# \\
 0 & \longrightarrow & C_c^\circ(Z_0, Z_1) & \longrightarrow & C_c^\circ(Z_0) & \longrightarrow & C_c^\circ(Z_1) \longrightarrow 0
 \end{array}$$

We define the *coarse Alexander duality* morphism of pairs to be the equivalence class $A = [A^{R+E}]$ consisting of the homology maps induced by the leftmost column of the above diagram:

$$(H_{n-k}(Y_1^R, Y_0^R))_R \xrightarrow{A} H_c^k(Z_0, Z_1)$$

The maps A^R represent a well-defined morphism of pro-groups by similar reasoning to that used in [Section 5](#) for the original coarse Alexander duality morphism.

The definitions of the coarse Alexander duality morphism for a single coarse embedding and for a pair of coarse embeddings are related by the following commutative diagram of homology groups:

$$\begin{array}{ccccccc}
 & & & \tilde{H}_{n-k-1}(Y_0^R) & \longrightarrow & \tilde{H}_{n-k-1}(Y_1^R) & \\
 & \nearrow & & \uparrow \partial & & \uparrow \partial & \searrow \\
 \cdots & \xrightarrow{\partial} & H_{n-k}(Y_1^R, Y_0^R) & \longrightarrow & H_{n-k}(X, Y_0^R) & \longrightarrow & H_{n-k}(X, Y_1^R) \xrightarrow{\partial} \cdots \\
 & & \downarrow & & \downarrow & & \downarrow \\
 \cdots & \xrightarrow{\delta} & H_c^k(Z_0, Z_1) & \longrightarrow & H_c^k(Z_0) & \longrightarrow & H_c^k(Z_1) \xrightarrow{\delta} \cdots
 \end{array}$$

The lower half of this diagram of homology maps is induced by the preceding commutative diagram of cochain maps. The wavy top row of the diagram consists of the reduced homology long exact sequence of the pair (Y_1^R, Y_0^R) . Each map labeled ∂ in the diagram is the connecting homomorphism of a homology long exact sequence. One can easily verify that this homology diagram commutes. Indeed the lower half commutes due to the well-known fact that a morphism between short exact sequences of chain complexes always induces a morphism between the corresponding homology long exact sequences. (See, for instance, [24, Theorem 4.5.4].) Similarly each region in the upper half of the diagram commutes by naturality of the connecting homomorphism ∂ with respect to inclusion maps.

We now condense the above homology diagram, showing only the top and bottom rows. Observe that the vertical maps in the condensed diagram below are, by definition, equal to the maps in the coarse

Alexander duality morphism:

$$(6.1) \quad \begin{array}{ccccccc} \cdots & \longrightarrow & H_{n-k}(Y_1^R, Y_0^R) & \xrightarrow{\partial} & \tilde{H}_{n-k-1}(Y_0^R) & \longrightarrow & \tilde{H}_{n-k-1}(Y_1^R) \longrightarrow \cdots \\ & & \downarrow A_R & & \downarrow A_R & & \downarrow A_R \\ \cdots & \xrightarrow{\delta} & H_c^k(Z_0, Z_1) & \longrightarrow & H_c^k(Z_0) & \longrightarrow & H_c^k(Z_1) \xrightarrow{\delta} \cdots \end{array}$$

Theorem 6.2 (coarse Alexander duality for pairs) *There exists a commutative diagram*

$$\begin{array}{ccccccc} \cdots & \longrightarrow & \check{H}_{n-k}^D(fZ_1^c, fZ_0^c) & \xrightarrow{\partial} & \check{H}_{n-k-1}^D(fZ_0^c) & \longrightarrow & \check{H}_{n-k-1}^D(fZ_1^c) \longrightarrow \cdots \\ & & \cong \downarrow A & & \cong \downarrow A & & \cong \downarrow A \\ \cdots & \xrightarrow{\delta} & H_c^k(Z_0, Z_1) & \longrightarrow & H_c^k(Z_0) & \longrightarrow & H_c^k(Z_1) \xrightarrow{\delta} \cdots \end{array}$$

in which the first row is an exact sequence induced by the long exact sequence of the pair (Y_1^R, Y_0^R) , the second row is a long exact sequence of the pair (Z_0, Z_1) , and the vertical maps are coarse Alexander duality isomorphisms.

Proof Passing to the inverse limit with respect to R in (6.1) gives the desired commutative diagram. By coarse Alexander duality, the second and third columns of (6.1) are pro-isomorphisms. Hence the second and third vertical maps in the limiting diagram are isomorphisms. If we knew that the limiting diagram had exact rows, we would be done by the traditional five lemma. Since the second row is known to be an exact sequence, we only need to show that the top row of the limiting diagram is an exact sequence.

The inverse sequences in the second and third columns of (6.1) are stable since each is pro-isomorphic to a rudimentary inverse system. By Lemma 2.7 (the five lemma for stability), the inverse sequence in the first column is stable as well. Since the first row of (6.1) is an exact sequence of stable inverse systems, its inverse limit is again exact by Lemma 2.9. $\square$

We now describe a setting in which inclusions of pairs induce homomorphisms between the corresponding Čech homology groups of relative ends.

Definition 6.3 A *coarse inclusion* is an inclusion of simplicial complexes $B_0 \subseteq A_0$ such that the inclusion map is a coarse embedding. A *coarse inclusion of pairs* is an inclusion of simplicial pairs $(B_0, B_1) \rightarrow (A_0, A_1)$ such that each of the inclusions

$$\begin{array}{ccc} B_1 & \longrightarrow & B_0 \\ \downarrow & & \downarrow \\ A_1 & \longrightarrow & A_0 \end{array}$$

is a coarse embedding.

Proposition 6.4 Suppose $(B_0, B_1) \rightarrow (A_0, A_1)$ is a coarse inclusion of pairs such that A_0, A_1, B_0, B_1 are uniformly acyclic with bounded geometry, and suppose $f: A_0 \rightarrow X$ is a coarse embedding into a coarse PD(n) space X .

(1) There exists a commutative diagram

$$\begin{array}{ccc} \check{H}_i^D(fA_1^c, fA_0^c) & \longrightarrow & \check{H}_i^D(fB_1^c, fB_0^c) \\ \cong \downarrow A & & \cong \downarrow A \\ H_c^{n-i}(A_0, A_1) & \longrightarrow & H_c^{n-i}(B_0, B_1) \end{array}$$

in which the vertical maps are Alexander duality isomorphisms and the horizontal maps are induced by inclusions of spaces.

(2) (naturality) The connecting homomorphism in the Čech homology long exact sequence is natural with respect to coarse inclusions, in the sense that the following diagram commutes:

$$\begin{array}{ccc} \check{H}_i^D(fA_1^c, fA_0^c) & \xrightarrow{\partial} & \check{H}_{i-1}^D(fA_0^c) \\ \downarrow & & \downarrow \\ \check{H}_i^D(fB_1^c, fB_0^c) & \xrightarrow{\partial} & \check{H}_{i-1}^D(fB_0^c) \end{array}$$

(3) (excision) Suppose $(B_0, B_1) = (\overline{A_0 - U}, \overline{A_1 - U})$ for some subcomplex U of A_0 contained in the interior of A_1 . Then the map induced by inclusion

$$\check{H}_i^D(fA_1^c, fA_0^c) \rightarrow \check{H}_i^D(fB_1^c, fB_0^c)$$

is an isomorphism.

Proof As in the definition of coarse Alexander duality for pairs, the cochain maps

$$\bar{P}: C_{n-\circ}(X) \rightarrow C_c^\circ(X) \quad \text{and} \quad f^\#: C_c^\circ(X) \rightarrow C_c^\circ(A_0)$$

induce morphisms that make the following diagram commute:

$$(6.5) \quad \begin{array}{ccc} \text{pro-}H_i^D(fA_1^c, fA_0^c) & \longrightarrow & \text{pro-}H_i^D(fB_1^c, fB_0^c) \\ \downarrow A & & \downarrow A \\ H_c^{n-i}(A_0, A_1) & \longrightarrow & H_c^{n-i}(B_0, B_1) \end{array}$$

We obtain (1) by passing to the inverse limit in the above diagram.

The diagram

$$\begin{array}{ccc} \text{pro-}H_i^D(fA_1^c, fA_0^c) & \xrightarrow{\partial} & \text{pro-}\tilde{H}_{i-1}^D(fA_0^c) \\ \downarrow & & \downarrow \\ \text{pro-}H_i^D(fB_1^c, fB_0^c) & \xrightarrow{\partial} & \text{pro-}\tilde{H}_{i-1}^D(fB_0^c) \end{array}$$

commutes by naturality of the connecting homomorphism in simplicial homology. Again, (2) follows by passing to the inverse limit.

Conclusion (3) follows from (1) as follows. In (6.5), the horizontal map in the second row is an isomorphism by excision for simplicial cohomology with compact supports. We conclude that the horizontal map in the first row of (6.5) is a pro-isomorphism, since the other three maps in (6.5) are pro-isomorphisms. $\square$

7 The Jordan cycle and adjacency graph

Let X be a coarse $\text{PD}(n)$ space. In this section, we examine the structure of a coarse embedding of a union of k halfspaces of dimension $n - 1$ into X . We show that such a union coarsely separates X into k deep components that are arranged in a cyclic pattern. The main tools used to understand this structure are the Jordan cycle and an associated dual graph to the system of complementary deep components. These objects are defined and studied below.

7.1 A topological adjacency graph

In order to motivate some of the machinery used in the coarse setting, we begin by examining a more classical situation in the topological category.

The classical Jordan–Brouwer separation theorem states that any subspace T of S^n homeomorphic to S^{n-1} separates S^n into exactly two complementary components, each of which has T as its frontier (see, for instance, [24, Theorem 4.7.15]). In this subsection, we extend this result to study embeddings of T_k , the space formed by gluing k copies of the $(n-1)$ -disc $D_1, \dots, D_k$ along their boundaries via homeomorphisms $S^{n-2} \rightarrow S^{n-2}$. In other words, T_k is homeomorphic to the join of S^{n-2} and a discrete set of k points. We fix an embedding $T_k \rightarrow S^n$, and identify T_k with its image.

We will see that T_k separates S^n into k complementary components that are arranged in a natural cyclic order about their common boundary sphere $\partial D_i \cong S^{n-2}$.

Definition 7.1 (adjacency graph) By Alexander duality, the complement of T_k in S^n has k components, each of which is path connected. Omitting one of the discs D_i from T_k yields a subspace T_{k-1}^i whose complement in S^n has $k - 1$ components. The inclusion of $S^n - T_k$ into $S^n - T_{k-1}^i$ induces a surjective map of components that is one-to-one except for a pair of components A_i and B_i of $S^n - T_k$ that lie in the same component of $S^n - T_{k-1}^i$. In other words, omission of the disc D_i from T_k has the effect of merging the two components A_i and B_i so that $A_i \cup B_i \cup \text{int}(D_i)$ is a component of $S^n - T_{k-1}^i$. The *adjacency graph* of complementary components of T_k is the graph with one vertex for each component of $S^n - T_k$ and one edge for each D_i representing the corresponding adjacency of components described above.

In fact, we get an equivalent set of $k - 1$ complementary components if we remove from D_i any open $(n-1)$ -ball U having diameter less than ϵ such that $D_i - U$ is homeomorphic to $I \times S^{n-2}$. Since $A_i \cup B_i \cup U$ is path connected, it contains a path from a point of A_i to a point of B_i that intersects U . Thus U intersects each of $\bar{A}_i$ and $\bar{B}_i$. Since $\epsilon > 0$ is arbitrary, we see that each point of D_i lies in $\bar{A}_i \cap \bar{B}_i$. On the other hand, the argument of the preceding paragraph shows that $\text{int}(D_i)$ does not meet the closure of any component other than A_i and B_i .

If the subspace $T_k \rightarrow S^n$ were piecewise linear with respect to standard PL structures, we might interpret the adjacency graph in terms of a pair of dual cell structures on S^n , but this interpretation is difficult to make precise for wild topological embeddings.

Definition 7.2 The *circuit* Circ_k is the graph with vertex set $\mathbb{Z}/k\mathbb{Z}$ and k edges $e_1, \dots, e_k$ such that e_i is adjacent to the vertices i and $i + 1$ (modulo k).

Proposition 7.3 Suppose $k \geq 2$. For any topological embedding $T_k \rightarrow S^n$, the corresponding adjacency graph Γ is isomorphic to the circuit Circ_k .

Proof We first claim that Γ is a connected graph. Indeed, if Γ had a separation, we could partition the components of $S^n - T_k$ into two disjoint sets corresponding to the separation. Taking the union of closures of components on each side of the separation gives a decomposition of S^n as the union of two closed sets whose intersection is homeomorphic to S^{n-2} . Such a decomposition is impossible by Alexander duality, since a subspace homeomorphic to S^{n-2} cannot disconnect S^n . Thus Γ is connected.

We now claim that Γ has no separating edge. Indeed, if it did, we could decompose S^n as the union of two closed sets whose intersection is a closed $(n-1)$ -disc, which is again impossible by Alexander duality. Thus Γ has no separating edge.

Note that the Euler characteristic of Γ is zero since Γ has an equal number of vertices and edges. Since Γ is a finite connected graph with k vertices, no separating edges, and Euler characteristic zero, the only possibility is that it is a circuit of length k . $\square$

7.2 A coarse adjacency graph

In principle, it seems reasonable to expect that the strategy above could be imitated in the coarse setting. But to perform this conversion would require a fair amount of machinery. Instead we replace the point-set topology arguments with entirely homological arguments. The homological approach has the advantage that it gives the adjacency graph a more homological interpretation. We will see that the adjacency graph is, in a certain sense, equivalent to a canonical homology class known as the Jordan cycle.

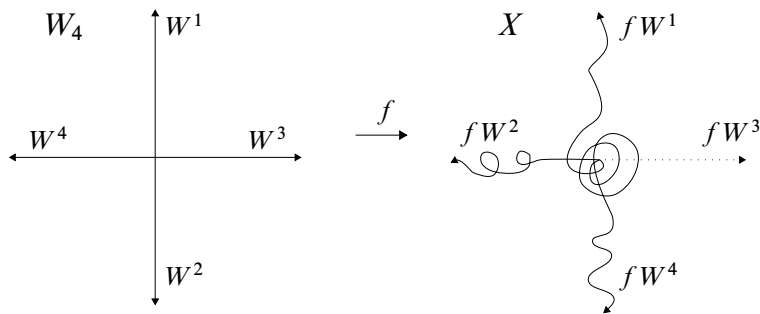


Figure 1: A coarse embedding of W_4 , a union of four 1-dimensional halfspaces glued along their boundaries, into a coarse PD(2) space X .

Standing assumptions 7.4 We use the following notation throughout this section as illustrated in Figure 1. Let X be a coarse PD(n) space. Let W^i be a bounded geometry metric simplicial complex homeomorphic to $\mathbb{R}_{\geq 0}^{n-1}$ for $i \in \{1, \dots, k\}$ with $k \geq 1$. Let

$$W_k = \bigcup_{i=1}^k W^i,$$

so that the halfspaces are glued along their boundaries by simplicial homeomorphisms. Let $L_i \subset W^i$ be the boundary of W^i , and let W_0 denote the image of these boundaries in W_k after gluing. Assume that W_0 is coarsely embedded in each halfspace W^i . Let $f: W_k \rightarrow X$ be a coarse embedding. Let $f_{\#}: C_o(W_k) \rightarrow C_o(X)$ be an approximating chain map as given by Lemma 3.9.

For each $I \subset \{1, \dots, k\}$, let W^I denote the union of halfspaces W^i with $i \in I$. If $I \subset J$, the inclusion $W^I \rightarrow W^J$ is a coarse embedding. Indeed, by a standard argument, any path connecting two points of W^I within W^J can be replaced with a path in W^I of controlled length using our assumption that $W_0 \rightarrow W^i$ is a coarse embedding for each i .

Remark 7.5 (cohomology calculations) If X is a coarse PD(n) space, then coarse Alexander duality allows one to calculate Čech homology associated to many coarse embeddings $Z \rightarrow X$ in terms of the compactly supported cohomology of Z . The cohomology of Z is easily calculated for several spaces of interest in this section. For example, it is well-known that $H_c^j(\mathbb{R}^\ell) \cong \mathbb{Z}$ if $j = \ell$ and is trivial otherwise. For halfspaces of dimension $n - 1$ the group $H_c^j(\mathbb{R}_{\geq 0}^{n-1}) = 0$ for all j . Similarly, $H_c^j(\mathbb{R}_{\geq 0}^\ell, \mathbb{R}^{\ell-1}) = \mathbb{Z}$ when $j = \ell$ and is trivial otherwise. A straightforward Mayer–Vietoris argument and induction on $k \geq 2$ shows that $H_c^j(W_k) = \mathbb{Z}^{k-1}$ if $j = n - 1$ and is trivial otherwise.

By the preceding calculation, the image fW_k coarsely separates X into k deep components corresponding to the relative ends $\mathcal{E}(X, fW_k)$. Intuitively, two relative ends may be considered adjacent if they meet along the image of a halfspace fW^i for some $i \in \{1, \dots, k\}$.

To make this idea more precise, we consider the effect of deleting a single halfspace from W_k . As explained below, this deletion reduces the number of relative ends by one in a precisely controlled manner that allows us to formalize the definition of adjacency.

Consider $W_{k-1}^i := \overline{W_k - W^i}$, the subcomplex obtained by removing the i^{th} halfspace from W_k . Whenever $2 \leq i \leq k$, coarse Alexander duality for the pair $W_{k-1}^i \hookrightarrow W_k$ gives the following commutative diagram with exact rows:

$$(7.6) \quad \begin{array}{ccccccc} 0 \rightarrow \check{H}_1^D((fW_{k-1}^i)^c, fW_k^c) & \xrightarrow{\partial} & \check{H}_0^D(fW_k^c) & \rightarrow & \check{H}_0^D((fW_{k-1}^i)^c) & \rightarrow & 0 \\ & \cong \downarrow A & & \cong \downarrow A & & \cong \downarrow A & \\ 0 \longrightarrow H_c^{n-1}(W_k, W_{k-1}^i) & \longrightarrow & H_c^{n-1}(W_k) & \longrightarrow & H_c^{n-1}(W_{k-1}^i) & \longrightarrow & 0 \end{array}$$

Note that $H_c^j(W_k, W_{k-1}^i) = H_c^j(\mathbb{R}_{\geq 0}^{n-1}, \mathbb{R}^{n-2})$ by excision, so (7.6) reduces to

$$0 \rightarrow \mathbb{Z} \rightarrow \mathbb{Z}^{k-1} \rightarrow \mathbb{Z}^{k-2} \rightarrow 0.$$

By Proposition 4.8, the map induced by inclusion $\mathcal{E}(X, fW_k) \rightarrow \mathcal{E}(X, fW_{k-1}^i)$ is surjective from a set of cardinality k to a set of cardinality $k-1$. In particular, for each i there is a unique pair of distinct relative ends in $\mathcal{E}(X, fW_k)$ with the same image in $\mathcal{E}(X, fW_{k-1}^i)$. We consider these relative ends to be adjacent as explained in the following definition.

Definition 7.7 (adjacency graph) The *adjacency graph* Γ has vertex set equal to the set of relative ends $V\Gamma = \mathcal{E}(X, fW_k)$. In the degenerate case that $k=1$, the graph Γ contains a single loop edge with both endpoints at the unique vertex. We now assume $k \geq 2$. For each i , we attach an unoriented edge e_i that connects the pair of distinct vertices that have the same image under the mapping

$$\mathcal{E}(X, fW_k) \rightarrow \mathcal{E}(X, fW_{k-1}^i).$$

In particular, the graph Γ contains k vertices and k unoriented edges $e_1, \dots, e_k$.

The edges of Γ can be oriented as follows. Let v_i and v'_i denote the vertices adjacent to e_i . Since $v_i - v'_i$ generates the kernel of the map from $\check{H}_0^D(fW_k^c)$ to $\check{H}_0^D((fW_{k-1}^i)^c)$ induced by the inclusion $fW_k^c \hookrightarrow (fW_{k-1}^i)^c$, then $v_i - v'_i$ also generates the image of the homomorphism

$$\partial: \check{H}_1^D((fW_{k-1}^i)^c, fW_k^c) \rightarrow \check{H}_0^D(fW_k^c).$$

Therefore, a choice of orientation for e_i is formally equivalent to a choice of generator of the cyclic group $\check{H}_1^D((fW_{k-1}^i)^c, fW_k^c)$.

Using coarse Alexander duality, we show in this section that the adjacency graph is isomorphic to the circuit Circ_k .

7.3 The adjacency chain complex

We have given a geometric definition above for the adjacency graph. However, in order to better understand the homology of this graph, we will study it from a more formal homological point of view. The construction in this subsection is analogous to the formal definition of the cellular chain complex that is used to develop cellular homology in terms of singular homology groups (see [12, Section 2.2]). Similarly we use Čech homology groups to define a formal adjacency chain complex in Definition 7.9. In Proposition 7.12, we show that this chain complex is isomorphic to the standard cellular chain complex of the geometrically defined adjacency graph.

Coarse Alexander duality for the pair $W_0 \hookrightarrow W_k$ yields the following commutative diagram with exact rows:

$$(7.8) \quad \begin{array}{ccccccc} 0 & \longrightarrow & \check{H}_1^D(fW_0^c) & \xrightarrow{\iota} & \check{H}_1^D(fW_0^c, fW_k^c) & \xrightarrow{\partial} & \check{H}_0^D(fW_k^c) \longrightarrow 0 \\ & & \cong \downarrow A & & \cong \downarrow A & & \cong \downarrow A \\ 0 & \longrightarrow & H_c^{n-2}(W_0) & \xrightarrow{\delta} & H_c^{n-1}(W_k, W_0) & \longrightarrow & H_c^{n-1}(W_k) \longrightarrow 0 \end{array}$$

which reduces to $0 \rightarrow \mathbb{Z} \rightarrow \mathbb{Z}^k \rightarrow \mathbb{Z}^{k-1} \rightarrow 0$.

Definition 7.9 Let $\mathcal{C}(\Gamma)$ denote the cellular chain complex $\{C_i(\Gamma), \partial\}$ of the graph Γ . We also find it convenient to consider the following *adjacency chain complex* $\mathcal{C}^a$ with $C_i^a = 0$ for $i \neq 0, 1$ and with $C_1^a \rightarrow C_0^a$ equal to

$$\check{H}_1^D(fW_0^c, fW_k^c) \xrightarrow{\partial} \check{H}_0^D(fW_k^c),$$

where ∂ denotes the composition of the connecting homomorphism ∂ of (7.8) with the canonical inclusion $\check{H}_0^D \rightarrow \check{H}_0^D$. We note that the two nontrivial terms of $\mathcal{C}^a$ are each isomorphic to $\mathbb{Z}^k$.

The group $C_0^a = \check{H}_0^D(fW_k^c) \cong \mathbb{Z}^k$ is freely generated by the set of relative ends $\mathcal{E}(X, fW_k)$ by Proposition 4.8. Hence, this basis for C_0^a corresponds to the vertex set in the adjacency graph. A natural basis for the group $C_1^a = \check{H}_1^D(fW_0^c, fW_k^c)$ is given by the following notation and lemma.

Notation 7.10 Let

$$\check{H}_1^D(fW_0^c, fW_k^c) \xrightarrow{\iota_i} \check{H}_1^D(fW_0^c, (fW^i)^c)$$

be the homomorphism induced by inclusion.

Lemma 7.11 The maps $\iota_1, \dots, \iota_k$ induce an isomorphism

$$\check{H}_1^D(fW_0^c, fW_k^c) \xrightarrow{I} \prod_{i=1}^k \check{H}_1^D(fW_0^c, (fW^i)^c).$$

Proof By [Proposition 6.4\(1\)](#), it suffices to show that maps induced by inclusion induce an isomorphism of cohomology with compact supports

$$H_c^{n-1}(W_k, W_0) \rightarrow \prod_{i=1}^k H_c^{n-1}(W^i, W_0).$$

This claim follows inductively by a straightforward Mayer–Vietoris argument. Indeed, for each $\ell = 2, \dots, k$ inclusions induce an isomorphism

$$H_c^{n-1}(W_\ell, W_0) \rightarrow H_c^{n-1}(W_{\ell-1}^\ell, W_0) \times H_c^{n-1}(W^\ell, W_0),$$

since the identity $H_c^*(W_{\ell-1}^\ell \cap W^\ell, W_0) = H_c^*(W_0, W_0) = 0$ holds in all dimensions. $\square$

Proposition 7.12 *The chain complex $\mathcal{C}^a$ is isomorphic to the cellular chain complex $\mathcal{C}(\Gamma)$ of the adjacency graph Γ .*

Proof The vertex set $V\Gamma = \mathcal{E}(X, fW_k)$ of the adjacency graph Γ freely generates the group $\check{H}_0^D(fW_k^c) = C_0^a$ by [Proposition 4.8](#), yielding an isomorphism $C_0(\Gamma) \rightarrow C_0^a$.

By the definition of Γ , its group of 1-chains $C_1(\Gamma)$ is isomorphic to the product

$$\prod_{i=1}^k \check{H}_1^D((fW_{k-1}^i)^c, fW_k^c),$$

and the boundary operator $\partial: C_1(\Gamma) \rightarrow C_0(\Gamma)$ is induced by the maps

$$\partial: \check{H}_1^D((fW_{k-1}^i)^c, fW_k^c) \rightarrow \check{H}_0^D(fW_k^c)$$

for $i \in \{1, 2, \dots, k\}$. By our standing assumptions, we may apply the excision result of [Proposition 6.4\(3\)](#) to the inclusion $(W^i, W_0) \hookrightarrow (W_k, W_{k-1}^i)$. Thus this inclusion induces an isomorphism

$$\check{H}_1^D((fW_{k-1}^i)^c, fW_k^c) \rightarrow \check{H}_1^D(fW_0^c, (fW^i)^c),$$

which defines an isomorphism $C_1(\Gamma) \rightarrow C_1^a$ induced by inclusion via the basis for C_1^a given in [Lemma 7.11](#).

In order to see that the two chain complexes are isomorphic, it suffices to show that following diagram commutes:

$$\begin{array}{ccc} C_1(\Gamma) & \xrightarrow{\partial} & C_0(\Gamma) \\ \downarrow \cong & & \downarrow \cong \\ C_1^a & \xrightarrow{\partial} & C_0^a \end{array}$$

It suffices to check commutativity on the basis elements of the product $C_1(\Gamma)$, which follows, for each i , from the commutativity of the diagram

$$\begin{array}{ccc}
 \check{H}_1^D(fW_0^c, fW_k^c) & \xrightarrow{\partial} & \check{H}_0^D(fW_k^c) \\
 \cong \uparrow I^{-1} & & \uparrow \text{id} \\
 \prod_{i=1}^k \check{H}_1^D(fW_0^c, (fW^i)^c) & & \\
 \uparrow i_* & & \\
 \check{H}_1^D(fW_0^c, (fW^i)^c) & & \\
 \cong \uparrow \text{Excision} & & \\
 \check{H}_1^D((fW_{k-1}^i)^c, fW_k^c) & \xrightarrow{\partial} & \check{H}_0^D(fW_k^c)
 \end{array}$$

where I^{-1} is the inverse of the isomorphism I in [Lemma 7.11](#), and the map in the right column is the identity map. To see that the diagram above commutes, notice that we have inclusions of pairs

$$((fW_{k-1}^i)^c, fW_k^c) \subset (fW_0^c, fW_k^c) \subset (fW_0^c, (fW^i)^c).$$

The map $i_* I^{-1}$ is the inverse of a map induced by inclusion, so the composition of the three maps in the left column of the diagram is a map induced by inclusion. Commutativity of the diagram now follows from the naturality of the connecting homomorphism, given by [Proposition 6.4](#). $\square$

7.4 The adjacency graph is a circuit

The homology of the adjacency graph may be seen by examining the adjacency chain complex and using (7.8). One sees that the adjacency graph is connected and homotopy equivalent to a circuit; in other words, it deformation retracts onto a core graph that is isomorphic to a circuit. In this subsection, we determine that the adjacency graph is actually isomorphic to a circuit by showing that such a deformation retraction cannot collapse any edge. The main technique used in this proof is the Jordan cycle, a 1-cycle that coarsely links with the set fW_0 . We show that the adjacency graph is, in a sense, equivalent to the Jordan cycle.

Definition 7.13 Let the *Jordan cycle* σ be a generator of $\check{H}_1^D(fW_0^c) \cong \mathbb{Z}$.

The Jordan cycle maps into the group $C_1^a = \check{H}_1^D(fW_0^c, fW_k^c)$ via (7.8). In the next lemma we show that under the natural coordinate system on $\check{H}_1^D(fW_0^c, fW_k^c)$ given in [Lemma 7.11](#) the Jordan cycle maps to the diagonal.

Lemma 7.14 If σ denotes the Jordan cycle and

$$\check{H}_1^D(fW_0^c) \xrightarrow{\iota} \check{H}_1^D(fW_0^c, fW_k^c)$$

is the map induced by inclusion, then $I \circ \iota(\sigma) = (1, 1, \dots, 1)$, for some choice of generators (each denoted 1) for each infinite cyclic group $\check{H}_1^D(fW_0^c, (fW^i)^c)$.

Proof The lemma follows by examining the following commutative diagram:

$$\begin{array}{ccc}
 \langle \sigma \rangle \cong \check{H}_1^D(fW_0^c) & \xrightarrow{\iota} & \check{H}_1^D(fW_0^c, fW_k^c) \\
 \downarrow \cong \iota_A & \swarrow \iota_i & \downarrow \cong I=(\iota_1, \dots, \iota_k) \\
 \check{H}_1^D(fW_0^c, (fW^i)^c) & \xleftarrow{\pi_i} \prod_{i=1}^k \check{H}_1^D(fW_0^c, (fW^i)^c) &
 \end{array}$$

The lower right triangle commutes by the definition of I in terms of the universal mapping property for direct products. The upper left triangle commutes since the three maps are each induced by inclusion. By [Lemma 7.11](#), the map I is an isomorphism. Applying [Theorem 6.2](#) to the pair $W_0 \hookrightarrow W^i$ for $1 \leq i \leq k$ gives the exact sequence

$$(7.15) \quad 0 \longrightarrow \check{H}_1^D(fW_0^c) \xrightarrow{\iota_A} \check{H}_1^D(fW_0^c, (fW^i)^c) \longrightarrow 0$$

so that the map ι_A is an isomorphism for each i . It follows that $I \circ \iota(\sigma) = (1, 1, \dots, 1)$. $\square$

Theorem 7.16 *The adjacency graph is isomorphic to a circuit.*

Proof The homology of the adjacency graph is equal to the homology of the adjacency chain complex $\mathcal{C}^a$ by [Proposition 7.12](#), which is given by $H_0(\mathcal{C}^a) = \mathbb{Z}$ and $H_1(\mathcal{C}^a) = \mathbb{Z}$ by (7.8). Therefore, the adjacency graph Γ deformation retracts to a circuit Γ_1 . By (7.8), the infinite cyclic group $H_1(\mathcal{C}^a)$ is generated by the image of the Jordan cycle $\sigma \in \check{H}_1^D(fW_0^c)$. In the coordinates on C_1^a given by the basis in [Lemma 7.11](#), the Jordan cycle maps to the diagonal $(1, 1, \dots, 1)$ by [Lemma 7.14](#). As described in the proof of [Proposition 7.12](#), this basis for C_1^a is in bijective correspondence with edges of the adjacency graph. Therefore, each edge of Γ also belongs to the circuit Γ_1 , which implies that $\Gamma = \Gamma_1$ is a circuit. $\square$

7.5 Maps between adjacency graphs

Notation 7.17 Let $1 \leq i \leq k$, and let $W_{k-1}^i := \overline{W_k - W^i}$ be the subcomplex obtained by removing the i^{th} halfspace from W_k . Let Γ be the adjacency graph of the coarse embedding $f: W_k \rightarrow X$, and let Γ_i be the adjacency graph of the restriction $f|_{W_{k-1}^i}: W_{k-1}^i \rightarrow X$.

Theorem 7.18 *For each $k \geq 2$, the map of vertex sets induced by inclusion extends to a map of graphs $\phi: \Gamma \rightarrow \Gamma_i$, which collapses the circuit $\Gamma = \text{Circ}_k$ to the circuit $\Gamma_i = \text{Circ}_{k-1}$ by collapsing the i^{th} edge to a point.*

Proof When $k = 2$ the result is trivial, so we assume $k > 2$. For each i , the map $\mathcal{E}(X, fW_k) \rightarrow \mathcal{E}(X, fW_{k-1}^i)$ collapses the pair of vertices corresponding to W^i to a single vertex. Whenever $i \neq j$,

consider $W_{k-2}^{i,j} := \overline{W_k - (W^i \cup W^j)}$. Then the following diagram of maps induced by inclusion is commutative:

$$\begin{array}{ccc} \mathcal{E}(X, fW_k) & \longrightarrow & \mathcal{E}(X, fW_{k-1}^i) \\ \downarrow & & \downarrow \\ \mathcal{E}(X, fW_{k-1}^j) & \longrightarrow & \mathcal{E}(X, fW_{k-2}^{i,j}) \end{array}$$

Therefore, the map $\phi: \Gamma \rightarrow \Gamma_i$ sends the pair of vertices corresponding to W^j to the pair of vertices corresponding to W^j . $\square$

The geometric structure of the map between graphs given in [Theorem 7.18](#) is detailed in the next lemma and will be needed in the next section.

Lemma 7.19 *Let $\phi: \Gamma \rightarrow \Gamma_i$ be the map defined in [Theorem 7.18](#). There exists a constant $R_0 > 0$ such that the following holds. Suppose $u = \phi(v_i) = \phi(v'_i) \in V\Gamma_i$ for $v_i \neq v'_i \in V\Gamma$. Let $L \in \mathcal{N}(fW_{k-1}^i)$, as in [Definition 4.1](#), satisfy $N_{R_0}(fW_{k-1}^i) \subseteq L$. Then, all deep components in $\pi_0(\overline{X - L})$ are stable, and there is a constant M such that $f(W_k - N_M(W_{k-1}^i)) \subset U$, where U is the deep component in $\pi_0(\overline{X - L})$ associated to u .*

Proof As observed in [Remark 7.5](#), the cohomology group $H_c^{n-1}(W_{k-1}^i)$ is free abelian of finite rank. Thus by coarse Alexander duality ([Theorem 5.1](#)), the dual inverse system $\{H_0(\overline{X - L})\}_{L \in \mathcal{N}(fW_{k-1}^i)}$ is stable and its inverse limit is free abelian of finite rank. By [Proposition 4.8](#), there exists $R_0 < \infty$ such that if $L \in \mathcal{N}(fW_{k-1}^i)$ satisfies $N_{R_0}(fW_{k-1}^i) \subseteq L$, then all deep components in $\pi_0(\overline{X - L})$ are stable, and the natural projection

$$\mathcal{E}(X, fW_{k-1}^i) \rightarrow \pi_0(\overline{X - L})$$

is a bijection onto the set of deep components of $\overline{X - L}$. [Proposition 4.8](#) also provides a neighborhood $L' \in \mathcal{N}(fW_{k-1}^i)$ such that any component in $\pi_0(\overline{X - L})$ that is not deep is contained in L' . Choose $M > 0$ so that f induces a map

$$\overline{W_k - N_M(W_{k-1}^i)} \rightarrow \overline{X - L'} \rightarrow \overline{X - L}.$$

Then f maps each component in $\pi_0(\overline{W_k - N_M(W_{k-1}^i)})$ into a single deep component in $\pi_0(\overline{X - L})$. We will prove that U is the deep component associated to u .

Let $K \in \mathcal{N}(fW_k)$ such that all deep components in $\pi_0(\overline{X - K})$ are stable and let V and V' be the deep components in $\pi_0(\overline{X - K})$ associated to v_i and v'_i , respectively. Since K is a subcomplex in $\mathcal{N}(fW_k)$, there is a number R_1 such that $K \subset N_{R_1}(f(W_k))$. Using the hypothesis f is a coarse embedding, we can find a positive constant R_2 such that $f(N_M(W_{k-1}^i)) \subset N_{R_2}(fW_{k-1}^i)$. Let $L_1 \in \mathcal{N}(fW_{k-1}^i)$ such that $N_{R_1+R_2}(L) \subset L_1$, and let U_1 be the deep component in $\pi_0(\overline{X - L_1})$ that is associated to u . Therefore, $N_{R_1}(U_1)$ lies in the deep component U' of $\pi_0(\overline{X - L})$ that is associated to u . We now prove

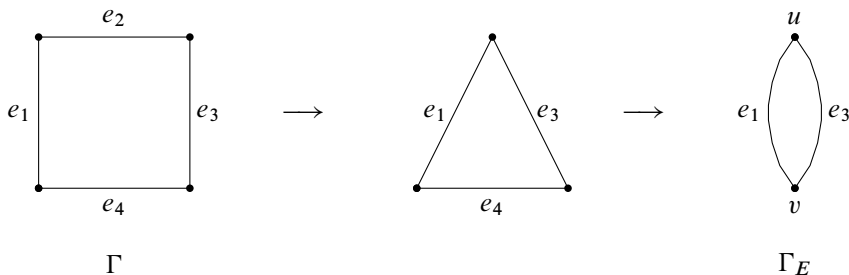


Figure 2: Collapse of adjacency graphs.

that $f(W_k - N_M(W_{k-1}^i))$ has a nonempty intersection with U' and this will imply that $U = U'$ is the deep component in $\pi_0(\overline{X - L})$ that is associated to u .

Let $K_1 \in \mathcal{N}(fW_k)$ such that $L_1 \subset K_1$ and $K \subset K_1$. Let V_1 and V'_1 be the deep components in $\pi_0(\overline{X - K_1})$ associated to v_i and v'_i respectively. Therefore, $V_1 \subset V$, $V'_1 \subset V'$, and both V_1 and V'_1 lie in the component U_1 . Let α be a path in U_1 joining a point in V_1 to a point in V'_1 . Therefore, α is a path between two deep components V and V' in $\pi_0(\overline{X - K})$, which implies that $\alpha \cap K$ is nonempty. Since $K \subset N_{R_1}(fW_k)$, there is a point w in W_k such that $f(w) \in N_{R_1}(\alpha) \subset N_{R_1}(U_1) \subset U'$. Also,

$$f(w) \in N_{R_1}(U_1) \subset \overline{X - N_{R_2}(L)} \subset \overline{X - N_{R_2}(fW_{k-1}^i)}.$$

Thus, $w \in W_k - N_M(W_{k-1}^i)$ by the choice of R_2 , which implies that $f(w)$ lies in $f(W_k - N_M(W_{k-1}^i))$. Therefore, the intersection $f(W_k - N_M(W_{k-1}^i)) \cap U'$ is nonempty, so $U = U'$ is the deep component in $\pi_0(\overline{X - L})$ associated to u . $\square$

Example 7.20 Consider the coarse embedding $f: W_k \rightarrow X$ for the case $k = 4$, which will be used in the next section. Let Γ be the adjacency graph of the embedding as in Figure 2. Each edge e_i is associated to the halfspace W^i . Let $E = W^1 \cup W^3 = \overline{W_4 - (W^2 \cup W^4)}$ be the subcomplex of W_4 obtained by removing the second halfspace W^2 and the fourth halfspace W^4 . The adjacency graph Γ_E of the embedding $f|_E: E \rightarrow X$ is obtained as in Figure 2. The relative ends $\mathcal{E}(X, fE)$ has two elements u and v , which are used to label the vertices of Γ_E as in Figure 2. By Lemma 7.19 there is a positive constant R such that the following holds. For each $L \in \mathcal{N}(fE)$ satisfying $N_R(fE) \subseteq L$ all deep components in $\pi_0(\overline{X - L})$ are stable and there is a constant M such that $f(W^2 - N_M(E)) \subset U$ (resp. $f(W^4 - N_M(E)) \subset V$), where U (resp. V) is the deep component in $\pi_0(\overline{X - L})$ that is associated to u (resp. v).

8 Amalgams with Klein bottle groups are not Kleinian

Definition 8.1 (isolated flats) A k -flat is an isometrically embedded copy of $\mathbb{E}^k$ for $k \geq 2$. Let Y be a CAT(0) space, and let $\text{Flat}(Y)$ denote the space of all flats in Y with the topology of Hausdorff convergence on bounded sets. A CAT(0) space Y with a proper, cocompact, isometric action by a group

G has *isolated flats* if it contains an equivariant collection $\mathcal{F}$ of flats such that $\mathcal{F}$ is closed and isolated in $\text{Flat}(Y)$ and each flat $F \subseteq X$ is contained in a uniformly bounded tubular neighborhood of some $F' \in \mathcal{F}$. We refer the reader to [13] for more details on spaces with isolated flats.

Definition 8.2 (class of groups) Let S_g denote a closed orientable surface of genus g greater than one. Let T^2 denote the 2-torus, and let K^2 denote the Klein bottle. Let $\mathcal{G}_T$ denote the collection of amalgamated free products of the form $\pi_1(S_g) *_{\langle a=b \rangle} \pi_1(T^2)$, where a is the homotopy class of an essential simple closed curve on S_g and b is the homotopy class of an essential simple closed curve on T^2 . Let $\mathcal{G}_K$ denote the collection of amalgamated free products of the form $\pi_1(S_g) *_{\langle a=c \rangle} \pi_1(K^2)$, where a is the homotopy class of an essential simple closed curve on S_g and c is the homotopy class of an essential simple closed curve on K^2 which is orientation reversing. In particular, the element $c \in \pi_1(K^2)$ acts by a glide reflection on the universal cover of K^2 equipped with a Euclidean metric.

Construction 8.3 (model geometry) Let $G = \pi_1(S_g) *_{\langle a=c \rangle} \pi_1(K^2) \in \mathcal{G}_K$. For $\ell > 0$, there exists a hyperbolic metric on S_g and a Euclidean metric on K^2 so the geodesic representatives of a and c each have length ℓ . Glue a circle of length ℓ to a and c by isometries to form a 2-complex X . Then $G \cong \pi_1(X)$.

Observe that the universal cover $\tilde{X}$ of X is a CAT(0) space with isolated flats, since it consists of flat subspaces separated by CAT(−1) regions, as discussed in [18].

Definition 8.4 A group G is *Kleinian* if it is isomorphic to a discrete subgroup of $\text{Isom}(\mathbb{H}^3)$. Note that for the purposes of this paper a Kleinian group may contain isometries that reverse orientation.

Theorem 8.5 If $G \in \mathcal{G}_K$, then G is not a Kleinian group.

Proof Let $G = \pi_1(S_h) *_{\langle a=c \rangle} \pi_1(K^2)$, where a is the homotopy class of an essential simple closed curve on S_h and c is the homotopy class of an essential simple closed curve on K^2 which is orientation reversing. Then, $G \cong \pi_1(X)$ as defined in Construction 8.3. Suppose towards a contradiction that G is a Kleinian group.

Let $P = \pi_1(K^2) \leq G$. Since P has a finite-index subgroup isomorphic to $\mathbb{Z}^2$, each conjugate of P in G acts parabolically on $\mathbb{H}^3$. Since the universal cover $\tilde{X}$ has isolated flats, the Hausdorff distance between any two flat subspaces of $\tilde{X}$ is infinite. In particular, the stabilizers of distinct flats are not commensurable. It follows that each conjugate of P in G fixes a distinct parabolic point in $\partial\mathbb{H}^3$, the visual boundary of $\mathbb{H}^3$. Indeed, if $p \in \partial\mathbb{H}^3$ is stabilized by a parabolic isometry of G , then $\text{Stab}_G(p)$ does not contain any loxodromic elements since G acts discretely on $\mathbb{H}^3$. So, $\text{Stab}_G(p)$ is a virtually abelian group of rank at most two. In particular, if P and gPg^{-1} for $g \in G$ fix the same parabolic point p , then P and gPg^{-1} are commensurable in $\text{Stab}_G(p)$, a contradiction.

We will show there exists a horosphere F stabilized by P that does not intersect any of its G -translates. Each conjugate of P in G stabilizes a family of horospheres centered at a parabolic point on $\partial\mathbb{H}^3$. Since

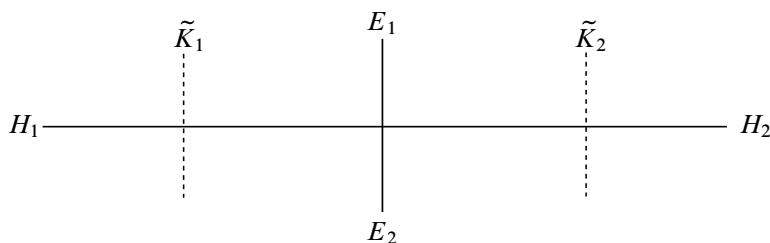


Figure 3: In $\tilde{X}$, a Euclidean plane $E_1 \cup E_2$ intersects a hyperbolic plane $H_1 \cup H_2$ in a line. The image depicts the cyclic order of the image of these half-spaces under a G -equivariant coarse embedding to $\mathbb{H}^3$. If Euclidean planes $\tilde{K}_1$ and $\tilde{K}_2$ are chosen very far from the plane $E_1 \cup E_2$, they map to opposite sides of the horospherical image of the plane $E_1 \cup E_2$ in $\mathbb{H}^3$, a contradiction.

G acts on $\mathbb{H}^3$ discretely and P acts on each horosphere cocompactly, if F is any horosphere stabilized by P , then there exists a compact subset $C \subset F$ so that $P \cdot C$ covers F . Any horosphere gF for $g \in G$ satisfying $gF \cap F \neq \emptyset$ can be translated using P to a horosphere stabilized by a conjugate of P that intersects C . Since G acts on $\mathbb{H}^3$ properly, there are only finitely many horospheres stabilized by a conjugate of P which intersect C . Thus, the horosphere F can be shrunk so that it is disjoint from all of its G -translates. Moreover, all other G -translates of F lie in one component of $\mathbb{H}^3 \setminus F$.

To reach a contradiction, we use coarse separation results and the action of G on $\tilde{X}$ to prove the horosphere F separates two of its G -translates. Let $\ell \subset \tilde{X}$ be the line stabilized by $\langle a \rangle$. The line ℓ is contained in $\tilde{S}_h$, a copy of the universal cover of S_h , and $\tilde{K}_0$, a copy of the universal cover of the Klein bottle. Let H_1 and H_2 be the half-planes in $\tilde{S}_h$ bounded by ℓ , and let E_1 and E_2 be the half-planes in $\tilde{K}_0$ bounded by ℓ . See Figure 3. Let $W = H_1 \cup H_2 \cup E_1 \cup E_2$.

Since $G = \pi_1(X)$ acts properly by isometries on $\mathbb{H}^3$, there is a G -equivariant coarse embedding $f: \tilde{X} \rightarrow \mathbb{H}^3$. There is a constant A such that for each lift $\tilde{K}$ of the Klein bottle in $\tilde{X}$ there is a G -translate gF in $\mathbb{H}^3$ such that the Hausdorff distance between $f(\tilde{K})$ and gF is less than A . Let Γ be the adjacency graph of the coarse embedding $f|_W: W \rightarrow \mathbb{H}^3$ as defined in Section 7. The graph Γ is a 4-cycle whose edges are associated to (and labeled by) H_1 , H_2 , E_1 , and E_2 . Since the action of a on $\tilde{X}$ stabilizes W -permuting its four halfplanes — and the coarse embedding $W \rightarrow \mathbb{H}^3$ is a -equivariant, the action of a induces a graph automorphism of Γ .

Moreover, the action on Γ stabilizes each edge labeled by H_i and interchanges edges labeled by E_1 and E_2 . Therefore, two edges of the 4-cycle Γ labeled by H_1 and H_2 are opposite.

Recall that $\tilde{K}_0 = E_1 \cup E_2$. By Example 7.20 and the fact the Hausdorff distance between $f\tilde{K}_0$ and F is finite, the relative end $\mathcal{E}(\mathbb{H}^3, F)$ consists of two elements. Moreover, there exists a positive constant R such that for each $L \in \mathcal{N}(F)$ satisfying $N_R(F) \subseteq L$, all deep components in $\pi_0(\overline{\mathbb{H}^3 - L})$ are stable, and there is a constant M such that $f(H_i - N_M(\tilde{K}_0)) \subset U_i$, where U_1 and U_2 are deep components in $\pi_0(\overline{\mathbb{H}^3 - L})$, each associated to one of the two elements of $\mathcal{E}(\mathbb{H}^3, F)$.

As explained in Example 7.20, the adjacency graph 4-cycle associated to $W \rightarrow \mathbb{H}^3$ collapses onto the adjacency graph 2-cycle associated to $\tilde{K}_0$ by a map that collapses each edge H_i to a point. In particular,

the two vertices of Γ adjacent to H_1 map to the same side of $\tilde{K}_0$, which is different from the side containing the two vertices adjacent to H_2 . In particular, $U_1 \neq U_2$.

Let V_1 and V_2 be the two components of $\mathbb{H}^3 - F$. Assume that V_i is associated to the same relative end as U_i for each i . Let L be a subcomplex in $\mathcal{N}(F)$ such that $N_{R+A}(F) \subseteq L$. Observe that $N_A(U_i) \subset V_i$ for each i . By the above discussion, there is a positive number M such that $f(H_i - N_M(\tilde{K}_0)) \subset U_i$ for each i . Let $\tilde{K}_1$ and $\tilde{K}_2$ be two lifts of the Klein bottle in $\tilde{X}$ such that each intersection $\tilde{K}_i \cap H_i$ contains a point not in $N_M(\tilde{K}_0)$. Therefore, $f\tilde{K}_i$ intersects with U_i for each i . Let g_1F and g_2F be two G -translates of F such that the Hausdorff distance between g_iF and $f\tilde{K}_i$ is less than A for each i . Therefore, each intersection $g_iF \cap V_i$ is nonempty which implies that $g_iF \subset V_i$ for each i . In other words, the horosphere F separates its G -translates g_1F and g_2F which is a contradiction. Therefore, G is not a Kleinian group. $\square$

Proposition 8.6 *Let S_g be a closed orientable surface of genus $g \geq 2$, and let a be the homotopy class of an essential simple closed curve in S_g . Then the amalgamated product*

$$G = \pi_1(S_g) *_{\langle a \rangle} (\langle a \rangle \times \mathbb{Z}).$$

is isomorphic to a geometrically finite Kleinian group.

Proof The group G is the fundamental group of a compact 3-manifold M formed from the product $S_g \times [-1, 1]$ by removing the interior of a closed tubular neighborhood N of the curve $a_0 \times \{0\}$, where a_0 represents the homotopy class a . Since $(M, \partial N)$ is a compact atoroidal Haken pared 3-manifold, the result follows directly from Thurston's hyperbolization theorem (see [15, Theorem 1.43]). $\square$

Proposition 8.7 *The group $G = \pi_1(S_g) *_{\langle a=c \rangle} \pi_1(K^2) \in \mathcal{G}_K$, where a is the homotopy class of a nonseparating simple closed curve on S_g has an index two subgroup that is a geometrically finite Kleinian group.*

Proof Suppose $G = \pi_1(S_g) *_{\langle a=c \rangle} \pi_1(K^2) \in \mathcal{G}_K$, where a is the homotopy class of a nonseparating simple closed curve on S_g . Then $G \cong \pi_1(X)$, where X is the union of S_g and K^2 glued along essential simple closed curves a_0 and c_0 on S_g and K^2 , respectively. The curve a_0 is a nonseparating simple closed curve on S_g in the homotopy class of $a \in \pi_1(S_g)$, and c_0 is an orientation-reversing simple closed curve on K^2 in the homotopy class of $c \in \pi_1(K^2)$.

To prove the proposition, we show that X has a double cover $\hat{X}$ formed from a closed orientable surface S_{2g-1} and a torus T^2 by identifying an essential simple closed curve from each surface. Since the loop c_0 reverses orientation, its preimage in the canonical orientable double cover $T^2 \rightarrow K^2$ is a single simple closed curve $\hat{c}$ double covering c_0 . Since the simple closed curve a_0 does not separate S_g , it represents a nontrivial element of $H_1(S_g; \mathbb{Z}/2\mathbb{Z}) = (\mathbb{Z}/2\mathbb{Z})^{2g}$. Thus there exists a homomorphism

$$\pi_1(S_g) \rightarrow H_1(S_g; \mathbb{Z}) \rightarrow H_1(S_g; \mathbb{Z}/2\mathbb{Z}) \rightarrow \mathbb{Z}/2\mathbb{Z}$$

mapping a to the nontrivial element of $\mathbb{Z}/2\mathbb{Z}$. In the double cover $S_{2g-1} \rightarrow S_g$, the curve a_0 has preimage a single simple closed curve $\hat{a}$ double covering a_0 . Form a 2-complex $\hat{X}$ from T^2 and S_{2g-1} by identifying $\hat{c}$ with $\hat{a}$. Then $\hat{X}$ double covers X . By [Proposition 8.6](#) the group $\pi_1(\hat{X})$ is isomorphic to a geometrically finite Kleinian group. $\square$

Remark 8.8 If $G = \pi_1(S_g) *_{\langle a=c \rangle} \pi_1(K^2) \in \mathcal{G}_K$ where a is the homotopy class of a separating simple closed curve on S_g , then G has an index four subgroup that is geometrically finite and Kleinian. The proof above does not apply since if a is separating, it has trivial image in the abelianization $H_1(S_g)$ of $\pi_1(S_g)$, and hence also in the abelian group $\mathbb{Z}/2\mathbb{Z}$. In particular, a is contained in every index two subgroup of $\pi_1(S_g)$, and any loop representing a has a disconnected preimage in every double cover of S_g . However, G and $G' = \pi_1(S_g) *_{\langle a=b \rangle} \pi_1(T^2) \in \mathcal{G}_K$ do have isomorphic index four subgroups, and the group G' is a geometrically finite Kleinian group. We leave the details as an exercise for the reader.

References

- [1] **M Bestvina**, *Local homology properties of boundaries of groups*, Michigan Math. J. 43 (1996) 123–139 [MR](#) [Zbl](#)
- [2] **M Bestvina, M Kapovich, B Kleiner**, *Van Kampen’s embedding obstruction for discrete groups*, Invent. Math. 150 (2002) 219–235 [MR](#) [Zbl](#)
- [3] **J Block, S Weinberger**, *Large scale homology theories and geometry*, from “Geometric topology” (W H Kazez, editor), AMS/IP Stud. Adv. Math. 2.1, Amer. Math. Soc., Providence, RI (1997) 522–569 [MR](#) [Zbl](#)
- [4] **K S Brown**, *Cohomology of groups*, Grad. Texts in Math. 87, Springer (1982) [MR](#) [Zbl](#)
- [5] **J Dydak**, *Local n -connectivity of quotient spaces and one-point compactifications*, from “Shape theory and geometric topology” (S Mardešić, J Segal, editors), Lecture Notes in Math. 870, Springer, Berlin (1981) 48–72 [MR](#) [Zbl](#)
- [6] **S Eilenberg, N Steenrod**, *Foundations of algebraic topology*, Princeton Univ. Press (1952) [MR](#) [Zbl](#)
- [7] **B Farb, R Schwartz**, *The large-scale geometry of Hilbert modular groups*, J. Differential Geom. 44 (1996) 435–478 [MR](#) [Zbl](#)
- [8] **R Geoghegan**, *The shape of a group: connections between shape theory and the homology of groups*, from “Geometric and algebraic topology” (H Toruńczyk, S Jackowski, S Spież, editors), Banach Center Publ. 18, PWN, Warsaw (1986) 271–280 [MR](#) [Zbl](#)
- [9] **R Geoghegan**, *Topological methods in group theory*, Grad. Texts in Math. 243, Springer (2008) [MR](#) [Zbl](#)
- [10] **R Geoghegan, E Swenson**, *On semistability of CAT(0) groups*, Groups Geom. Dyn. 13 (2019) 695–705 [MR](#) [Zbl](#)
- [11] **C R Guilbault**, *Ends, shapes, and boundaries in manifold topology and geometric group theory*, from “Topology and geometric group theory” (M W Davis, J Fowler, J-F Lafont, I J Leary, editors), Springer Proc. Math. Stat. 184, Springer (2016) 45–125 [MR](#) [Zbl](#)
- [12] **A Hatcher**, *Algebraic topology*, Cambridge Univ. Press (2002) [MR](#) [Zbl](#)
- [13] **G C Hruska, B Kleiner**, *Hadamard spaces with isolated flats*, Geom. Topol. 9 (2005) 1501–1538 [MR](#) [Zbl](#)

- [14] **G C Hruska, E Stark, H C Tran**, *Surface group amalgams that (don't) act on 3-manifolds*, Amer. J. Math. 142 (2020) 885–921 [MR](#) [Zbl](#)
- [15] **M Kapovich**, *Hyperbolic manifolds and discrete groups*, Progr. Math. 183, Birkhäuser, Boston, MA (2001) [MR](#) [Zbl](#)
- [16] **M Kapovich, B Kleiner**, *Hyperbolic groups with low-dimensional boundary*, Ann. Sci. École Norm. Sup. 33 (2000) 647–669 [MR](#) [Zbl](#)
- [17] **M Kapovich, B Kleiner**, *Coarse Alexander duality and duality groups*, J. Differential Geom. 69 (2005) 279–352 [MR](#) [Zbl](#)
- [18] **M Kapovich, B Leeb**, *On asymptotic cones and quasi-isometry classes of fundamental groups of 3-manifolds*, Geom. Funct. Anal. 5 (1995) 582–603 [MR](#) [Zbl](#)
- [19] **M Kapovich, B Leeb**, *Quasi-isometries preserve the geometric decomposition of Haken manifolds*, Invent. Math. 128 (1997) 393–416 [MR](#) [Zbl](#)
- [20] **S Mardešić, J Segal**, *Shape theory: the inverse system approach*, North-Holland Math. Libr. 26, North-Holland, Amsterdam (1982) [MR](#) [Zbl](#)
- [21] **J Milnor**, *On the Steenrod homology theory*, from “Novikov conjectures, index theorems and rigidity, I” (S C Ferry, A Ranicki, J Rosenberg, editors), Lond. Math. Soc. Lect. Note Ser. 226, Cambridge Univ. Press (1995) 79–96 [MR](#) [Zbl](#)
- [22] **W D Neumann, G A Swarup**, *Canonical decompositions of 3-manifolds*, Geom. Topol. 1 (1997) 21–40 [MR](#) [Zbl](#)
- [23] **J Roe**, *Lectures on coarse geometry*, Univ. Lect. Ser. 31, Amer. Math. Soc., Providence, RI (2003) [MR](#) [Zbl](#)
- [24] **E H Spanier**, *Algebraic topology*, McGraw-Hill, New York (1966) [MR](#) [Zbl](#)

Department of Mathematical Sciences, University of Wisconsin–Milwaukee
Milwaukee, WI, United States

Department of Mathematics, Wesleyan University
Middletown, CT, United States

FCCI Insurance Group
Sarasota, FL, United States

chruska@uwm.edu, estark@wesleyan.edu, hungtran280687@gmail.com

Received: 5 May 2022 Revised: 26 January 2024

K-cowaist on complete foliated manifolds

GUANGXIANG SU
XIANGSHENG WANG

Let (M, F) be a connected (not necessarily compact) foliated manifold carrying a complete Riemannian metric g^{TM} . We generalize Gromov's K-cowaist using the coverings of M , as well as defining a closely related concept called the $\hat{A}$ -cowaist. Let k^F be the associated leafwise scalar curvature of $g^F = g^{TM}|_F$. We obtain some estimates on k^F using these two concepts. In particular, assuming that the generalized K-cowaist is infinity and either TM or F is spin, we show that $\inf(k^F) \leq 0$.

[58J20](#); [53C12](#), [53C21](#)

1 Introduction

1.1 Main results

Let M be a closed connected oriented smooth Riemannian manifold. Let $E \rightarrow M$ be a Hermitian vector bundle with a Hermitian connection ∇^E and R^E be the curvature of ∇^E . If $\dim M$ is even, Gromov [7, Section 4; 9, Section 4.1.4] defines the K-cowaist¹ of M by

$$\text{K-cw}_2(M) = \sup_E (\|R^E\|^{-1}),$$

where $E \rightarrow M$ is a unitary bundle for which (at least) one characteristic (Chern) number of E does not vanish. Gromov also generalizes the definition of the K-cowaist to open manifolds by sticking to bundles $E \rightarrow M$ trivialized at infinity and using the characteristic numbers coming from the cohomology with compact supports. Moreover, if $\dim M$ is odd, Gromov defines

$$\text{K-cw}_2(M) = \sup_k \text{K-cw}_2(M \times \mathbb{R}^k),$$

where one takes those $k \geq 0$ such that $\dim M + k$ is even.

The K-cowaist is closely related to scalar curvature. In [7, Section 5 $\frac{1}{4}$], Gromov proves that every complete Riemannian spin manifold of dimension n with the scalar curvature $k^{g^{TM}} \geq \varepsilon^{-2}$ satisfies

$$\text{K-cw}_2(M) \leq \text{const}_n \varepsilon^2.$$

¹In [7], K-cowaist was called K-area. But recently, Gromov [9] suggests that K-cowaist should be a more proper name for this concept.

In [7, Section 9²], Gromov also defines the K-cowaist for foliated manifolds. In this paper, for the case that M is a foliated manifold again, we further generalize the definition of K-cowaist by considering the coverings of M . As in [7], we also study the relation between the leafwise scalar curvature and this generalized K-cowaist.

We now explain it in detail. Let M be a connected oriented (not necessarily compact) manifold carrying a (not necessarily complete) Riemannian metric g^{TM} . Let $F \subseteq TM$ be an integrable subbundle of TM and $g^F = g^{TM}|_F$ be the restricted metric on F . In the following, we assume that both $\dim M$ and $\text{rk}(F)$ are even. If $\dim M$ is odd and $\text{rk}(F)$ is even, we replace M by $M \times S^1$. If $\text{rk}(F)$ is odd, we replace F by $F \oplus TS^1$ and M by $M \times S^1 \times S^1$ or $M \times S^1$ depending on whether $\dim M$ is even or odd.

We take $\tilde{\pi}: \tilde{M} \rightarrow M$ to be a covering of M and $\tilde{F}$ to be the lifted foliation on $\tilde{M}$. Then $\tilde{M}$ and $\tilde{F}$ carry the lifted metrics $g^{T\tilde{M}}$ and $g^{\tilde{F}}$.

Let (E, g^E, ∇^E) be a Hermitian vector bundle over $\tilde{M}$ with the Hermitian metric g^E and the Hermitian connection ∇^E . We assume that E is trivial at infinity.

Let $R^E = (\nabla^E)^2$ be the curvature of ∇^E . Hence, for any $x \in \tilde{M}$ and $\alpha, \beta \in T_x \tilde{M}$, $R^E(\alpha \wedge \beta) \in \text{End}(T_x \tilde{M})$. Recall that $\|R_{\tilde{F}}^E\|$ is defined by² (cf [7]),

$$(1-1) \quad \|R_{\tilde{F}}^E\| = \sup_{x \in \tilde{M}} \sup_{\substack{\alpha, \beta \in \tilde{F}_x \\ \alpha \perp \beta, |\alpha \wedge \beta|=1}} |R^E(\alpha \wedge \beta)|.$$

Now, we can define a pair of closely related concepts.

Definition 1.1 With above notation, if the vector bundle E satisfies that some Chern number of E is nonzero, we define the (covering) K-cowaist of (M, F) by

$$(1-2) \quad \text{K-ccw}_2(M, F) = \sup_{\tilde{M}, E} (\|R_{\tilde{F}}^E\|^{-1}).$$

Similarly, if the vector bundle E satisfies the inequality

$$\int_{\tilde{M}} \hat{A}(T\tilde{M})(\text{ch}(E) - \text{rk}(E)) \neq 0,$$

we define the (covering) $\hat{A}$ -cowaist of (M, F) by

$$(1-3) \quad \hat{A}\text{-ccw}_2(M, F) = \sup_{\tilde{M}, E} (\|R_{\tilde{F}}^E\|^{-1}).$$

From the index theory viewpoint, among the two concepts given as above, the $\hat{A}$ -cowaist perhaps relates to the scalar curvature more directly. In fact, we can use the $\hat{A}$ -cowaist to give a quantitative estimate on the leafwise scalar curvature.

²In the literature, there are different ways to define the norm of R^E (or $R_{\tilde{F}}^E$); see eg [3; 12]. Here, we adopt the same definition as in [14]. However, for the content discussed in this paper, all these norms work.

Theorem 1.2 *Let M be a connected oriented (not necessarily compact) manifold carrying a complete Riemannian metric g^{TM} . Let $F \subseteq TM$ be an integrable subbundle of TM with the restricted metric $g^F = g^{TM}|_F$. Let k^F be the associated leafwise scalar curvature of F . If either TM or F is spin, then*

$$\inf(k^F) \leq \frac{2 \operatorname{rk}(F)(\operatorname{rk}(F) - 1)}{\widehat{\mathbf{A}}\text{-ccw}_2(M, F)}.$$

Note that in the definition of $\widehat{\mathbf{A}}$ -cowaist, we don't need that g^{TM} is complete. But for the above theorem, the completeness is necessary. As in [15], we only give the proof of Theorem 1.2 for the TM spin case in detail. The F spin case can be proved similarly as in [15, Section 2.5].

To obtain a similar estimate using the K-cowaist, we notice the following reinterpretation of the result in [7, Section 5 $\frac{3}{8}$]. One can also see [3], [12] or [14] for a more detailed proof of this result.

Proposition 1.3 $\mathbf{K}\text{-ccw}_2(M, F) = +\infty$ implies $\widehat{\mathbf{A}}\text{-ccw}_2(M, F) = +\infty$.

Due to this proposition, we have the following corollary of Theorem 1.2.

Theorem 1.4 *Under the same assumptions of Theorem 1.2, if we further assume $\mathbf{K}\text{-ccw}_2(M, F) = +\infty$, then $\inf(k^F) \leq 0$.*

Since $\mathbf{K}\text{-cw}_2(M) = +\infty$ implies $\mathbf{K}\text{-ccw}_2(M, F) = +\infty$, as a corollary of Theorem 1.4, one can show the following result.

Theorem 1.5 (Gromov [9, page 258, footnote 277]) *Complete manifolds X with infinite $\mathbf{K}\text{-cw}_2(X)$, carry no spin foliations F , where the induced Riemannian metrics in the leaves satisfy $k^F \geq \sigma > 0$.*

1.2 A discussion about the definitions of $\widehat{\mathbf{A}}$ -cowaist

Compared to other similar concepts in the literature, a feature of the definition of $\widehat{\mathbf{A}}$ -cowaist is that the supremum is calculated using bundles over any coverings of M rather than bundles over M alone. Whether it is necessary to take the supremum on this larger set turns out to be a delicate problem. For simplicity, we will assume $F = TM$ in this subsection.

To facilitate our discussion, we use the following definition which resembles the definition of $\widehat{\mathbf{A}}$ -cowaist but does not use the covering space. Let M be a manifold and (E, g^E, ∇^E) be a Hermitian vector bundle over M with the Hermitian metric g^E and the Hermitian connection ∇^E . If M is noncompact, we also assume that E is trivial at infinity. Define

$$\widehat{\mathbf{A}}\text{-area}(M) = \sup_E \{ \|R^E\|^{-1} \mid \int_M \widehat{\mathbf{A}}(TM)(\operatorname{ch}(E) - \operatorname{rk}(E)) \neq 0 \}.$$

We discuss several cases separately.

Case 1: M is noncompact In this case, $\hat{\text{A}}\text{-ccw}_2(M, TM)$ is strictly larger than $\hat{\text{A}}\text{-area}(M)$ in general. As an example, let M be an annulus in $\mathbb{R}^2$ and $F = TM$. By [8, page 33 (c)], we have

$$(1-4) \quad \text{K-cw}_2(M) = \text{area}(M).$$

Since $\dim M = 2$, for any vector bundle E over M , we have

$$\int_M \hat{\text{A}}(TM)(\text{ch}(E) - \text{rk}(E)) = \int_M c_1(E).$$

Hence, $\text{K-cw}_2(M)$ and $\hat{\text{A}}\text{-area}(M)$ coincide in this case. Using (1-4), we can see that

$$\hat{\text{A}}\text{-ccw}_2(M, TM) = \sup_{\tilde{M}} \text{area}(\tilde{M}) = +\infty > \hat{\text{A}}\text{-area}(M).$$

Case 2: M is compact and the universal covering of M is compact First, we note that if M is compact, $\hat{\text{A}}\text{-area}(M)$ is just a small variant of [5, Definition 1.6].³

Since the universal covering of M is compact, any covering space of M , $\tilde{M}$, is also compact. Then by the same proof of [8, page 33 (d)], we have

$$\hat{\text{A}}\text{-area}(M) = \hat{\text{A}}\text{-area}(\tilde{M}).$$

Therefore, in this case, $\hat{\text{A}}\text{-ccw}_2(M, TM)$ is equal to $\hat{\text{A}}\text{-area}(M)$.

Case 3: M is compact and the universal covering of M is noncompact This case is the most difficult and we do not have a definite answer at the moment. In fact, whether $\hat{\text{A}}\text{-area}(M)$ is equal to $\hat{\text{A}}\text{-ccw}_2(M, TM)$ in this case relates closely to Gromov's question [8, page 34, Question 23]: is there a closed manifold M such that $\text{K-cw}_2(M) < \infty$ and the universal covering of M satisfies $\text{K-cw}_2(\tilde{M}) = \infty$?

The main difficulty in this case is that the pull-back and push-forward construction for vector bundles do not work well for the noncompact spaces. If we put some restrictions on the covering spaces, maybe some partial results are still possible. For example, if we assume $\pi_1(M)$ is residually finite, we have the following simple extension of [7, page 26, (v')].

Proposition 1.6 *If M is a closed manifold and $\pi_1(M)$ is residually finite, for the universal covering space $\tilde{M}$ of M ,*

$$\text{K-ccw}_2(M) \geq \text{K-cw}_2(M) \geq \text{K-cw}_2(\tilde{M}) = \text{K-ccw}_2(\tilde{M}).$$

Proof Clearly, we only need to show that

$$(1-5) \quad \text{K-cw}_2(M) \geq \text{K-cw}_2(\tilde{M}).$$

If $\tilde{M}$ is compact, (1-5) follows from the push-forward inequality [7, Section 4 $\frac{3}{5}$]. We try to show that the push-forward argument also works for noncompact $\tilde{M}$.

Take a vector bundle E over $\tilde{M}$ which is trivial outside a compact set K . Since $\pi_1(M)$ is residually finite, we can find a finite covering of M , N , such that the covering map from $\tilde{M}$ to N is injective on K .

³In fact, if M carries a metric with the positive scalar curvature, these two definitions are the same.

As a result, we can push forward the vector bundle E to a vector bundle E_N over N . Since N is a finite covering of M , we have

$$\text{K-cw}_2(N) = \text{K-cw}_2(M).$$

Therefore,

$$\|R^E\|^{-1} = \|R^{E_N}\|^{-1} \leq \text{K-cw}_2(N) = \text{K-cw}_2(M),$$

from which we obtain (1-5). $\square$

We also note that the K-cowaist also generalizes [4, Definition 5.1].

2 Proof of Theorem 1.2

In this section we prove Theorem 1.2. Our strategy closely follows the proof of [13, Theorem 1.2]. Note that in [13], there is a map $f: M \rightarrow S^n(1)$, which enables us to construct suitable bundles over a closed manifold associated with the noncompact manifold M . Compared to [13], the new idea in this paper is that we will show that in the current situation, the auxiliary bundles needed for the proof, as well as the endomorphisms between the these bundles, can be constructed *without using the map f* .

We argue by contradiction. Assume that

$$\inf(k^F) > \frac{2\text{rk}(F)(\text{rk}(F) - 1)}{\hat{\text{A-ccw}}_2(M, F)}.$$

Then, by the definition of $\hat{\text{A-ccw}}_2(M, F)$, there exists

- a covering manifold $\tilde{\pi}: \tilde{M} \rightarrow M$ with the lifted foliation $\tilde{F}$ and the lifted metrics $g^{T\tilde{M}}$ and $g^{\tilde{F}}$;
- a Hermitian vector bundle E_0 over $\tilde{M}$ with the Hermitian metric g^{E_0} and the Hermitian connection ∇^{E_0} , which is trivial at infinity and satisfies

$$(2-1) \quad \int_{\tilde{M}} \hat{\text{A}}(T\tilde{M})(\text{ch}(E_0) - \text{rk}(E_0)) \neq 0;$$

- a constant $\kappa > 0$ such that

$$(2-2) \quad \tilde{\pi}^*(k^F) - 2\text{rk}(F)(\text{rk}(F) - 1)\|R_{\tilde{F}}^{E_0}\| > \kappa \quad \text{on } \tilde{M}.$$

As explained in Introduction, we only prove the TM spin case in detail. In the following, we assume that TM is spin.

To begin, we note that if both M and $\tilde{M}$ are compact, by [16, Section 1.1], one gets a contradiction easily. Therefore, in the following, we assume that $\tilde{M}$ is noncompact. For the rest of the proof we will only deal with quantities associated with $\tilde{M}$ and $\tilde{F}$. To simplify the notation, we will denote the foliation $(\tilde{M}, \tilde{F})$ by (M, F) and the metrics $(g^{T\tilde{M}}, g^{\tilde{F}})$ by (g^{TM}, g^F) .

Roughly speaking, we will prove Theorem 1.2 in three steps.

- We construct a closed manifold $\hat{M}_{H_{3m}}$ and a $\mathbb{Z}_2$ -graded bundle $\hat{E}$ over it. We also construct a fiber bundle $\hat{M}_{\mathcal{H}_{3m}, R}$ over $\hat{M}_{H_{3m}}$ associated with the foliation F .

- (ii) We construct a deformed Dirac operator on $\widehat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ and obtain some estimates about it.
- (iii) We construct a closed manifold using $\widehat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ and an operator $P_{R,\beta,\gamma,+}^{\mathcal{E}_{3m},R}$ using the deformed Dirac operator. We will show that (2-1) implies the index of $P_{R,\beta,\gamma,+}^{\mathcal{E}_{3m},R}$ is not zero while (2-2) implies the index of $P_{R,\beta,\gamma,+}^{\mathcal{E}_{3m},R}$ is zero. Thus we obtain a contradiction.

As we have said at the beginning of this section, the main difference between the proof of Theorem 1.2 and [13, Theorem 1.2] is the first step.

Step 1 Let $(E_1 = M \times \mathbb{C}^k, g^{E_1}, \nabla^{E_1})$, with $k = \text{rk}(E_0)$, be the trivial vector bundle on M . Then, let $E = E_0 \oplus E_1$ be a $\mathbb{Z}_2$ -graded Hermitian vector bundle over M with a $\mathbb{Z}_2$ -graded metric $g^E = g^{E_0} \oplus g^{E_1}$ and a $\mathbb{Z}_2$ -graded Hermitian connection $\nabla^E = \nabla^{E_0} \oplus \nabla^{E_1}$.

Since $(E_0, g^{E_0}, \nabla^{E_0})$ is trivial at infinity, there exists a compact subset K^4 and an isomorphism

$$(2-3) \quad \psi: (E_0|_{M \setminus K}, g^{E_0}, \nabla^{E_0}) \rightarrow ((M \setminus K) \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}}) = (E_1|_{M \setminus K}, g^{E_1}, \nabla^{E_1}).$$

Following [10, Theorem 1.17], we choose a fixed point $x_0 \in M$ and let $d: M \rightarrow \mathbb{R}^+$ be a regularization of the distance function $\text{dist}(x, x_0)$ such that

$$|\nabla d|(x) \leq \frac{3}{2},$$

for any $x \in M$.

Set

$$B_m = \{x \in M \mid d(x) \leq m\}, \quad m \in \mathbb{N}$$

and choose a sufficiently large m such that $K \subset B_m$. Note that B_m is compact by completeness of g^{TM} .

To construct the desired closed manifold $\widehat{M}_{H_{3m}}$, following [10], we take a compact hypersurface $H_{3m} \subseteq M \setminus B_{3m}$, which cuts M into two parts such that the compact part, denoted by $M_{H_{3m}}$, contains B_{3m} . Then $M_{H_{3m}}$ is a compact smooth manifold with boundary H_{3m} . Let $g^{TH_{3m}}$ be the induced metric on H_{3m} . For a sufficiently small $\varepsilon' > 0$, on the product manifold $H_{3m} \times [-\varepsilon', 1 + \varepsilon']$, we construct a metric as follows.

Near the boundary $H_{3m} \times \{-\varepsilon'\}$ of $H_{3m} \times [-\varepsilon', 1 + \varepsilon']$, ie on $H_{3m} \times [-\varepsilon', \varepsilon']$, by using the geodesic normal coordinate of $H_{3m} \subseteq M$, for a small ε' , there is an isomorphism between $H_{3m} \times [-\varepsilon', \varepsilon']$ and a neighborhood of H_{3m} , denoted by U , in M because H_{3m} is compact. Moreover, we can require that under this isomorphism, $U \cap M_{H_{3m}}$ is mapped to $H_{3m} \times [-\varepsilon', 0]$.

Now, we define the metric on $H_{3m} \times [-\varepsilon', \varepsilon']$ to be the pull-back metric obtained from that of U . In the same way, we can construct a metric on $H_{3m} \times [1 - \varepsilon', 1 + \varepsilon']$. Meanwhile, the metric on $H_{3m} \times [1/3, 2/3]$ is defined to be the product metric of $g^{TH_{3m}}$ and the standard metric on $[1/3, 2/3]$. Finally, the metric on $H_{3m} \times [-\varepsilon', 1 + \varepsilon']$ is a smooth extension of the metrics on the above three pieces.

⁴We can and will choose K to be the closure of an open subset.

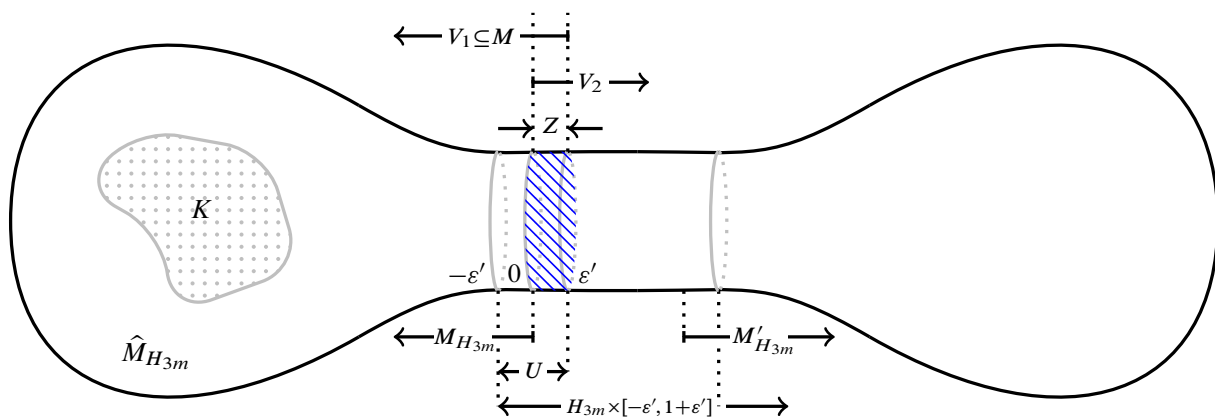


Figure 1: Gluing construction.

Using the isometry between $H_{3m} \times [-\varepsilon', \varepsilon']$ and U , $M_{H_{3m}} \cup U$ and $H_{3m} \times [-\varepsilon', 1 + \varepsilon']$ can be glued into a smooth Riemannian manifold with boundary,

$$(M_{H_{3m}} \cup U) \bigcup_U (H_{3m} \times [-\varepsilon', 1 + \varepsilon']).$$

More precisely, the resulting manifold is a quotient space of $(M_{H_{3m}} \cup U) \bigcup (H_{3m} \times [-\varepsilon', 1 + \varepsilon'])$, two points of $M_{H_{3m}} \cup U$ and $H_{3m} \times [-\varepsilon', 1 + \varepsilon']$, respectively, correspond to the same point in the resulting manifold if and only if the two points are related by the isomorphism between U and $H_{3m} \times [-\varepsilon', \varepsilon']$.

Let $M'_{H_{3m}}$ be another copy of $M_{H_{3m}}$ with the same metric and the opposite orientation. In a similar way, we can glue $M'_{H_{3m}}$ and $(M_{H_{3m}} \cup U) \bigcup_U (H_{3m} \times [-\varepsilon', 1 + \varepsilon'])$ to obtain a closed manifold $\hat{M}_{H_{3m}}$. Note that $M_{H_{3m}} \cup U$, $M'_{H_{3m}}$ and $H_{3m} \times [1/3, 2/3]$ all have natural isometric embeddings into $\hat{M}_{H_{3m}}$. As a result, we will view these manifolds as submanifolds of $\hat{M}_{H_{3m}}$ in the following. Figure 1 is an illustration of this gluing construction.

Take $V_1 := M_{H_{3m}} \cup U^\circ$ and $V_2 = \hat{M}_{H_{3m}} \setminus M_{H_{3m}}$ to be two open subsets of $\hat{M}_{H_{3m}}$. Then

$$V_1 \cup V_2 = \hat{M}_{H_{3m}}, \quad V_1 \cap V_2 = Z := H_{3m} \times (0, \varepsilon').$$

As we have said, we will treat $V_1 = M_{H_{3m}} \cup U^\circ$ as a submanifold of both M and $\hat{M}_{H_{3m}}$. This means that, although E is a $\mathbb{Z}_2$ -graded vector bundle defined over M , its restriction on $M_{H_{3m}} \cup U^\circ$, that is $E|_{M_{H_{3m}} \cup U^\circ} = E|_{V_1}$, is a $\mathbb{Z}_2$ -graded vector bundle defined over a submanifold of $\hat{M}_{H_{3m}}$. We are going to extend $E|_{V_1}$ to a $\mathbb{Z}_2$ -graded bundle $\hat{E} = \hat{E}_0 \oplus \hat{E}_1$ over $\hat{M}_{H_{3m}}$ which satisfies

$$(2-4) \quad (\hat{E}_0|_{V_2}, g^{\hat{E}_0}, \nabla^{\hat{E}_0}) \simeq (V_2 \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}}) \simeq (\hat{E}_1|_{V_2}, g^{\hat{E}_1}, \nabla^{\hat{E}_1}).$$

We will construct the $\hat{E}_0$ and $\hat{E}_1$ separately. The construction of $\hat{E}_1$ is straightforward. Since $E_1|_{M_{H_{3m}}} = M_{H_{3m}} \times \mathbb{C}^k$, we can take $(\hat{E}_1, g^{\hat{E}_1}, \nabla^{\hat{E}_1})$ to be $(\hat{M}_{H_{3m}} \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}})$, which satisfies (2-4).

To construct $(\hat{E}_0, g^{\hat{E}_0}, \nabla^{\hat{E}_0})$, we glue two vector bundles as in [1]. Choose the trivial bundle

$$(E'_0, g^{E'_0}, \nabla^{E'_0}) := (V_2 \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}})$$

over V_2 . Recall that $Z \subset U$ can be viewed as a submanifold of M and $Z \cap K = \emptyset$. Hence, by the definition of ψ , (2-3), we have an isomorphism between $(E_0|_Z, g^{E_0}, \nabla^{E_0})$ (ie $((E_0|_{V_1})|_Z, g^{E_0}, \nabla^{E_0})$) and $(Z \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}})$ (ie $(E'_0|_Z, g^{E'_0}, \nabla^{E'_0})$). In other words, the restriction of ψ on Z , denoted by $\psi|_Z$, induces an isomorphism between $((E_0|_{V_1})|_Z, g^{E_0}, \nabla^{E_0})$ and $(E'_0|_Z, g^{E'_0}, \nabla^{E'_0})$. We define $\hat{E}_0$ to be

$$\hat{E}_0 := E_0|_{V_1} \bigcup_{\psi|_Z} E'_0|_{V_2}.$$

By definition, ψ , thus $\psi|_Z$, preserves the metric and the connection. As a result, $\hat{E}_0$ inherits a metric and a connection from those of $E_0|_{V_1}$ and $E'_0|_{V_2}$. Moreover, the property of gluing construction implies

$$\begin{aligned} (\hat{E}_0|_{M_{H_{3m}}}, g^{\hat{E}_0}, \nabla^{\hat{E}_0}) &\simeq (E_0|_{M_{H_{3m}}}, g^{E_0}, \nabla^{E_0}), \\ (\hat{E}_0|_{V_2}, g^{\hat{E}_0}, \nabla^{\hat{E}_0}) &\simeq (E'_0, g^{E'_0}, \nabla^{E'_0}) = (V_2 \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}}). \end{aligned}$$

Therefore, $\hat{E}_0$ also satisfies (2-4).

Let $i: E_0|_{V_1} \hookrightarrow \hat{E}_0$ and $j: E'_0|_{V_2} \hookrightarrow \hat{E}_0$ be the canonical embeddings in the gluing construction. The definition of gluing construction implies the composition of maps

$$j|_Z^{-1} \circ i|_Z: (E_0|_Z, g^{E_0}, \nabla^{E_0}) \xrightarrow{i|_Z} (\hat{E}_0|_Z, g^{\hat{E}_0}, \nabla^{\hat{E}_0}) \xrightarrow{j|_Z^{-1}} (E'_0|_Z, g_{\text{st}}, \nabla_{\text{st}}) = (Z \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}})$$

is just $\psi|_Z$. Since $\hat{E}_1|_{V_2} = V_2 \times \mathbb{C}^k$, we can define

$$\nu := j^{-1}: (\hat{E}_0|_{V_2}, g^{\hat{E}_0}, \nabla^{\hat{E}_0}) \xrightarrow{\sim} (E'_0, g^{E'_0}, \nabla^{E'_0}) = (V_2 \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}}) = (\hat{E}_1|_{V_2}, g^{\hat{E}_1}, \nabla^{\hat{E}_1}).$$

At the same time, ψ induces the map

$$\begin{aligned} \psi|_{V_1} \circ i^{-1}: (\hat{E}_0|_{V_1}, g^{\hat{E}_0}, \nabla^{\hat{E}_0}) &\xrightarrow{\sim} (E_0|_{V_1}, g^{E_0}, \nabla^{E_0}) \rightarrow (E_1|_{V_1}, g^{E_1}, \nabla^{E_1}) = (V_1 \times \mathbb{C}^k, g_{\text{st}}, \nabla_{\text{st}}) \\ &= (\hat{E}_1|_{V_1}, g^{\hat{E}_1}, \nabla^{\hat{E}_1}). \end{aligned}$$

Therefore, we have

$$(2-5) \quad (\psi|_{V_1} \circ i^{-1})|_Z = \psi|_Z \circ i|_Z^{-1} = j|_Z^{-1} \circ i|_Z \circ i|_Z^{-1} = j|_Z^{-1} = \nu|_Z.$$

Note that the bundle endomorphism $\psi|_{V_1} \circ i^{-1}$ is just a smooth section of $\text{Hom}(\hat{E}_0, \hat{E}_1)$ over V_1 . Similarly, the bundle endomorphism ν is a smooth section of $\text{Hom}(\hat{E}_0, \hat{E}_1)$ over V_2 . Consequently, by (2-5), the following section ω of $\text{Hom}(\hat{E}_0, \hat{E}_1)$ is well defined and smooth:

$$\omega_x = \begin{cases} (\psi|_{V_1} \circ i^{-1})_x & \text{if } x \in V_1, \\ \nu_x & \text{if } x \in V_2. \end{cases}$$

Moreover, by the property of ψ and ν , we know that ω preserves the metrics and connections on $\hat{M}_{H_{3m}} \setminus K$.

Take ω^* to be the adjoint of ω with respect to $g^{\hat{E}_0}$ and $g^{\hat{E}_1}$. Set

$$W = \omega + \omega^*: \Gamma(\hat{E}) \rightarrow \Gamma(\hat{E}),$$

which is an odd and self-adjoint bundle endomorphism of $\hat{E}$. There exists a constant $\delta > 0$ such that

$$(2-6) \quad W^2 \geq \delta \quad \text{on } \hat{M}_{H_{3m}} \setminus K.$$

After the construction of $\hat{M}_{H_{3m}}$ and $\hat{E}$, we now explain how to construct a fiber bundle $\hat{\mathcal{M}}_{\mathcal{H}_{3m}, R}$ over $\hat{M}_{H_{3m}}$ associated with the foliation F . Let $F^\perp$ be the orthogonal complement to F , ie we have the orthogonal splitting

$$(2-7) \quad TM = F \oplus F^\perp, \quad g^{TM} = g^F \oplus g^{F^\perp}.$$

Following [6, Section 5] (also cf [15, Section 2.1]), let $\pi: \mathcal{M} \rightarrow M$ be the Connes fibration over M such that for any $x \in M$, $\mathcal{M}_x = \pi^{-1}(x)$ is the space of Euclidean metrics on the linear space $T_x M / F_x$. Let $T^V \mathcal{M}$ denote the vertical tangent bundle of the fibration $\pi: \mathcal{M} \rightarrow M$. Then it carries a natural metric $g^{T^V \mathcal{M}}$ such that any two points $p, q \in \mathcal{M}_x$ with $x \in M$ can be joined by a unique geodesic along $\mathcal{M}_x$. Let $d^{\mathcal{M}_x}(p, q)$ denote the length of this geodesic.

By using the Bott connection on TM/F (cf [15, (1.2)]), which is leafwise flat, one lifts F to an integrable subbundle $\mathcal{F}$ of $T\mathcal{M}$. Then g^F lifts to a Euclidean metric $g^{\mathcal{F}} = \pi^* g^F$ on $\mathcal{F}$.

Let $\mathcal{F}_1^\perp \subseteq T\mathcal{M}$ be a subbundle, which is transversal to $\mathcal{F} \oplus T^V \mathcal{M}$, such that we have a splitting

$$T\mathcal{M} = (\mathcal{F} \oplus T^V \mathcal{M}) \oplus \mathcal{F}_1^\perp.$$

Then $\mathcal{F}_1^\perp$ can be identified with $T\mathcal{M}/(\mathcal{F} \oplus T^V \mathcal{M})$ and carries a canonically induced metric $g^{\mathcal{F}_1^\perp}$. We let $\mathcal{F}_2^\perp$ denote $T^V \mathcal{M}$.

The metric $g^{F^\perp}$ in (2-7) determines a canonical embedded section $s: M \hookrightarrow \mathcal{M}$. For any $p \in \mathcal{M}$, set $\rho(p) = d^{\mathcal{M}_{\pi(p)}}(p, s(\pi(p)))$.

For any $\beta, \gamma > 0$, following [15, (2.15)], let $g_{\beta, \gamma}^{T\mathcal{M}}$ be the metric on $T\mathcal{M}$ defined by the orthogonal splitting

$$(2-8) \quad T\mathcal{M} = \mathcal{F} \oplus \mathcal{F}_1^\perp \oplus \mathcal{F}_2^\perp, \quad g_{\beta, \gamma}^{T\mathcal{M}} = \beta^2 g^{\mathcal{F}} \oplus \frac{g^{\mathcal{F}_1^\perp}}{\gamma^2} \oplus g^{\mathcal{F}_2^\perp}.$$

For any $R > 0$, let $\mathcal{M}_R$ be the smooth manifold with boundary defined by

$$\mathcal{M}_R = \{p \in \mathcal{M} \mid \rho(p) \leq R\}.$$

Set $\mathcal{H}_{3m} = \pi^{-1}(H_{3m})$ and

$$\mathcal{M}_{\mathcal{H}_{3m}, R} = (\pi^{-1}(M_{H_{3m}})) \cap \mathcal{M}_R, \quad \mathcal{H}_{3m, R} = \mathcal{H}_{3m} \cap \mathcal{M}_R.$$

Consider another copy $\mathcal{M}'_{\mathcal{H}_{3m}, R}$ of $\mathcal{M}_{\mathcal{H}_{3m}, R}$ carrying the metric $g^{T\mathcal{M}'_{3m}, R}$ defined by (2-8) with $\beta = \gamma = 1$. Meanwhile, let $g^{T\mathcal{H}_{3m}, R}$ be the induced metric on $\mathcal{H}_{3m, R}$ by (2-8) with $\beta = \gamma = 1$ and dt^2 be the standard metric on $[1/3, 2/3]$. As we have done for $\hat{M}_{H_{3m}}$, we can glue $\mathcal{M}_{\mathcal{H}_{3m}, R}$, $\mathcal{M}'_{\mathcal{H}_{3m}, R}$ and $\mathcal{H}_{3m, R} \times [-\varepsilon', 1 + \varepsilon']$

together to get a manifold $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$; cf [13, Section 2.2]. But, unlike $\hat{M}_{H_{3m}}$, $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ is a smooth manifold with boundary. Moreover, we can define a smooth metric $g^{T\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}}$ on $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ such that

$$g^{T\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}}|_{\mathcal{M}_{\mathcal{H}_{3m},R}} = g_{\beta,\gamma}^{T\mathcal{M}_{3m,R}}, \quad g^{T\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}}|_{\mathcal{M}'_{\mathcal{H}_{3m},R}} = g^{T\mathcal{M}'_{3m,R}},$$

$$g^{T\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}}|_{\mathcal{H}_{3m,R} \times [1/3, 2/3]} = g^{T\mathcal{H}_{3m,R}} \oplus dt^2.$$

The map $\pi: \mathcal{M}_{\mathcal{H}_{3m},R} \rightarrow M_{H_{3m}}$ can be extended to $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R} \rightarrow \hat{M}_{H_{3m}}$ and we still denote the extended map by π . As before, we pull the bundles on $\hat{M}_{H_{3m}}$ back to $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$, that is, we take

$$(\mathcal{E}_{3m,R}, \nabla^{\mathcal{E}_{3m,R}}, g^{\mathcal{E}_{3m,R}}) = \pi^*(\hat{E}, \nabla^{\hat{E}}, g^{\hat{E}}).$$

As usual, $R^{\mathcal{E}_{3m,R}} = (\nabla^{\mathcal{E}_{3m,R}})^2$ is the curvature of $\nabla^{\mathcal{E}_{3m,R}}$.

Step 2 Recall that we have assumed that TM is oriented and spin, which implies that $\mathcal{F} \oplus \mathcal{F}_1^\perp = \pi^*(TM)$ is spin. Without loss of generality, as in [15, pages 1062–1063], we assume further that F is oriented and $\text{rk}(F^\perp)$ is divisible by 4. Then $F^\perp$ is also oriented and $\dim \mathcal{M}$ is even.

It is clear that $\mathcal{F} \oplus \mathcal{F}_1^\perp, \mathcal{F}_2^\perp$ over $\mathcal{M}_{\mathcal{H}_{3m},R}$ can be extended to $(\mathcal{H}_{3m,R} \times [-\varepsilon', 1 + \varepsilon']) \cup \mathcal{M}'_{\mathcal{H}_{3m},R}$ such that we have the orthogonal splitting⁵

$$(2-9) \quad T\hat{\mathcal{M}}_{\mathcal{H}_{3m},R} = (\mathcal{F} \oplus \mathcal{F}_1^\perp) \oplus \mathcal{F}_2^\perp \quad \text{on } \hat{\mathcal{M}}_{\mathcal{H}_{3m},R}.$$

Let $S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp)$ denote the spinor bundle over $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ with respect to the metric $g^{T\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}}|_{\mathcal{F} \oplus \mathcal{F}_1^\perp}$ (thus with respect to $\beta^2 g^{\mathcal{F}} \oplus (g^{\mathcal{F}_1^\perp}/\gamma^2)$ on $\mathcal{M}_{\mathcal{H}_{3m},R}$). Let $\Lambda^*(\mathcal{F}_2^\perp)$ denote the exterior algebra bundle of $\mathcal{F}_2^{\perp,*}$, with the $\mathbb{Z}_2$ -grading given by the natural even/odd parity.

Let

$$(2-10) \quad D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}: \Gamma(S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp)) \rightarrow \Gamma(S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp))$$

be the sub-Dirac operator on $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ constructed as in [15, (2.16)]. It is clear that one can canonically define the twisted sub-Dirac operator (twisted by $\mathcal{E}_{3m,R}$) on $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$,

$$(2-11) \quad D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}}: \Gamma(S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{3m,R}) \rightarrow \Gamma(S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{3m,R}).$$

Let $\tilde{f}: [0, 1] \rightarrow [0, 1]$ be a smooth function such that $\tilde{f}(t) = 0$ for $0 \leq t \leq \frac{1}{4}$, while $\tilde{f}(t) = 1$ for $\frac{1}{2} \leq t \leq 1$. For any $p \in \mathcal{M}_{\mathcal{H}_{3m},R}$, we connect p and $s(\pi(p))$ by the unique geodesic in $\mathcal{M}_{\pi(p)}$. Let $\sigma(p) \in \mathcal{F}_2^\perp|_p$ denote the unit vector tangent to this geodesic. Then

$$\tilde{\sigma} = \tilde{f} \left(\frac{\rho}{R} \right) \sigma$$

is a smooth section of $\mathcal{F}_2^\perp|_{\mathcal{M}_{\mathcal{H}_{3m},R}}$. It extends to a smooth section of $\mathcal{F}_2^\perp|_{\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}}$, which we still denote by $\tilde{\sigma}$. It is easy to see that we may and we will assume that $\tilde{\sigma}$ is transversal to (and thus nowhere zero on) $\partial\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$. Note that the Clifford action $\hat{c}(\tilde{\sigma})$ (cf [15, (1.47)]) now acts on $S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{3m,R}$ over $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$.

⁵ $\mathcal{F}$ restricted to $(\mathcal{H}_{3m,R} \times [-\varepsilon', 1 + \varepsilon']) \cup \mathcal{M}'_{\mathcal{H}_{3m},R}$ needs no longer to be integrable.

With $\hat{c}(\tilde{\sigma})$ and W , for $\varepsilon > 0$, we introduce the following deformation of $D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m, R}}$ on $\hat{\mathcal{M}}_{\mathcal{H}_{3m, R}}$, which combines the deformations in [15, (2.21)] and [16, (1.11)]:

$$(2-12) \quad D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m, R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta}.$$

For the deformed operator (2-12), the following estimate holds, which is an analog of [13, Lemma 2.1]. Let $h: [0, 1] \rightarrow [0, 1]$ be a smooth function such that $h(t) = 1$ for $0 \leq t \leq \frac{3}{4}$, while $h(t) = 0$ for $\frac{7}{8} \leq t \leq 1$.

Lemma 2.1 *There exist $c_0, \varepsilon, m, R > 0$ such that, when $\beta, \gamma > 0$ are small enough (which may depend on m and R),*

- (i) *for any $s \in \Gamma(S_{\beta, \gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{3m, R})$ supported in the interior of $\hat{\mathcal{M}}_{\mathcal{H}_{3m, R}}$,*⁶

$$\left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m, R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta} \right) s \right\| \geq \frac{c_0}{\beta} \|s\|;$$

- (ii) *for any $s \in \Gamma(S_{\beta, \gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{3m, R})$ supported in the interior of $\mathcal{M}_{\mathcal{H}_{3m, R}} \setminus \mathcal{M}_{\mathcal{H}_{3m, R}/2}$,*

$$\left\| \left(h\left(\frac{\rho}{R}\right) D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m, R}} h\left(\frac{\rho}{R}\right) + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta} \right) s \right\| \geq \frac{c_0}{\beta} \|s\|.$$

Proof We follow the same strategy as [13, Lemma 2.1] to prove this lemma. Especially, the proof of (ii) is the same as the proof of [13, Lemma 2.1(ii)]. Here, we only show how to modify the arguments in [13, Lemma 2.1] to prove (i).

As in [13, (2.14)–(2.15)], on $\hat{M}_{H_{3m}}$, by using the regularized distance function $d(x)$, we can find cut-off functions $\psi_{m,1}, \psi_{m,2}: \hat{M}_{H_{3m}} \rightarrow [0, 1]$ satisfying, for $i = 1, 2$,

$$(2-13) \quad \psi_{m,1}^2 + \psi_{m,2}^2 = 1,$$

$$(2-14) \quad |\nabla \psi_{m,i}|(x) \leq \frac{C}{m} \quad \text{for any } x \in \hat{M}_{H_{3m}}.$$

Then we pull $\psi_{m,1}$ and $\psi_{m,2}$ back to cut-off functions $\varphi_{m,1}$ and $\varphi_{m,2}$ defined on $\hat{\mathcal{M}}_{\mathcal{H}_{3m, R}}$. By [13, (2.24)], we know that

$$(2-15) \quad \varphi_{m,1} = \begin{cases} 1 & \text{if } x \in \pi^{-1}(B_m), \\ 0 & \text{if } x \in \hat{\mathcal{M}}_{\mathcal{H}_{3m, R}} \setminus \pi^{-1}(B_{2m}), \end{cases} \quad \varphi_{m,2} = \begin{cases} 0 & \text{if } x \in \pi^{-1}(B_m), \\ 1 & \text{if } x \in \hat{\mathcal{M}}_{\mathcal{H}_{3m, R}} \setminus \pi^{-1}(B_{2m}). \end{cases}$$

Using $\varphi_{m,i}$, for any $s \in \Gamma(S_{\beta, \gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{3m, R})$ supported in the interior of $\hat{\mathcal{M}}_{\mathcal{H}_{3m, R}}$, by (2-13), we have the following estimate (cf [13, (2.26)]):

$$(2-16) \quad \begin{aligned} & \sqrt{2} \left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m, R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta} \right) s \right\| \\ & \geq \left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m, R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta} \right) (\varphi_{m,1} s) \right\| + \left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m, R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta} \right) (\varphi_{m,2} s) \right\| \\ & \quad - \|c_{\beta, \gamma}(d\varphi_{m,1})s\| - \|c_{\beta, \gamma}(d\varphi_{m,2})s\|, \end{aligned}$$

⁶The norms below depend on β and γ . If there is no confusion, we omit the subscripts for simplicity.

where for each $i \in \{1, 2\}$, we identify $d\varphi_{m,i}$ with the gradient of $\varphi_{m,i}$ and $c_{\beta,\gamma}(\cdot)$ means the Clifford action with respect to the metric (2-8).

For the last two terms in the right-hand side of (2-16), we can use the estimate [13, (2.30)],

$$(2-17) \quad |c_{\beta,\gamma}(d\varphi_{m,i})s|(x) = \left(O\left(\frac{1}{\beta m}\right) + O_{m,R}(\gamma)\right)|s|(x), \quad x \in \mathcal{M}_{\mathcal{H}_{3m},R},$$

where the subscripts in $O_{m,R}(\cdot)$ mean that the big O constant may depend on m and R .

For the first two terms in the right-hand side of (2-16), by a direct computation, we have

$$(2-18) \quad \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta}\right)^2 = \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta}\right)^2 + \left[D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}}, \frac{\varepsilon \pi^* W}{\beta}\right] + \frac{\varepsilon^2 (\pi^* W)^2}{\beta^2}.$$

Meanwhile, since W is a constant endomorphism, ie parallel with respect to the connection, outside K , we know

$$(2-19) \quad \left[D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}}, \frac{\varepsilon \pi^* W}{\beta}\right] = 0 \quad \text{on } \hat{\mathcal{M}}_{\mathcal{H}_{3m},R} \setminus \pi^{-1}(K).$$

Therefore, for the second term on the right-hand side of (2-16), by (2-6), (2-15), (2-18) and (2-19), one has

$$(2-20) \quad \left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta}\right)(\varphi_{m,2s}) \right\|^2 \\ = \left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta}\right)(\varphi_{m,2s}) \right\|^2 + \left\| \frac{\varepsilon \pi^* W}{\beta}(\varphi_{m,2s}) \right\|^2 \geq \frac{\delta \varepsilon^2}{\beta^2} \|\varphi_{m,2s}\|^2.$$

The main difference between the proof of (i) and [13, Lemma 2.1(i)] lies in the estimate of the first term in the right-hand side of (2-16). Let $\text{rk}(F) = \text{rk}(\mathcal{F}) = q$, $\text{rk}(\mathcal{F}_1^\perp) = q_1$ and $\text{rk}(\mathcal{F}_2^\perp) = q_2$. Since on $\mathcal{M}_{\mathcal{H}_{3m},R}$, $g^{\mathcal{F}} = \pi^* g^F$, for a local orthonormal basis $\{f_1, \dots, f_q\}$ of $(\mathcal{F}, g^{\mathcal{F}})$, we can choose it to be lifted from a local orthonormal basis of (F, g^F) . Moreover, we choose $h_1, \dots, h_{q_1}$ (resp. $e_1, \dots, e_{q_2}$) to be a local orthonormal basis of $(\mathcal{F}_1^\perp, g^{\mathcal{F}_1^\perp})$ (resp. $(\mathcal{F}_2^\perp, g^{\mathcal{F}_2^\perp})$). Then

$$\{f_1, \dots, f_q, h_1, \dots, h_{q_1}, e_1, \dots, e_{q_2}\}$$

is a local orthonormal frame for $T\mathcal{M}_{\mathcal{H}_{3m},R}$. Then, by [13, (2.40)], we have

$$(2-21) \quad \left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta}\right)(\varphi_{m,1s}) \right\|^2 \\ \geq \left(\frac{\pi^* k^F}{4\beta^2} \varphi_{m,1s}, \varphi_{m,1s}\right) + \left(\frac{1}{2\beta^2} \sum_{i,j=1}^q R^{\mathcal{E}_{3m,R}}(f_i, f_j) c_{\beta,\gamma}(\beta^{-1} f_i) c_{\beta,\gamma}(\beta^{-1} f_j) \varphi_{m,1s}, \varphi_{m,1s}\right) \\ + \left(\left[D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}}, \frac{\hat{c}(\tilde{\sigma})}{\beta}\right] \varphi_{m,1s}, \varphi_{m,1s}\right) + \left(\left[D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}}, \frac{\varepsilon \pi^* W}{\beta}\right] \varphi_{m,1s}, \varphi_{m,1s}\right) \\ + \left(\frac{\varepsilon^2 (\pi^* W)^2}{\beta^2} \varphi_{m,1s}, \varphi_{m,1s}\right) + \left(O_{m,R}\left(\frac{1}{\beta} + \frac{\gamma^2}{\beta^2}\right) \varphi_{m,1s}, \varphi_{m,1s}\right).$$

To estimate the right-hand side of (2-21), we proceed term by term.

(a) For the first two terms, by (1-1) and (2-2),

$$(2-22) \quad \left(\frac{1}{2\beta^2} \sum_{i,j=1}^q R^{\mathcal{E}_{3m,R}}(f_i, f_j) c_{\beta,\gamma}(\beta^{-1} f_i) c_{\beta,\gamma}(\beta^{-1} f_j) \varphi_{m,1s}, \varphi_{m,1s} \right) + \left(\frac{\pi^* k^F}{4\beta^2} \varphi_{m,1s}, \varphi_{m,1s} \right) \\ \geq \frac{\kappa}{4\beta^2} \|\varphi_{m,1s}\|^2.$$

Note that (2-22) is the counterpart of [13, (2.37)–(2.38)] in our situation.

(b) For the third term, by [15, Lemma 2.1], on $\mathcal{M}_{\mathcal{H}_{3m},R} \setminus s(M_{H_{3m}})$, we have

$$(2-23) \quad \left[D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}}, \frac{\hat{c}(\tilde{\sigma})}{\beta} \right] = O_m\left(\frac{1}{\beta^2 R}\right) + O_{m,R}\left(\frac{1}{\beta}\right).$$

(c) For the fourth term, since $\nabla^{\mathcal{E}_{3m,R}}$ (resp. $\pi^* W$) is a pull-back connection (resp. bundle endomorphism) via π , by [13, (2.34)] and (2-19), we have

$$(2-24) \quad \left[D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}}, \frac{\varepsilon \pi^* W}{\beta} \right] = \begin{cases} O(\varepsilon/\beta^2) + O_R(\varepsilon\gamma/\beta) & \text{on } \pi^{-1}(K), \\ 0 & \text{on } \hat{\mathcal{M}}_{\mathcal{H}_{3m},R} \setminus \pi^{-1}(K). \end{cases}$$

(d) For the fifth term, by (2-6), we have

$$(2-25) \quad \left(\frac{\varepsilon^2 (\pi^* W)^2}{\beta^2} \varphi_{m,1s}, \varphi_{m,1s} \right) \geq \frac{\varepsilon^2 \delta}{\beta^2} \|\varphi_{m,1s}\|_{\pi^{-1}(B_{2m} \setminus K)}^2,$$

where the subscript on the norm means the integral on $\pi^{-1}(B_{2m} \setminus K)$.

Now, as in [13, Lemma 2.1], we split every term on the right-hand side of (2-21) into integrals on $\pi^{-1}(K)$ and $\pi^{-1}(B_{2m} \setminus K)$ separately. At the same time, we choose ε small enough that

$$\frac{\kappa}{8\beta^2} \|s\|_{\pi^{-1}(K)}^2 + O\left(\frac{\varepsilon}{\beta^2}\right) \|s\|_{\pi^{-1}(K)}^2 \geq 0.$$

Then, by (2-22)–(2-25), we have

$$(2-26) \quad \left\| \left(D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}} + \frac{\hat{c}(\tilde{\sigma})}{\beta} + \frac{\varepsilon \pi^* W}{\beta} \right) (\varphi_{m,1s}) \right\|^2 \\ \geq \min\left\{ \frac{\kappa}{8}, \delta \varepsilon^2 \right\} \frac{\|\varphi_{m,1s}\|^2}{\beta^2} + O_R\left(\frac{\varepsilon\gamma}{\beta}\right) \|\varphi_{m,1s}\|_{\pi^{-1}(K)}^2 + O_{m,R}\left(\frac{\gamma^2}{\beta^2}\right) \|\varphi_{m,1s}\|^2 \\ + O_{m,R}\left(\frac{1}{\beta}\right) \|\varphi_{m,1s}\|^2 + O_m\left(\frac{1}{\beta^2 R}\right) \|\varphi_{m,1s}\|^2.$$

By (2-16), (2-20) and (2-26), by taking m sufficiently large and then taking R sufficiently large, one finds that there exist $c_0, \varepsilon, m, R > 0$ such that when $\beta, \gamma > 0$ are small enough, the estimate in (i) of Lemma 2.1 holds. $\square$

Step 3 Let $\partial\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ bound another oriented manifold $\mathcal{N}_{3m,R}$ such that

$$\tilde{\mathcal{N}}_{3m,R} = \hat{\mathcal{M}}_{\mathcal{H}_{3m},R} \cup \mathcal{N}_{3m,R}$$

is an oriented closed manifold. Let $g^{T\tilde{\mathcal{N}}_{3m,R}}$ be a smooth metric on $T\tilde{\mathcal{N}}_{3m,R}$ such that

$$g^{T\tilde{\mathcal{N}}_{3m,R}}|_{\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}} = g^{T\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}}.$$

The existence of $g^{T\tilde{\mathcal{N}}_{3m,R}}$ is clear.

Let $\mathcal{Q}$ be a Hermitian vector bundle over $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ such that

$$(S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{m,R})_- \oplus \mathcal{Q}$$

is a trivial vector bundle over $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$. Then

$$(S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{m,R})_+ \oplus \mathcal{Q}$$

is also a trivial vector bundle near $\partial\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ under the identification $\hat{c}(\tilde{\sigma}) + \pi^*\omega + \text{Id}_{\mathcal{Q}}$.

Since the above two vector bundles are both trivial near $\partial\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$, by extending them via the trivial bundle over $\tilde{\mathcal{N}}_{3m,R} \setminus \hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$, we get a $\mathbb{Z}_2$ -graded Hermitian vector bundle $\xi = \xi_+ \oplus \xi_-$ over $\tilde{\mathcal{N}}_{3m,R}$ and an odd self-adjoint endomorphism $\mathcal{W}' = \omega' + \omega'^* \in \Gamma(\text{End}(\xi))$ (with $\omega': \Gamma(\xi_+) \rightarrow \Gamma(\xi_-)$, ω'^* being the adjoint of ω') such that

$$\xi_{\pm} = (S_{\beta,\gamma}(\mathcal{F} \oplus \mathcal{F}_1^\perp) \hat{\otimes} \Lambda^*(\mathcal{F}_2^\perp) \hat{\otimes} \mathcal{E}_{m,R})_{\pm} \oplus \mathcal{Q}$$

over $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$, $\mathcal{W}'$ is invertible on $\mathcal{N}_{3m,R}$ and

$$(2-27) \quad \mathcal{W}' = \hat{c}(\tilde{\sigma}) + \pi^*W + \begin{pmatrix} 0 & \text{Id}_{\mathcal{Q}} \\ \text{Id}_{\mathcal{Q}} & 0 \end{pmatrix}$$

on $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$, which is invertible on $\hat{\mathcal{M}}_{\mathcal{H}_{3m},R} \setminus \mathcal{M}_{\mathcal{H}_{3m},R/2}$.

Recall that $h(\rho/R)$ vanishes near $\mathcal{M}_{\mathcal{H}_{3m},R} \cap \partial\mathcal{M}_R$. We extend it to a function on $\tilde{\mathcal{N}}_{3m,R}$ which equals zero on $\mathcal{N}_{3m,R}$ and an open neighborhood of $\partial\hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$ in $\tilde{\mathcal{N}}_{3m,R}$, and we denote the resulting function on $\tilde{\mathcal{N}}_{3m,R}$ by $\tilde{h}_R$.

Let $\pi_{\tilde{\mathcal{N}}_{3m,R}}: T\tilde{\mathcal{N}}_{3m,R} \rightarrow \tilde{\mathcal{N}}_{3m,R}$ be the projection of the tangent bundle of $\tilde{\mathcal{N}}_{3m,R}$. Let $\gamma^{\tilde{\mathcal{N}}_{3m,R}} \in \text{Hom}(\pi_{\tilde{\mathcal{N}}_{3m,R}}^* \xi_+, \pi_{\tilde{\mathcal{N}}_{3m,R}}^* \xi_-)$ be the symbol defined by

$$(2-28) \quad \gamma^{\tilde{\mathcal{N}}_{3m,R}}(p, u) = \pi_{\tilde{\mathcal{N}}_{3m,R}}^* (\sqrt{-1}\tilde{h}_R^2 c_{\beta,\gamma}(u) + \omega'(p)) \quad \text{for } p \in \tilde{\mathcal{N}}_{3m,R}, u \in T_p\tilde{\mathcal{N}}_{3m,R}.$$

By (2-27) and (2-28), $\gamma^{\tilde{\mathcal{N}}_{3m,R}}$ is singular only if $u = 0$ and $p \in \mathcal{M}_{\mathcal{H}_{3m},R/2}$. Thus $\gamma^{\tilde{\mathcal{N}}_{3m,R}}$ is an elliptic symbol.

On the other hand, it is clear that $\tilde{h}_R D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathcal{E}_{3m,R}} \tilde{h}_R$ is well defined on $\tilde{\mathcal{N}}_{3m,R}$ if we define it to be zero on $\tilde{\mathcal{N}}_{3m,R} \setminus \hat{\mathcal{M}}_{\mathcal{H}_{3m},R}$.

Let $A: L^2(\xi) \rightarrow L^2(\xi)$ be a second-order positive elliptic differential operator on $\tilde{\mathcal{N}}_{m,R}$ preserving the $\mathbb{Z}_2$ -grading of $\xi = \xi_+ \oplus \xi_-$, such that its symbol equals to $|\eta|^2$ at $\eta \in T\tilde{\mathcal{N}}_{3m,R}$.⁷ As in [15, (2.33)], let $P_{R,\beta,\gamma}^{\mathbb{E}_{3m,R}}: L^2(\xi) \rightarrow L^2(\xi)$ be the zeroth order pseudodifferential operator on $\tilde{\mathcal{N}}_{3m,R}$ defined by

$$(2-29) \quad P_{R,\beta,\gamma}^{\mathbb{E}_{3m,R}} = A^{-\frac{1}{4}} \tilde{h}_R D_{\mathcal{F} \oplus \mathcal{F}_1^\perp, \beta, \gamma}^{\mathbb{E}_{3m,R}} \tilde{h}_R A^{-\frac{1}{4}} + \frac{W'}{\beta}.$$

Let $P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}: L^2(\xi_+) \rightarrow L^2(\xi_-)$ be the obvious restriction. Then the principal symbol of $P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}$, which we denote by $\gamma(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}})$, is homotopic through elliptic symbols to $\gamma^{\tilde{\mathcal{N}}_{3m,R}}$. Thus $P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}$ is a Fredholm operator. Moreover, the index of the symbol $\gamma^{\tilde{\mathcal{N}}_{3m,R}}$ can be calculated by the Atiyah–Singer index theorem directly (cf [2] and [11, Proposition III.11.24]). Therefore,

$$(2-30) \quad \begin{aligned} \text{ind}(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}) &= \text{ind}(\gamma(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}})) = \text{ind}(\gamma^{\tilde{\mathcal{N}}_{3m,R}}) \\ &= \langle \hat{A}(T\hat{M}_{H_{3m}})(\text{ch}(\hat{E}_0) - \text{ch}(\hat{E}_1)), [\hat{M}_{H_{3m}}] \rangle \\ &= \langle \hat{A}(TM)(\text{ch}(E_0) - \text{ch}(E_1)), [M] \rangle \neq 0, \end{aligned}$$

where the inequality comes from (2-1).

For any $0 \leq t \leq 1$, set

$$(2-31) \quad P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(t) = P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}} + \frac{(t-1)\omega'}{\beta} + A^{-\frac{1}{4}} \frac{(1-t)\omega'}{\beta} A^{-\frac{1}{4}}.$$

Then $P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(t)$ is a smooth family of zeroth-order pseudodifferential operators such that the corresponding symbol $\gamma(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(t))$ is elliptic for $0 < t \leq 1$. Thus $P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(t)$ is a continuous family of Fredholm operators for $0 < t \leq 1$ with $P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(1) = P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}$.

Then, by using Lemma 2.1, the exactly same arguments in [13, Proposition 2.2] show that for suitable $\varepsilon, m, R, \beta, \gamma > 0$,

$$\dim(\ker(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(0))) = \dim(\ker(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(0)^*)) = 0.$$

As a result, when $t = 0$, $P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(0)$ is also Fredholm and has a vanishing index. By the property of Fredholm index, we have

$$\text{ind}(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}) = \text{ind}(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(1)) = \text{ind}(P_{R,\beta,\gamma,+}^{\mathbb{E}_{3m,R}}(0)) = 0,$$

which contradicts (2-30) and the proof of Theorem 1.2 is completed.

Acknowledgements G Su was partially supported by NSFC grant 12425106, 11931007 and 12271266, the Nankai Zhide Foundation and the Fundamental Research Funds for the Central Universities grant 63243068. X Wang was partially supported by NSFC grant 12471049 and 12101361, the project of Young Scholars of Shandong University.

⁷To be more precise, here A also depends on the defining metric. We omit the corresponding subscript/superscript only for convenience.

References

- [1] **M F Atiyah**, *K-theory*, Benjamin, New York (1967) [MR](#) [Zbl](#)
- [2] **M F Atiyah, I M Singer**, *The index of elliptic operators, I*, Ann. of Math. 87 (1968) 484–530 [MR](#) [Zbl](#)
- [3] **C Bär, B Hanke**, *Boundary conditions for scalar curvature*, from “Perspectives in scalar curvature, II”, World Sci., Hackensack, NJ (2023) 325–377 [MR](#) [Zbl](#)
- [4] **M-T Benameur, J L Heitsch**, *Enlargeability, foliations, and positive scalar curvature*, Invent. Math. 215 (2019) 367–382 [MR](#) [Zbl](#)
- [5] **S Cecchini, R Zeidler**, *Scalar and mean curvature comparison via the Dirac operator*, Geom. Topol. 28 (2024) 1167–1212 [MR](#) [Zbl](#)
- [6] **A Connes**, *Cyclic cohomology and the transverse fundamental class of a foliation*, from “Geometric methods in operator algebras”, Pitman Res. Notes Math. Ser. 123, Longman, Harlow (1986) 52–144 [MR](#) [Zbl](#)
- [7] **M Gromov**, *Positive curvature, macroscopic dimension, spectral gaps and higher signatures*, from “Functional analysis on the eve of the 21st century, II”, Progr. Math. 132, Birkhäuser, Boston, MA (1996) 1–213 [MR](#) [Zbl](#)
- [8] **M Gromov**, *101 questions, problems and conjectures around scalar curvature*, preprint (2017) Available at <https://www.ihes.fr/~gromov/positivescalarcurlvature/596/>
- [9] **M Gromov**, *Four lectures on scalar curvature*, from “Perspectives in scalar curvature, I”, World Sci., Hackensack, NJ (2023) 1–514 [MR](#) [Zbl](#)
- [10] **M Gromov, H B Lawson, Jr**, *Positive scalar curvature and the Dirac operator on complete Riemannian manifolds*, Inst. Hautes Études Sci. Publ. Math. 58 (1983) 83–196 [MR](#) [Zbl](#)
- [11] **H B Lawson, Jr, M-L Michelsohn**, *Spin geometry*, Princeton Math. Ser. 38, Princeton Univ. Press (1989) [MR](#) [Zbl](#)
- [12] **M Listing**, *Homology of finite K-area*, Math. Z. 275 (2013) 91–107 [MR](#) [Zbl](#)
- [13] **G Su, X Wang, W Zhang**, *Nonnegative scalar curvature and area decreasing maps on complete foliated manifolds*, J. Reine Angew. Math. 790 (2022) 85–113 [MR](#) [Zbl](#)
- [14] **X Wang**, *On a relation between the K-cowaist and the $\hat{A}$ -cowaist*, Proc. Amer. Math. Soc. 151 (2023) 4983–4990 [MR](#) [Zbl](#)
- [15] **W Zhang**, *Positive scalar curvature on foliations*, Ann. of Math. 185 (2017) 1035–1068 [MR](#) [Zbl](#)
- [16] **W Zhang**, *Positive scalar curvature on foliations: the enlargeability*, from “Geometric analysis”, Progr. Math. 333, Birkhäuser (2020) 537–544 [MR](#) [Zbl](#)

Chern Institute of Mathematics & LPMC, Nankai University
Tianjin, China

School of Mathematics, Shandong University
Jinan, China

guangxiangsu@nankai.edu.cn, xiangsheng@sdu.edu.cn

Received: 22 June 2022

Line bundle twists for unitary bordism are ghosts

THORSTEN HERTL

We prove that the canonical twist $\zeta: K(\mathbb{Z}, 3) \rightarrow BGL_1(M\mathrm{Spin}^c)$ does not extend to a twist for unitary bordism by showing that every continuous map $f: K(\mathbb{Z}, 3) \rightarrow BGL_1(MU)$ loops to a nullhomotopic map.

[55N22](#), [55P42](#), [55R37](#); [14L05](#), [19L50](#)

1 Introduction

Twisted (co)homologies capture geometric information and invariants that are only available for untwisted (co)homologies under additional orientation assumptions. Prominent examples of this phenomenon are Thom classes. A vector bundle $E \rightarrow X$ of rank d has a Thom class $t_E \in H^d(E, E \setminus 0; \mathbb{Z})$ if and only if E is orientable. However, it always has a Thom class in $H^d(E, E \setminus 0; \mathbb{Z}^{w_1(E)})$ if the coefficients are twisted by the first Stiefel–Whitney class $w_1(E) \in H^1(X; \mathbb{Z}_2) = \mathrm{Hom}(\pi_1(X), \mathrm{Aut}(\mathbb{Z}))$.

Orientability of E , however, does not suffice to guarantee a Thom class in K -theory. Atiyah, Bott and Shapiro showed in [4] that an oriented vector bundle E has a Thom class in K -theory if the structure group reduces to $\mathrm{Spin}^c(d)$. They also constructed a Thom class in this case, which is referred to as *Atiyah–Bott–Shapiro orientation*. To produce K -theoretic Thom classes for all oriented vector bundles, Donovan and Karoubi [10], later generalised by Atiyah and Segal [5], constructed twisted K -theory groups using operator-theoretic methods. Since $K(\mathbb{Z}, 3) = BPU(H)$ is the classifying space of the projective unitary group of a separable Hilbert space H , each $\alpha \in H^3(X, \mathbb{Z})$ gives rise to a bundle of Fredholm operators over X . The group $K_\alpha(X)$ is the set of homotopy classes of sections of this bundle. We have now, for each oriented bundle $E \rightarrow X$, a Thom isomorphism

$$K^*(X) \cong K_{W_3(E)}^{*+\mathrm{rk}(E)}(E, E \setminus 0),$$

where $W_3(E) \in H^3(X; \mathbb{Z})$ is the third integral Stiefel–Whitney class; see for example Carey and Wang [8] for details.

A homotopy-theoretic approach to twisted (co)homology in terms of twisted spectra is provided by the work of Ando, Blumberg and Gepner [1] and Ando, Blumberg, Gepner, Hopkins and Rezk [2; 3]. Their construction goes roughly as follows: An A_∞ -ring spectrum R has a space of units $\mathrm{GL}_1(R)$, which deloops to a classifying space $BGL_1(R)$. A *twist* is a map $\xi: X \rightarrow BGL_1(R)$ from which we can

construct a Thom spectrum $R_{X,\xi}$, which is an R -module spectrum. The associated twisted (co)homology groups are then given by

$$R_k(X, \xi) := \pi_0 R\text{-mod}(\Sigma^k R, R_{X,\xi}) \quad \text{and} \quad R^k(X, \xi) := \pi_0 R\text{-mod}(R_{X,\xi}, \Sigma^k R).$$

If the twist is nullhomotopic, then this construction gives the usual (untwisted) (co)homology theory associated to R . In practice, twists of interest factor through specific reference spaces; in the case of K -theory, this reference space is $K(\mathbb{Z}, 3)$.

Our motivation for the work presented here arose from the results of Hebestreit and Joachim [11]; see also Hebestreit and Savage [12]. The authors construct point-set models for twisted Spin^c bordism and twisted K -theory over $K(\mathbb{Z}, 3)$ as well as a model for a twisted Atiyah–Bott–Shapiro orientation. It is shown in [11, Appendix C] that the point-set model also arises from the twist $\zeta: K(\mathbb{Z}, 3) \rightarrow BGL_1(M\text{Spin}^c)$ that comes from the fibre inclusion $U(1) \rightarrow \text{Spin}^c$; see Section 4 for a detailed description of this twist.

One application of these models concerns the bordism-determines-homology question. The classical results of Conner and Floyd [9] and Hopkins and Hovey [13] state that K -theory is completely described by the Todd orientation $\text{Td}: MU \rightarrow K$ and by the Atiyah–Bott–Shapiro orientation $\alpha_{\text{ABS}}: M\text{Spin}^c \rightarrow K$ in the sense that

$$MU_*(X) \otimes_{MU_*(\text{pt})} K_*(\text{pt}) \cong K_*(X) \cong M\text{Spin}^c_*(X) \otimes_{M\text{Spin}^c_*(\text{pt})} K_*(\text{pt})$$

for all spectra X . In the forthcoming preprint of Baum, Joachim, Khorami and Schick [7] the latter isomorphism is proved by geometric means in the twisted setup using the $K(\mathbb{Z}, 3)$ -twisted ABS orientation.

Since the Todd orientation factors through the ABS orientation under the canonical map $MU \rightarrow M\text{Spin}^c$, it is natural to ask whether one can also extend the first isomorphism to the $K(\mathbb{Z}, 3)$ -twisted setup so that it reduces to the isomorphism of [7]. The first step would be to construct a twisted unitary bordism spectrum $MU_{K(\mathbb{Z}, 3), ?}$ over $K(\mathbb{Z}, 3)$ and a comparison map $MU_{K(\mathbb{Z}, 3), ?} \rightarrow M\text{Spin}^c_{K(\mathbb{Z}, 3), \zeta}$ that extends the twisted ABS orientation “to the left”. Our first result shows that this is impossible:

Theorem A *The twist $\zeta: K(\mathbb{Z}, 3) \rightarrow BGL_1(M\text{Spin}^c)$ does not factor through unitary bordism. More precisely, there are no continuous maps S and T that make the following diagram homotopy commutative:*

$$\begin{array}{ccc} K(\mathbb{Z}, 3) & \xrightarrow{\zeta} & BGL_1(M\text{Spin}^c) \\ & \searrow T & \uparrow S \\ & & BGL_1(MU) \end{array}$$

Theorem A is a consequence of the following purely homotopy-theoretical result:

Theorem B *Every continuous map $f: K(\mathbb{Z}, 3) \rightarrow BGL_1(MU)$ loops to a nullhomotopic map*

$$\Omega f: K(\mathbb{Z}, 2) \rightarrow GL_1(MU).$$

Its (slightly more general) stable counterpart reads as follows:

Lemma C Each map of spectra $\Sigma_+^\infty K(\mathbb{Z}, 2) \rightarrow MU$ that is a ring map up to homotopy induces a homomorphism between the Pontrjagin rings $MU_*(K(\mathbb{Z}, 2))$ and $MU_*(MU)$ that vanishes on $\widetilde{MU}_*(K(\mathbb{Z}, 2))$.

Partial results strongly indicating [Theorem A](#) were obtained by Joel Meier in his unpublished Master's thesis [\[16\]](#). He showed that there is no H -map $K(\mathbb{Z}, 2) \rightarrow SO/U$ whose composition with the fibre comparison map $SO/U \rightarrow K(\mathbb{Z}, 2) = \text{hofib}(B\text{Spin}^c \rightarrow BSO)$ is homotopic to the identity. Such a map, when it deloops, would produce a nontrivial map T in [Theorem A](#). In this case, S would arise from a map of spectra, a condition we do not assume.

[Lemma C](#) has a purely algebraic application. A formal group law over a ring R is a formal power series $F \in R[[x, y]]$ that satisfies the same conditions as the Taylor series expansion of the multiplication of a one-dimensional Lie group. Let L be the Lazard ring, the ring over which the universal formal group law F_{univ} is defined. [Lemma C](#) and Quillen's result [\[17\]](#) that L is isomorphic to $MU_*(\text{pt})$ imply the next result:

Theorem D There is no nontrivial homomorphism between the universal formal group law F_{univ} and the multiplicative formal group $F_{\text{mult}}(x, y) := x + y + xy$ on L . More precisely, if $g \in L[[x]]$ is a formal power series with $g(0) = 0$ and

$$g(F_{\text{univ}}(x, y)) = F_{\text{mult}}(g(x), g(y)),$$

then $g = 0$.

Outline of the proof [Theorem A](#) follows from [Theorem B](#) once we know that ζ is sufficiently nontrivial. We will translate the statement of [Theorem B](#) into a stable homotopy-theoretical one and use the Adams–Novikov spectral sequence to deduce it from [Lemma C](#). The main task is therefore the proof of [Lemma C](#), which is purely algebraic. We will show that $MU_*(\text{pt})$ -algebra homomorphisms between the Pontrjagin rings $MU_*(K(\mathbb{Z}, 2))$ and $MU_*(MU)$ are uniquely determined by the image of a single element in low degrees. Algebraic arguments then imply that any algebra homomorphism $MU_*(\mathbb{C}P^\infty) \rightarrow MU_*(MU)$ factors through $MU_*(\text{pt})$. Using the natural homology co-operations, we derive that no such algebra homomorphism can come from a map of spectra.

Organisation We first give a short summary of the theory of orthogonal spectra and set up conventions in [Section 2](#). [Section 3](#) is devoted to the proof of [Theorem B](#) and [Lemma C](#). There, we also outline how the questions we address are related to the question of whether the K -theory spectrum is a retract as a ring spectrum of the periodic complex bordism spectrum MUP . The proof of [Theorem A](#) will be carried out in [Section 4](#). [Section 5](#), in which we prove [Theorem D](#), is logically independent of [Section 4](#).

Acknowledgements The motivation to study these questions arose from the Master's thesis of Joel Meier [\[16\]](#). I would like to thank him for several discussions and reading an earlier draft of this article. I also thank Thorben Kastenzholz for an illuminating example, a referee for suggesting to study the connection to MUP and another referee for beneficial suggestions. Finally, I thank my two PhD advisors, Thomas Schick and Wolfgang Steimle, for their valuable advice.

2 Foundations and conventions

Throughout the entire article we work in the category of compactly generated Hausdorff spaces. We assume that the reader is familiar with the notion of orthogonal spectra at the level of [6]. The fact important to us is that orthogonal ring spectra form a symmetric monoidal model category [6, Theorem 6.4.8] with the sphere spectrum $\mathbb{S}$ as unit. The monoid objects are precisely (unital) ring spectra. The category of ring spectra also carries a model structure such that forgetting the ring structure is a Quillen right adjoint [6, Theorem 6.6.12]. Suspension and evaluation at the initial space form a strong symmetric monoidal adjunction

$$(1) \quad \Sigma_+^\infty : \mathbf{Top} \xrightleftharpoons[\Omega^\infty]{\Sigma^\infty} \mathbf{Sp}^{\text{orth}} : \Omega^\infty.$$

See [6, Lemmas 6.3.21 and 6.3.18]. We refer to the natural bijection between $[X, \Omega^\infty Y]$ and $[\Sigma_+^\infty X, Y]$ with Ad . Consequently, the suspension spectrum of a topological monoid is a ring spectrum and the subspace $R_0^\times$ of all elements of the initial space R_0 of a ring spectrum R that induce invertible elements in $\pi_0(R)$ is a topological monoid. The suspension of a monoid homomorphism is a ring map and, conversely, each ring map $f : R \rightarrow S$ restricts to a monoid homomorphism $f_0 : R_0^\times \rightarrow S_0^\times$.

The left adjoint Σ_+^∞ maps the canonical cylinder object $(-) \times [0, 1]$ to $(-) \wedge [0, 1]_+$ so that H -spaces give rise to homotopy ring spectra and H -maps give rise to homotopy ring maps.

If R is a cofibrant and fibrant object in the model category of orthogonal ring spectra, then its *space of units* $\text{GL}_1(R)$ is the set of all path components of $R_0 = \Omega^\infty R$ that represent invertible elements in $\pi_0(R) = \pi_0(R_0)$. It is a group-like topological monoid.

For a topological monoid G , we denote by $|G_\bullet|$ the geometric realisation of the singular simplicial set $S_\bullet(G)$. It is again a topological monoid and the counit $\varepsilon : |G_\bullet| \rightarrow G$ is a continuous monoid-homomorphism and a weak homotopy equivalence.

For a group-like topological monoid, we define its classifying space $BG := |B_\bullet(e, |G_\bullet|, e)|$ to be the (thin) geometric realisation of the bar construction of $|G_\bullet|$ in sense of [15, Section 11]. It comes with a quasifibration $EG \rightarrow BG$, where EG is a pointed contractible free $|G_\bullet|$ -right space. In this setup, Sugawara [20] constructed a *group completion map* $\iota : G \rightarrow \Omega BG$, which is an H -map [20, Lemma 11] and hence a weak homotopy equivalence (because G is group-like).

3 Proof of Theorem B

We first prove Theorem B under the assumption of Lemma C. To this end, we need good models for the involved players so that the spaces of units are strict monoids.

A model for $K(\mathbb{Z}, 2) \simeq BU(1)$ is given by $\mathbb{C}P^\infty$. The H -structure induced by the group structures is modelled by polynomial multiplication

$$\mu : ([\zeta_0 : \zeta_1 : \cdots], [\omega_0 : \omega_1 : \cdots]) \mapsto ([\zeta_0\omega_0 : \zeta_1\omega_0 + \zeta_0\omega_1 : \zeta_2\omega_0 + \zeta_1\omega_1 + \zeta_0\omega_2 : \cdots]).$$

It turns $\mathbb{C}P^\infty$ into a group-like monoid, whose homotopy inversion is modelled by complex conjugation.

Pick an orthogonal ring spectrum that represents unitary bordism and denote its cofibrant-fibrant replacement in the model category of ring spectra by MU .

Lemma 3.1 *Adjunction (1) induces an injection*

$$[\mathbb{C}P^\infty, GL_1 MU] \hookrightarrow [\Sigma_+^\infty \mathbb{C}P^\infty; MU]$$

and the image of the constant maps are precisely those homotopy classes that induce the zero map on $\widetilde{MU}_*(\mathbb{C}P^\infty)$.

Proof Since $GL_1(MU) \subseteq MU_0 = \Omega^\infty MU$ is a collection of path components, this inclusion induces an injective map on homotopy classes. As adjunction (1) is a Quillen adjunction, we get an injective map

$$[\mathbb{C}P^\infty, GL_1(MU)] \hookrightarrow [\mathbb{C}P^\infty, MU_0] \xrightarrow{\text{Ad}} [\Sigma_+^\infty \mathbb{C}P^\infty; MU].$$

From the Adams–Novikov spectral sequence it follows that the Hurewicz homomorphism

$$[\Sigma_+^\infty \mathbb{C}P^n; MU] \rightarrow \text{Hom}_{MU_* MU}(MU_* \mathbb{C}P^n; MU_* MU),$$

whose target is the set of all $MU_* MU$ -comodule maps, is bijective. Indeed, as described in Switzer's book [21, Chapter 19], the edge homomorphism of the Adams–Novikov spectral sequence agrees with the Hurewicz homomorphism. Since $MU_* MU$ is a free $MU_* MU$ -comodule and $MU_*(\mathbb{C}P^n)$ is a free $MU_*(\text{pt})$ -module [21, Proposition 16.30], the Adams–Novikov spectral sequence is concentrated in the column

$$E_{0,t}^2 = \text{Hom}_{MU_* MU}^t(MU_* \mathbb{C}P^n; MU_* MU),$$

by [21, Proposition 19.7], which forces the spectral sequence to collapse for algebraic reasons. As the spectral sequence converges to $[\Sigma_+^\infty \mathbb{C}P^n; MU]_*$, the edge homomorphism must be an isomorphism.

Since $[\Sigma_+^\infty \mathbb{C}P^\infty; MU] = MU^0(\mathbb{C}P^\infty) = \lim_{n \rightarrow \infty} MU^0(\mathbb{C}P^n)$, see [21, Proof of Proposition 16.29], we deduce that the Hurewicz-homomorphism

$$[\Sigma_+^\infty \mathbb{C}P^\infty; MU] \rightarrow \text{Hom}_{MU_* MU}(MU_* \mathbb{C}P^\infty; MU_* MU),$$

is bijective, too.

The constant maps

$$\mathbb{C}P^\infty \xrightarrow{\text{const}} \text{pt} \xrightarrow{c} MU_0 = \Omega^\infty MU$$

correspond under the adjunction to the maps of spectra

$$\Sigma_+^\infty \mathbb{C}P^\infty \xrightarrow{\Sigma_+^\infty \text{const}} \Sigma_+^\infty \text{pt} = \mathbb{S} \xrightarrow{\text{Ad}(c)} MU,$$

so they induce the zero map on $\widetilde{MU}_*(\mathbb{C}P^\infty) = MU_*(\mathbb{C}P^\infty, \text{pt}) = \ker \text{const}_*$. The bijection of the Hurewicz map conversely shows that every $MU_*(MU)$ comodule map that vanishes on $\widetilde{MU}_*(\mathbb{C}P^\infty) \subseteq MU_*(\mathbb{C}P^\infty)$ must come from a map of spectra that is homotopic to the adjoint of a constant map. $\square$

Proof of Theorem B Let $f: B\mathbb{C}P^\infty \rightarrow BGL_1(MU)$ be a twist and let Ωf be its induced map on loop spaces. The monoids $\mathbb{C}P^\infty$ and $GL_1(MU)$ are group-like, so the group completion maps are

weak homotopy equivalences. Since postcomposition with weak equivalences induces bijections between homotopy classes (see [21, Theorem 6.31]) there are unique homotopy classes such that each representative makes the following diagram homotopy commutative:

$$\begin{array}{ccc}
 \mathbb{C}P^\infty & \dashrightarrow & \mathrm{GL}_1(MU) \\
 \varepsilon \uparrow \simeq & & \simeq \uparrow \varepsilon \\
 |\mathbb{C}P^\bullet| & \dashrightarrow & |\mathrm{GL}_1(MU)_\bullet| \\
 \iota \downarrow \simeq & & \simeq \downarrow \iota \\
 \Omega B\mathbb{C}P^\infty & \xrightarrow{\Omega f} & \Omega B\mathrm{GL}_1(MU)
 \end{array}$$

The maps of the group-competition zigzags and Ωf are H -maps, so the homotopy class of the lift $\mathbb{C}P^\infty \dashrightarrow \mathrm{GL}_1(MU)$ must consist of H -maps. We denote any representative of the lifted homotopy class again by Ωf .

Under adjunction (1) the H -map Ωf corresponds to a homotopy ring map

$$\mathrm{Ad}(\Omega f): \Sigma_+^\infty \mathbb{C}P^\infty \rightarrow MU.$$

Lemma 3.1 implies that Ωf is nullhomotopic if $\mathrm{Ad}(\Omega f)$ is. But $\mathrm{Ad}(\Omega f)$ is nullhomotopic by the second half of Lemma 3.1 and Lemma C. $\square$

It remains to prove Lemma C, which is reformulated in Lemma 3.4 below. Before we carry out the proof, we recall and develop the required structural results.

We recall from Switzer [21, Proposition 16.29] that $MU^*(\mathbb{C}P^\infty) \cong MU^*[\![c]\!]$, where $c \in \widetilde{MU}^2(\mathbb{C}P^\infty)$ is the universal first Chern class. As a $MU_*(\mathrm{pt})$ -left module, $MU_*(\mathbb{C}P^\infty)$ is isomorphic to the free module $MU_*(\mathrm{pt})\{1, \beta_1, \beta_2, \dots\}$, where β_j is dual to c^j ; see [21, Proposition 16.30]. Since $MU_*(\mathrm{pt})$ is a polynomial ring over $\mathbb{Z}$ (see [21, Theorem 20.25]) it is torsion free.

We recall further that $(MU_*MU, MU_*(\mathrm{pt}))$ is a Hopf-algebroid. This means, in particular, that the multiplication of MU turns MU_*MU into a commutative ring, that the diagonal map $\Delta: MU \rightarrow MU \wedge MU$ induces a coaction on MU_*MU , and that these two structures are compatible. However, $MU_*(\mathrm{pt})$ acts canonically from the left and from the right on $MU_*(MU)$ and these actions are not the same. There are therefore a left unit η_L and a right unit η_R on MU_*MU . The left unit sends $\lambda \in MU_*(\mathrm{pt})$ to $\lambda \cdot 1$, where 1 is the unit of the Pontrjagin ring. For more details, see [21, page 414 ff] or [14, Proposition 4.5.3]. We will consider MU_*MU as a left-module over $MU_*(\mathrm{pt})$. As an $MU_*(\mathrm{pt})$ -algebra, on which the ring $MU_*(\mathrm{pt})$ acts from the left, $MU_*(MU)$ is a polynomial algebra

$$MU_*(MU) \cong MU_*(\mathrm{pt})[b_1, b_2, \dots],$$

where $b_j \in MU_{2j}(MU)$; see [21, Theorem 17.16] or [14, Proposition 4.4.4]. The next lemma on the Pontrjagin ring structure of $MU_*(\mathbb{C}P^\infty)$ is a direct consequence of [18, Theorem 3.4].

Lemma 3.2 *The additive generators $\{1, \beta_1, \beta_2, \dots\}$ of the Pontrjagin ring $MU_*(\mathbb{C}P^\infty)$ satisfy the multiplicative relations*

$$\beta_i \bullet \beta_j := MU_*(\mu)(\beta_i \otimes \beta_j) = \sum_{k \geq 0} \left(\sum_{\substack{a_1 + \dots + a_k = i \\ b_1 + \dots + b_k = j}} \prod_{r=1}^k \alpha_{a_r b_r} \right) \beta_k,$$

where the $\alpha_{ij} \in MU_{2(i+j-1)}(\text{pt})$ are the coefficients in the unitary bordism formal group law.

To make calculations manageable, we eventually will carry them out in complex K -theory and integral homology. Let us first recall the connection between unitary bordism and complex K -theory following [21, page 423 ff and page 433 ff]. If $t \in K_2(\text{pt}) = K^{-2}(\text{pt})$ denotes the Bott-element, then Bott periodicity yields $K_*(\text{pt}) \cong \mathbb{Z}[t, t^{-1}]$. The Pontrjagin ring $K_*(K)$ embeds into $K_*(K) \otimes \mathbb{Q} \cong \mathbb{Q}[u, v, u^{-1}, v^{-1}]$. It is a Hopf-algebroid with left unit $\eta_L(t) = u$ and right unit $\eta_R(t) = v$. We do not need its coaction here. As in the case of unitary bordism, $K_*(\mathbb{C}P^\infty)$ is a free $K_*(\text{pt})$ -left module and $K_*(MU)$ is isomorphic to a polynomial ring over $K_*(\text{pt})$. More precisely, we have an isomorphism of $K_*(\text{pt})$ -left modules

$$K_*(\mathbb{C}P^\infty) \cong K_*(\text{pt})\{1, t^1 Y_1, t^2 Y_2, \dots\},$$

with $Y_j \in K_0(\mathbb{C}P^\infty)$ and an isomorphism of $K_*(\text{pt})$ -algebras

$$K_*(MU) \cong K_*(\text{pt})[b_1^K, b_2^K, \dots],$$

where the underlying ring $K_*(\text{pt})$ acts from the left. In [21] the elements b_j^K are denoted by Y'_j .

The Todd orientation is a map of spectra $\text{Td}: MU \rightarrow K$ that classifies the Thom class of the universal complex vector bundle; see [21, page 434] or [9, page 29]. It induces a natural transformation of homology theories $MU_* \rightarrow K_*$ that maps β_j to $t^j Y_j$ and b_j to b_j^K , respectively [21, page 434]. It follows from the geometric description that Td maps the generator $[\mathbb{C}P^1] \in \Omega_2^U(\text{pt})$ to the Bott element $t \in K_2(\text{pt})$. Alternatively, it follows from [9, Corollary 6.5] and the fact that the Todd-genus of $\mathbb{C}P^1$ is 1.

Like Lemma 3.2, the next lemma is also a consequence of [18, Theorem 3.4], but the given formula is now explicit.

Lemma 3.3 *The additive generators $\{1, t Y_1, t^2 Y_2, \dots\}$ of the Pontrjagin ring $K_*(\mathbb{C}P^\infty)$ satisfy the multiplicative relations*

$$t^i Y_i \bullet t^j Y_j = t^{i+j} \cdot \sum_{k=\max\{i,j\}}^{i+j} \binom{k}{2k-(i+j)} \binom{2k-(i+j)}{k-j} Y_k.$$

Proof The (standard) complex orientation of complex K -theory is given by

$$y^K := t^{-1} \cdot (\tau - 1) \in \tilde{K}^2(\mathbb{C}P^\infty),$$

where $1 = \varepsilon^1$ denotes the trivial complex line bundle. Since the monoid map $\mu: \mathbb{C}P^\infty \times \mathbb{C}P^\infty \rightarrow \mathbb{C}P^\infty$ classifies the tensor product of vector bundles, we have under the isomorphism $K^*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \cong K^*(\text{pt})[[x, y]]$ the equality

$$K^*(\mu)(y^K) = x + y + t \cdot xy.$$

See [14, page 148]. Using [18, Theorem 3.4], we deduce the following identity of formal power series:

$$\sum_{i,j \geq 0} (t^i Y_i \bullet t^j Y_j) x^i y^j = \sum_{k \geq 0} t^k Y_k (x + y + txy)^k = \sum_{k \geq 0} \sum_{l=0}^k \sum_{m=0}^l \binom{k}{l} \binom{l}{m} t^{2k-l} Y_k \cdot x^{k-l+m} y^{k-m}.$$

The system of equations $k - l + m = i$ and $k - m = j$ is solvable under the constraints $0 \leq m \leq l \leq k$ if and only if $0 \leq 2k - (i + j) \leq k$, or equivalently if $\lceil \frac{1}{2}(i + j) \rceil \leq k \leq i + j$. In this case the solutions are $m = k - j$ and $l = 2k - (i + j)$. Comparing coefficients now yields

$$t^i Y_i \bullet t^j Y_j = t^{i+j} \cdot \sum_{k=\lceil (i+j)/2 \rceil}^{i+j} \binom{k}{2k-(i+j)} \binom{2k-(i+j)}{k-j} Y_k.$$

The latter binomial coefficient is zero if $k - j > 2k - (i + j)$ or $k - i > 2k - (i + j)$. Thus the formula can be shortened to the claimed formula. $\square$

Lastly, the integral homology groups of $\mathbb{C}P^\infty$ and MU are given by $H_*(\mathbb{C}P^\infty; \mathbb{Z}) \cong \mathbb{Z}\{1, \beta_1, \dots\}$ and $H_*(MU; \mathbb{Z}) \cong \mathbb{Z}[b_1, b_2, \dots]$, where the latter is an isomorphism of rings. The multiplicative structure on the Pontrjagin ring $H_*(\mathbb{C}P^\infty; \mathbb{Z})$ is given by

$$\beta_i \bullet \beta_j = \binom{i+j}{i} \beta_{i+j}.$$

The (generalised) Hurewicz homomorphism that sends a singular stable complex manifold $[f: M^n \rightarrow X]$ to the image of the fundamental class $f_*[M] \in H_n(X, \mathbb{Z})$ is a natural transformation of homology theories $\text{Hur}: MU_*(-) \rightarrow H_*(-; \mathbb{Z})$. It has the property that it sends β_i^{MU} to β_i^H and b_i^{MU} to b_i^H .

Lemma 3.4 Every map of homotopy ring spectra $\varphi: \Sigma_+^\infty \mathbb{C}P^\infty \rightarrow MU$ induces a homomorphism that vanishes on $\widetilde{MU}_*(\mathbb{C}P^\infty)$.

Proof Since $\widetilde{MU}_*(\mathbb{C}P^\infty) = MU_*(\text{pt})\{\beta_1, \beta_2, \dots\} \subseteq MU_*(\mathbb{C}P^\infty)$ [21, Proposition 16.30] it suffices to show that $MU_*(\varphi)$ sends all β_n to zero. Since $MU_*(\varphi)$ is degree preserving, there are elements $\lambda_j \in MU_j(\text{pt})$ such that

$$MU_*(\varphi)(\beta_1) = \lambda_2 1 + \lambda_0 b_1 =: Q \in MU_2(MU).$$

From Lemma 3.2 we derive $\beta_1 \bullet \beta_1 = 2\beta_2 + \alpha_{11}\beta_1$ and inductively

$$\beta_1^{\bullet n} = n! \beta_n + n \zeta_{n-1} \beta_{n-1} + \dots + n \zeta_1 \beta_1$$

for some $n \zeta_j \in MU_{2(n-j)}(\text{pt})$. Since $MU_*(MU)$ is torsion-free it suffices to show $Q := MU_*(\varphi)(\beta_1) = 0$.

To show that $\lambda_0 = 0$, we consider the image of Q in integral homology. The following diagram commutes:

$$\begin{array}{ccc} MU_*(\mathbb{C}P^\infty) & \xrightarrow{MU_*(\varphi)} & MU_* MU \\ \downarrow \text{Hur} & & \downarrow \text{Hur} \\ H_*(\mathbb{C}P^\infty; \mathbb{Z}) & \xrightarrow{H_*(\varphi)} & H_*(MU; \mathbb{Z}) \end{array}$$

Since Hur is a natural transformation, we deduce

$$n! \cdot H_{2n}(\varphi)(\beta_n) = \text{Hur}(f(\beta^{\bullet n})) = \text{Hur}(Q^n) = \text{Hur}(Q)^n = \lambda_0^n,$$

which implies that $n \mid \lambda_0$ for all $n \in \mathbb{N}$, and hence $\lambda_0 = 0$. Thus $Q = \lambda_2 \cdot 1 \in MU_2(\text{pt})$.

To show that $\lambda_2 = 0$, we consider the image of Q in K -theory. The following diagram commutes:

$$\begin{array}{ccc} MU_*(\mathbb{C}P^\infty) & \xrightarrow{MU_*(\varphi)} & MU_*MU \\ \downarrow \text{Td} & & \downarrow \text{Td} \\ K_*(\mathbb{C}P^\infty) & \xrightarrow{K_*(\varphi)} & K_*(MU) \end{array}$$

Set $g := K_*(\varphi)$. Since Td is an isomorphism in degree two, the previous calculation implies

$$g(tY_1) = g(\text{Td}_2(\beta_1)) = \text{Td}_2(f(\beta_1)) = \text{Td}_2(\lambda_2 1_{MU_*MU}) = \text{Td}_2(\lambda_2) 1_{K_*(MU)} = \lambda_2 t \cdot 1_{K_*MU},$$

where we abuse notation and denote by λ_2 an element in $MU_2(\text{pt})$ and the multiple $\lambda_2 \in \mathbb{Z}$ of the generator $[\mathbb{C}P^1] \in \Omega_2^U(\text{pt})$ such that $\lambda_2 \cdot [\mathbb{C}P^1] = \lambda_2$. The Pontrjagin ring structure of $K_*(\mathbb{C}P^\infty)$ gives

$$2g(t^2Y_2) = g(tY_1 \bullet tY_1) - g(t^2Y_1) = g(tY_1)^2 - t g(tY_1) = t^2(\lambda_2^2 - \lambda_2).$$

The formulas for the coactions

$$\psi_{\mathbb{C}P^\infty}: K_*(\mathbb{C}P^\infty) \rightarrow K_*(K) \otimes_{K_*(\text{pt})} K_*(\mathbb{C}P^\infty) \quad \text{and} \quad \psi_{MU}: K_*(MU) \rightarrow K_*(K) \otimes_{K_*(\text{pt})} K_*(MU)$$

can be found in [21, Proposition 17.38]. Since the coactions are natural with respect to morphisms induced by maps of spectra, we must have

$$(\text{id} \otimes g)(\psi_{\mathbb{C}P^\infty}(t^2Y_2)) = \psi_{MU}(g(t^2Y_2)).$$

The following calculations,¹ in which $P = 1 + p_1 + p_2 + \dots$ denotes the formal series with

$$p_n = \frac{(v-u)(v-2u) \cdots (v-nu)}{(n+1)!} \in K_{2n}(K) \subseteq \mathbb{Q}[u, v, u^{-1}, v^{-1}]_{(2n)}$$

and $\cdot_r$ denotes the right action,

$$\begin{aligned} 2(\text{id} \otimes g)(\psi_{\mathbb{C}P^\infty}(t^2Y_2)) &= \sum_{i+j=2} (P^j)_{2i} \otimes 2 \cdot g(t^jY_j) \\ &= 1 \otimes t^2(\lambda_2^2 - \lambda_2) \cdot 1_{K_*MU} + p_1 \otimes 2\lambda_2 t \cdot 1_{K_*MU} + 0 \otimes 2 \\ &= 1_{K_*K} \cdot_r t^2(\lambda_2^2 - \lambda_2) \otimes 1_{K_*MU} + \frac{1}{2}(v-u) \cdot_r 2\lambda_2 t \otimes 1_{K_*MU} \\ &= (\lambda_2^2 - \lambda_2)v^2 \otimes 1_{K_*MU} + \lambda_2(v-u)v \otimes 1_{K_*MU} = (\lambda_2^2 v^2 - \lambda_2 uv) \otimes 1_{K_*MU} \end{aligned}$$

and

$$\begin{aligned} 2\psi_{MU}(g(t^2Y_2)) &= \psi_{MU}(t^2(\lambda_2^2 - \lambda_2) \cdot 1_{K_*MU}) = t^2(\lambda_2^2 - \lambda_2) \cdot \psi_{MU}(1) \\ &= t^2(\lambda_2^2 - \lambda_2) \cdot 1_{K_*K} \otimes 1_{K_*MU} = (\lambda_2^2 - \lambda_2)u^2 \otimes 1_{K_*MU} \end{aligned}$$

show that these two elements are equal only if $\lambda_2 = 0$. Thus $Q = 0$, and so $MU_*(\varphi)$ vanishes on $\widetilde{MU}_*(\mathbb{C}P^\infty)$ because $MU_*(MU)$ has no torsion. $\square$

¹Here $\lambda_2 \in \mathbb{Z}$ denotes only the multiple of a generator.

Remark 3.5 One might be tempted to prove the previous lemma using the $\ker \varepsilon$ -adic filtration of the augmentation map $\varepsilon: MU_*(\text{pt}) \rightarrow \mathbb{Z} = H_*(\text{pt}; \mathbb{Z})$ (which is nothing but the Hurewicz homomorphism) because the associated graded rings $\text{gr}^\bullet MU_*(\mathbb{C}P^\infty)$ and $\text{gr}^\bullet MU_* MU$ are isomorphic (as rings) to $\text{gr}^\bullet MU_*(\text{pt}) \otimes H_*(\mathbb{C}P^\infty; \mathbb{Z})$ and $\text{gr}^\bullet MU_*(\text{pt}) \otimes H_*(MU)$, respectively, and the Pontrjagin ring structure induced by the formal group law is the $\text{gr}^\bullet MU_*(\text{pt})$ -linear extension of the Pontrjagin ring structure of $H_*(\mathbb{C}P^\infty; \mathbb{Z})$. However, under these identification, the associated map $\text{gr}^\bullet(MU_*(\varphi))$ corresponds to $\text{id}_{MU_*(\text{pt})} \otimes H_*(\varphi)$, so it cannot rule out the possibility that $MU_*(\varphi)(\beta_1) = \lambda_2$.

Remark 3.6 The weaker statement that every homotopy ring map $\varphi: \Sigma_+^\infty \mathbb{C}P^\infty \rightarrow MU$ induces the zero map on all homotopy groups of positive degree has close connection to the question of whether K is a retract (as a ring spectrum) of the periodic complex bordism spectrum MUP .

Both statements can be proved using the fact that $H_*(K; \mathbb{Z}) \cong \mathbb{Q}[t, t^{-1}]$, where $t \in H_2(K; \mathbb{Z})$ is the image of the Bott element $\beta \in \pi_2(K)$ under the Hurewicz homomorphism; see [21, Theorem 16.25]. Since $H_*(MUP; \mathbb{Z}) \cong \bigoplus_{i \in \mathbb{Z}} H_*(MU; \mathbb{Z})$, it is clear that there are no unital ring maps $f: K \rightarrow MUP$ and $g: MUP \rightarrow K$ such that $gf \simeq \text{id}_K$.

The proof that there are no ring maps that induce nontrivial homomorphisms on all homotopy groups is a bit more elaborate: Since $\pi_*(MU)$ is torsion-free, $\pi_*(\varphi)$ factors through $\pi_*(\Sigma_+^\infty \mathbb{C}P^\infty)/\text{torsion}$. The Hurewicz homomorphisms

$$\pi_*(\Sigma_+^\infty \mathbb{C}P^\infty)/\text{torsion} \rightarrow H_*(\mathbb{C}P^\infty; \mathbb{Z}) \quad \text{and} \quad \pi_*(MU) \rightarrow H_*(MU; \mathbb{Z})$$

are injective because the rational Hurewicz homomorphism is a natural equivalence between the homology theories $\pi_*(-) \otimes \mathbb{Q}$ and $H_*(-; \mathbb{Q})$. Naturality implies that $\pi_*(\varphi)$ is uniquely determined by $H_*(\varphi)$.

By the Pontrjagin ring structure of $H_*(\mathbb{C}P^\infty; \mathbb{Z}) \cong \mathbb{Z}\{1, \beta_1, \dots\}$ the homomorphism $H_*(\varphi)$ is uniquely determined by the image of β_1 . Thus there is a $\lambda \in \mathbb{Z}$ such that $H_*(\varphi)(\beta_1) = \lambda \beta_1$ and if $\pi_*(\varphi) \neq 0$, then $\lambda \neq 0$, too.

By Snaith's theorem [19], $MUP \simeq \Sigma_+^\infty BU[b_1^{-1}]$, so we have the following factorisation:

$$\begin{array}{ccccccc} \Sigma_+^\infty \mathbb{C}P^\infty & \xrightarrow{\varphi} & MU & \hookrightarrow & MUP & \longrightarrow & MUP[\lambda^{-1}] \\ \downarrow & & & & & & \parallel \\ \Sigma_+^\infty \mathbb{C}P^\infty[\beta_1^{-1}] & \xrightarrow{\bar{\varphi}} & \Sigma_+^\infty BU[b_1^{-1}, \lambda^{-1}] & & & & \end{array}$$

As the second and third upper horizontal maps induce injective maps on integral homology, the ring homomorphism $H_*(\bar{\varphi})$ is not trivial. But this is a contradiction, because, by Snaith's theorem [19], $\Sigma_+^\infty \mathbb{C}P^\infty[\beta_1^{-1}] \simeq K$ and every ring homomorphism

$$\mathbb{Q}[t, t^{-1}] \cong H_*(K, \mathbb{Z}) \rightarrow H_*(MUP[\lambda^{-1}]; \mathbb{Z}) \cong H_*(MUP; \mathbb{Z}) \otimes \mathbb{Z}[\lambda^{-1}]$$

is necessarily the zero map.

4 Proof of Theorem A

Theorem B immediately implies Theorem A once we show that the twist $\zeta\colon K(\mathbb{Z}, 3) \rightarrow BGL_1(M\operatorname{Spin}^c)$ in Theorem A induces nontrivial morphisms on homology groups. According to [11, page 52 ff] the twist $\zeta\colon K(\mathbb{Z}, 3) \rightarrow BGL_1(\operatorname{Spin}^c)$ is the delooping of the map $BU(1) = K(\mathbb{Z}, 2) \rightarrow GL_1(M\operatorname{Spin}^c)$ that is the adjoint of

$$M\iota\colon \Sigma_+^\infty BU(1) \rightarrow M\operatorname{Spin}^c,$$

the map between Thom spectra induced by the fibre inclusion $\iota\colon U(1) \rightarrow \operatorname{Spin}^c$.

Recall the canonical homomorphism $\xi\colon \operatorname{Spin}^c \xrightarrow{2\colon 1} \operatorname{SO} \times U(1)$. The representation that is used in the construction of the universal $\operatorname{Spin}^c(n)$ vector bundle factors through $\operatorname{SO}(n)$. That means, by definition,

$$M\operatorname{Spin}_n^c = \operatorname{Th}(E\operatorname{Spin}^c(n) \times_{\operatorname{pr}_1 \circ \xi} \mathbb{R}^n).$$

Thus the homomorphism $\xi\colon \operatorname{Spin}^c(n) \rightarrow \operatorname{SO}(n) \times U(1)$ induces maps

$$\begin{array}{ccc} \operatorname{Th}(E\operatorname{Spin}^c(n) \times_{\operatorname{pr}_1 \circ \xi} \mathbb{R}^n) & \longrightarrow & \operatorname{Th}(E(\operatorname{SO}(n) \times U(1)) \times_{\operatorname{pr}_1} \mathbb{R}^n) \\ \parallel & & \parallel \\ & & \operatorname{Th}((E\operatorname{SO}(n) \times_{\operatorname{torsion}} \mathbb{R}^n) \times BU(1)) \\ & & \parallel \\ M\operatorname{Spin}_n^c & \xrightarrow{M_n\xi} & M\operatorname{SO}_n \wedge BU(1)_+ \end{array}$$

These maps are compatible with the inclusion and therefore give rise to a map of spectra

$$M\operatorname{Spin}^c \xrightarrow{M\xi} M\operatorname{SO} \wedge BU(1)_+.$$

We will show that the composition

$$\Sigma_+^\infty BU(1) \xrightarrow{M\iota} M\operatorname{Spin}^c \xrightarrow{M\xi} M\operatorname{SO} \wedge BU(1)_+$$

induces nontrivial morphisms on rational homology groups. By the Thom isomorphism theorem, we obtain the commutative diagram

$$\begin{array}{ccc} H_*(\Sigma_+^\infty BU(1); \mathbb{Q}) & \xrightarrow{H_*(M\xi \circ M\iota)} & H_*(M\operatorname{SO} \wedge BU(1)_+; \mathbb{Q}) \\ \Phi \cong \downarrow & & \cong \downarrow \Phi \\ & & H_*(B\operatorname{SO} \times BU(1); \mathbb{Q}) \\ & & \downarrow H_*(\operatorname{pr}_2) \\ H_*(BU(1); \mathbb{Q}) & \xrightarrow{H_*(\operatorname{pr}_2 \circ B\xi \circ B\iota)} & H_*(BU(1); \mathbb{Q}) \end{array}$$

where Φ denotes the Thom isomorphism maps. Note that $\operatorname{pr}_2 \circ B\xi \circ B\iota$ is homotopic to $B\varphi$, where φ denotes the composition

$$U(1) \xrightarrow{\iota} \operatorname{Spin}^c \xrightarrow{\xi} \operatorname{SO} \times U(1) \xrightarrow{\operatorname{pr}_2} U(1).$$

The homomorphism φ is given by $\varphi(z) = z^2$.

Lemma 4.1 *The map $B\varphi$ induces nonvanishing homomorphisms on all rational homology groups in even degree.*

Proof Note that $B\varphi$ is an H -map. It induces the multiplication with two on $H_2(BU(1); \mathbb{Q})$ by the Hurewicz theorem. Recall that the Pontrjagin ring structure of $H_*(BU(1); \mathbb{Q}) \cong \mathbb{Q}\{1, \beta_1, \beta_2, \dots\}$ is given by

$$\beta_i \cdot \beta_j = \binom{i+j}{i} \beta_{i+j}.$$

Using this formula, we derive inductively

$$H_{2n}(B\varphi; \mathbb{Q})(\beta_n) = 2^n \beta_n,$$

which is obviously nonzero. □

Corollary 4.2 *The map $M\iota: \Sigma_+^\infty BU(1) \rightarrow M\mathrm{Spin}^c$ induces nontrivial homomorphisms on homology groups in even degree.*

This corollary together with [Theorem B](#) immediately implies the main theorem.

Proof of Theorem A Assume that we would have a factorisation up to homotopy

$$\begin{array}{ccc} K(\mathbb{Z}, 3) & \xrightarrow{\xi} & B\mathrm{GL}_1(M\mathrm{Spin}^c) \\ & \searrow T & \uparrow S \\ & & B\mathrm{GL}_1(MU) \end{array}$$

Then their loopings were homotopic, too. The adjoint maps then would satisfy

$$M\iota = \mathrm{Ad}(\Omega\xi) \simeq \mathrm{Ad}(\Omega S \circ \Omega T).$$

But we know from [Theorem B](#) that ΩT , and hence $M\iota = \mathrm{Ad}(\Omega S \circ \Omega T)$ too, is nullhomotopic. Thus, $M\iota$ would be trivial on all homology groups, which is a contradiction. □

5 Proof of Theorem D

Let R be a commutative ring with unit. Recall that a (one-dimensional commutative) *formal group law* over R is a formal power series $F \in R[[x, y]]$ that satisfies

$$F(x, 0) = F(0, x) = x, \quad F(x, y) = F(y, x) \quad \text{and} \quad F(x, F(y, z)) = F(F(x, y), z).$$

A *homomorphism* between formal group laws F and G is a formal power series $g \in R[[x]]$ with $g(0) = 0$ and $g(F(x, y)) = G(g(x), g(y))$.

By the work of Lazard, there is a ring L that carries a universal group law $F_{\mathrm{univ}} \in L[[x, y]]$ on it in the sense that, for any other formal group law F over a ring R , there is a unique ring homomorphism $\phi: L \rightarrow R$ such that $\phi(F_{\mathrm{univ}}(x, y)) = F(x, y)$. Quillen [\[17\]](#) identified L as the unitary cobordism ring

$MU^*(\text{pt})$. Under this identification, the universal group law agrees with the unitary bordism formal group law, meaning it is given by the formal power series

$$MU^*(\mu)(c) = \sum_{i,j \geq 0} \alpha_{ij} c^i d^j \in MU^*(\text{pt})[[c, d]],$$

where we use the same notation as in [Lemma 3.2](#).

Quillen's result allows us to deduce [Theorem D](#) from [Lemma C](#) by showing that these two statements are, in fact, equivalent.

Proof of Theorem D As $MU^*(\text{pt})[[c]]_0 = MU^0(\mathbb{C}P^\infty) = [\Sigma_+^\infty \mathbb{C}P^\infty; MU]$, we can represent every formal power series of pure degree zero

$$a = a(c) = \sum_{k \geq 0} a_k c^k \quad \text{with } a_k \in MU^{-2k}(\text{pt}),$$

by a map of spectra $f_a: \Sigma_+^\infty \mathbb{C}P^\infty \rightarrow MU$ that is unique up to homotopy. The map f_a is, by definition, a homotopy ring map if and only if the following diagram commutes up to homotopy:

$$\begin{array}{ccc} \Sigma_+^\infty \mathbb{C}P^\infty \wedge \Sigma_+^\infty \mathbb{C}P^\infty & \xrightarrow{f_a \wedge f_a} & MU \wedge MU \\ \downarrow \Sigma_+^\infty \mu & & \downarrow m \\ \Sigma_+^\infty \mathbb{C}P^\infty & \xrightarrow{f_a} & MU \end{array} \quad \text{and} \quad \begin{array}{ccc} S & & \\ \eta_{\Sigma_+^\infty \mathbb{C}P^\infty} \downarrow & \searrow \eta_{MU} & \\ \Sigma_+^\infty \mathbb{C}P^\infty & \xrightarrow{f_a} & MU \end{array}$$

Here m denotes the ring structure of MU and $\eta_{(-)}$ denotes the corresponding unit. Translating these diagrams into algebra gives

$$\mu^*(a) = a(c) \cdot a(d) \quad \text{and} \quad a(0) = 1,$$

because m represents the cohomology cross product; see [\[21, page 270\]](#). Quillen's theorem [\[17\]](#) yields

$$\mu^*(a) = \sum_{k \geq 0} a_k \mu^*(c)^k = \sum_{k \geq 0} a_k (F_{\text{univ}}(c, d))^k = a(F_{\text{univ}}(c, d)),$$

so that $a(F_{\text{univ}}(c, d)) = a(c) \cdot a(d)$. If we set $g = a - 1$, then we get

$$g(F_{\text{univ}}(c, d)) = g(c) + g(d) + g(c)g(d) = F_{\text{mult}}(g(c), g(d)) \quad \text{and} \quad g(0) = 0.$$

In conclusion, an element $g \in MU^*(\text{pt})[[c]]$ is a homomorphism between the universal formal group law and the multiplicative formal group law if and only if $g + 1$ is represented by a homotopy ring map. But [Lemma C](#) applied to the representing homotopy ring map of g implies that $g + 1 = 1$, so that $g = 0$. $\square$

References

- [1] **M Ando, A J Blumberg, D Gepner**, *Twists of K-theory and Tmf*, from “Superstrings, geometry, topology, and C^* -algebras”, Proc. Sympos. Pure Math. 81, Amer. Math. Soc., Providence, RI (2010) 27–63 [MR](#) [Zbl](#)

- [2] **M Ando, A J Blumberg, D Gepner, M J Hopkins, C Rezk**, *An ∞ -categorical approach to R -line bundles, R -module Thom spectra, and twisted R -homology*, J. Topol. 7 (2014) 869–893 [MR](#) [Zbl](#)
- [3] **M Ando, A J Blumberg, D Gepner, M J Hopkins, C Rezk**, *Units of ring spectra, orientations and Thom spectra via rigid infinite loop space theory*, J. Topol. 7 (2014) 1077–1117 [MR](#) [Zbl](#)
- [4] **M F Atiyah, R Bott, A Shapiro**, *Clifford modules*, Topology 3 (1964) 3–38 [MR](#) [Zbl](#)
- [5] **M Atiyah, G Segal**, *Twisted K -theory*, Ukr. Mat. Visn. 1 (2004) 287–330 [MR](#) [Zbl](#)
- [6] **D Barnes, C Roitzheim**, *Foundations of stable homotopy theory*, Cambridge Stud. Adv. Math. 185, Cambridge Univ. Press (2020) [MR](#) [Zbl](#)
- [7] **P Baum, M Joachim, M Khorami, T Schick**, *Twisted Spin^c -bordism, D -branes and twisted K -homology*, in preparation
- [8] **A L Carey, B-L Wang**, *Thom isomorphism and push-forward map in twisted K -theory*, J. K-Theory 1 (2008) 357–393 [MR](#) [Zbl](#)
- [9] **P E Conner, E E Floyd**, *The relation of cobordism to K -theories*, Lecture Notes in Math. 28, Springer (1966) [MR](#) [Zbl](#)
- [10] **P Donovan, M Karoubi**, *Graded Brauer groups and K -theory with local coefficients*, Inst. Hautes Études Sci. Publ. Math. 38 (1970) 5–25 [MR](#) [Zbl](#)
- [11] **F Hebestreit, M Joachim**, *Twisted spin cobordism and positive scalar curvature*, J. Topol. 13 (2020) 1–58 [MR](#) [Zbl](#)
- [12] **F Hebestreit, S Sagave**, *Homotopical and operator algebraic twisted K -theory*, Math. Ann. 378 (2020) 1021–1059 [MR](#) [Zbl](#)
- [13] **M J Hopkins, M A Hovey**, *Spin cobordism determines real K -theory*, Math. Z. 210 (1992) 181–196 [MR](#) [Zbl](#)
- [14] **S O Kochman**, *Bordism, stable homotopy and Adams spectral sequences*, Fields Inst. Monogr. 7, Amer. Math. Soc., Providence, RI (1996) [MR](#) [Zbl](#)
- [15] **J P May**, *Classifying spaces and fibrations*, Mem. Amer. Math. Soc. 155, Amer. Math. Soc., Providence, RI (1975) [MR](#) [Zbl](#)
- [16] **J Meier**, *Investigating twisted unitary cobordism spectra*, unpublished thesis (2021)
- [17] **D Quillen**, *On the formal group laws of unoriented and complex cobordism theory*, Bull. Amer. Math. Soc. 75 (1969) 1293–1298 [MR](#) [Zbl](#)
- [18] **D C Ravenel, W S Wilson**, *The Hopf ring for complex cobordism*, J. Pure Appl. Algebra 9 (1977) 241–280 [MR](#) [Zbl](#)
- [19] **V Snaith**, *Localized stable homotopy of some classifying spaces*, Math. Proc. Cambridge Philos. Soc. 89 (1981) 325–330 [MR](#) [Zbl](#)
- [20] **M Sugawara**, *On a condition that a space is an H -space*, Math. J. Okayama Univ. 6 (1957) 109–129 [MR](#) [Zbl](#)
- [21] **R M Switzer**, *Algebraic topology: homotopy and homology*, Grundle Math. Wissen. 212, Springer (1975) [MR](#) [Zbl](#)

School of Mathematics and Statistics, The University of Melbourne
Parkville, Australia

thorsten.hertl@unimelb.edu.au

Received: 14 November 2022 Revised: 2 October 2023

The generalized Kauffman–Harary conjecture is true

RHEA PALAK BAKSHI
HUIZHENG GUO
GABRIEL MONTOYA-VEGA
SUJOY MUKHERJEE
JÓZEF H PRZYTICKI

For a reduced alternating diagram of a knot with a prime determinant p , the Kauffman–Harary conjecture states that every nontrivial Fox p -coloring of the knot assigns different colors to its arcs. We prove a generalization of the conjecture, stated nineteen years ago by Asaeda, Przytycki and Sikora: for every pair of distinct arcs in the reduced alternating diagram of a prime link with determinant δ , there exists a Fox δ -coloring that distinguishes them.

57K10; 57M12

1 History of the alternation conjecture

In 1998, Louis H Kauffman and Frank Harary formulated the following conjecture [7]:

Alternation conjecture *Let D be a reduced, alternating diagram of a knot K having determinant p , where p is prime. Then every nontrivial p -coloring of D assigns different colors to different arcs.*

This conjecture is now known as the Kauffman–Harary conjecture. It was proved for rational knots [9; 13], Montesinos knots [1], some Turk’s head knots [6], and algebraic knots [5]. In 2009, Thomas W Mattman and Pablo Solis proved this conjecture using the notion of pseudocolorings. A generalization of this conjecture, known as the generalized Kauffman–Harary (GKH) conjecture, was formulated by Marta M Asaeda, Adam S Sikora and the fifth author in 2004 [1]. They proved this conjecture for Montesinos links in the same paper. In this paper, we prove it in full generality.

The paper is structured as follows. In the next section we introduce the GKH conjecture and we prove it in Section 3. In Section 4, we reformulate and prove the conjecture for nonprime alternating links. We illustrate the results with some examples in Section 5. In the last section, we discuss pseudocolorings, followed by some open questions.

2 Preliminaries

In this section, we state the original GKH conjecture and then some alternative versions. The difference between the original and generalized Kauffman–Harary conjecture is that the former is restricted to links with prime determinant. Note that the only link whose determinant is prime is the Hopf link.

Generalized Kauffman–Harary conjecture If D is a reduced alternating diagram of a prime link L , then different arcs of D represent different elements of $H_1(M_L^{(2)}, \mathbb{Z})$, where $M_L^{(2)}$ denotes the double branched cover of S^3 branched along L .

The GKH conjecture was formulated in [1] using the homology of the double branched cover of S^3 branched along L . We use a diagrammatic version of this conjecture by using the universal¹ group of Fox colorings $\text{Col}(D)$ for a prime link L with diagram D .

Definition 2.1 The group $\text{Col}(D)$ is the abelian group whose generators are indexed by the arcs of D , denoted by $\text{arcs}(D)$, and whose relations are $2b - a - c = 0$, given by the crossings of D . More precisely,

$$\text{Col}(D) = \left\{ \text{arcs}(D) \mid \begin{array}{c} a \quad b \\ \diagdown \quad \diagup \\ b \quad c = 2b - a \end{array} \right\}.$$

It is known that $\text{Col}(D) = \mathbb{Z} \oplus H_1(M_L^{(2)}, \mathbb{Z})$ (see for example [14]).

Definition 2.2 Let $\text{Col}^{\text{trivial}}(D) \cong \mathbb{Z}$ be the group of trivial colorings of D . This group is embedded in $\text{Col}(D)$ and the quotient group $\text{Col}(D)/\text{Col}^{\text{trivial}}(D)$ is called the *reduced group of Fox colorings*. We denote it by $\text{Col}^{\text{red}}(D)$.

Notice that, for a diagram D of a link L , $\text{Col}^{\text{red}}(D) = H_1(M_L^{(2)}, \mathbb{Z})$, and, for nonsplit alternating links, this group is finite with nonzero determinant.

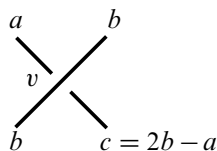
The first two statements of the following conjecture are equivalent to the original GKH conjecture, while part (c) offers an extension.

Conjecture 2.3 (alternative forms of the generalized Kauffman–Harary conjecture) Let D be a reduced alternating diagram of an alternating prime link and let $\delta(D)$ denote the absolute value of its determinant.

- (a) Let $\mathbb{Z}^{|\text{arcs}|}$ denote the free abelian group with basis set $\text{arcs}(D)$. The map $\mathbb{Z}^{|\text{arcs}|} \xrightarrow{\beta} \text{Col}(D)$ is injective on the arcs of D , that is, $\beta(a_i) \neq \beta(a_j)$ for $i \neq j$.
- (b) The diagram D has t Fox $\delta(D)$ -colorings $y_1, y_2, \dots, y_t$, such that, for every pair of distinct arcs a_i, a_j , there exists y_k such that $y_k(a_i) \neq y_k(a_j)$.
- (c) If $\text{Col}^{\text{red}}(D) = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \dots \oplus \mathbb{Z}_{n_s}$ with $n_{i+1} \mid n_i$, then there are s Fox n_1 -colorings such that, for each pair of arcs, one of these colorings distinguishes them.

Remark 2.4 Parts (a) and (b) of Conjecture 2.3 are equivalent to each other, since, for a finite group G , we have $G \cong \text{Hom}(G, \mathbb{Z}_{n_1}) \cong \text{Hom}(G, \mathbb{Z}_{\delta(D)})$, where $G = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \dots \oplus \mathbb{Z}_{n_s}$ with $n_{i+1} \mid n_i$ and $\delta(D) = n_1 n_2 \dots n_s$. In particular, $\text{Hom}(\text{Col}^{\text{red}}(D), \mathbb{Z}_{\delta(D)}) \cong \text{Hom}(\text{Col}^{\text{red}}(D), \mathbb{Z}_{n_1}) \cong \text{Col}^{\text{red}}(D)$. Thus, we can work with a group or its dual. To distinguish elements in the group, we often analyze its homomorphisms (dual elements) into the given ring. See [10], for example.

¹Analogous to the fundamental group and the fundamental quandle, this group is often called the fundamental group of Fox colorings.

Figure 1: Fox coloring relation at crossing v .

3 Proof of the generalized Kauffman–Harary conjecture

The proof of the GKH conjecture is organized as follows. First we define the crossing matrix $C'(D)$ and coloring matrix $L(D)$ of a link diagram D . Following [11], we prove that every column of the coloring matrix represents a nontrivial Fox $\delta(D)$ -coloring. Then, using the fact that the coloring matrix of the mirror image of D is the transpose of L , we prove part (b), and equivalently, part (a) of Conjecture 2.3. Additionally, we show that the columns of the coloring matrix generate the group $\text{Col}^{\text{red}}(D)$ and use this fact to prove part (c) of Conjecture 2.3.

Definition 3.1 A Fox k -coloring of a diagram D is a function $f: \text{arcs}(D) \rightarrow \mathbb{Z}_k$ satisfying the property that every arc is colored by an element of $\mathbb{Z}_k = \{0, 1, 2, 3, \dots, k-1\}$ in such a way that, at each crossing, the sum of the colors of the undercrossings is equal to twice the color of the overcrossing modulo k . That is, if at a crossing v the overcrossing is colored by b and the undercrossings are colored by a and c , then $2b - a - c \equiv 0$ modulo k . See Figure 1 for an illustration. The group of Fox k -colorings of a diagram D is denoted by $\text{Col}_k(D)$ and the number of Fox k -colorings is denoted by $\text{col}_k(D)$. Analogously to Definition 2.2, we quotient $\text{Col}_k(D)$ by the group of trivial colorings and denote this by $\text{Col}_k^{\text{red}}(D)$.

The matrix describing the space of colorings $\text{Col}(D)$ is referred to by Mattman and Solis [11] as the crossing matrix for a fixed arbitrary ordering of crossings and arcs. If the diagram D of n crossings has no trivial components, then this is an $n \times n$ matrix. For an alternating link diagram, we can order arcs in agreement with the ordering of crossings. Thus we have:

Definition 3.2 Fix an ordering of the crossings of a reduced, connected, alternating link diagram D of a nontrivial link. Then the set of arcs inherits the order of the set of crossings. In this way, the over-arc has the same index as the crossing. The crossing matrix² of D , denoted by $C'(D)$, is an $n \times n$ matrix such that each row corresponds to a crossing that gives the relation $2b - a - c = 0$ (see Figure 1). The entries of the matrix are defined as³

$$C'_{ij} = \begin{cases} 2 & \text{if } a_i \text{ is the over-arc at } c_i, \\ -1 & \text{if } a_j \text{ is an under-arc at } c_i \ (i \neq j), \\ 0 & \text{otherwise.} \end{cases}$$

²The alternative, more descriptive name could be *unreduced fundamental Fox colorings matrix*.

³It is possible that two under-arcs at a crossing are not distinct. Then the relation $2b - a - c = 0$ becomes $2b - 2a = 0$. For instance, this may occur for the Hopf link.

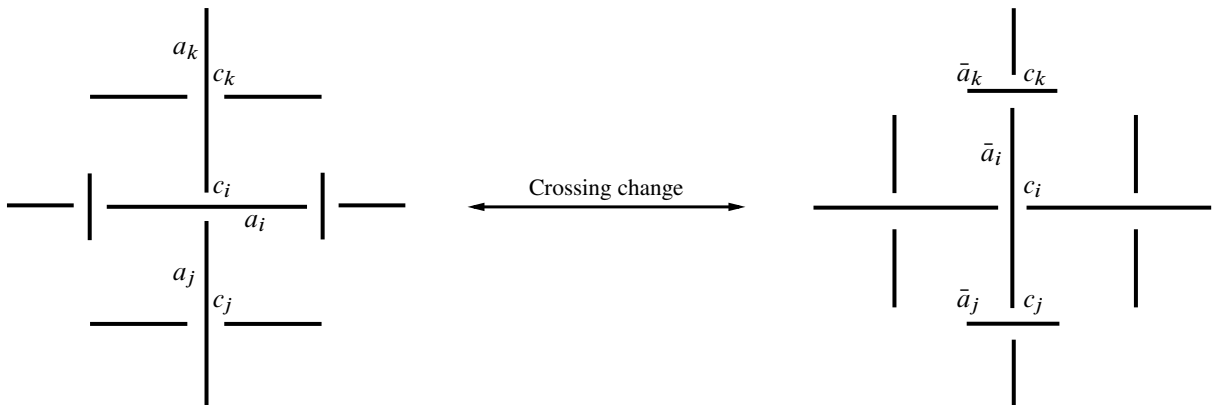


Figure 2: Neighborhood of the crossing c_i in D (on the left) and $\bar{D}$.

The following lemma holds only for alternating links and plays an important role in the proof of the GKH conjecture:

Lemma 3.3 *Let D be a reduced alternating link diagram with crossing matrix $C'(D)$ and let $\bar{D}$ be its mirror image. Then the matrix C'^T is a crossing matrix for $\bar{D}$.*

Proof Denote the crossings of the diagram D by $c_1, \dots, c_n$ and let the over-arc at the crossing c_i be denoted by a_i . Notice that, in the matrix $C'(D)$, all entries on the diagonal are 2. We obtain $\bar{D}$ by crossing-change operations and we keep the ordering and names of the crossings. Now let $\bar{a}_i$ denote the over-arc at the crossing c_i in the diagram $\bar{D}$. In the row corresponding to the crossing c_i , suppose the columns corresponding to the arcs a_j and a_k have -1 as entries. Then, in the matrix $C'(\bar{D})$, the column corresponding to $\bar{a}_i$ must have entries -1 in the rows corresponding to the crossings c_j and c_k ; see Figure 2. $\square$

Recall that, if $\delta(D) \neq 0$, then $\text{Col}^{\text{red}}(D)$ is a finite group whose invariant factor decomposition is $\text{Col}^{\text{red}}(D) = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \dots \oplus \mathbb{Z}_{n_s}$, with $n_{i+1} | n_i$ for all i . Notice that s is the minimum number of generators of this group and n_1 is the annihilator of the group. Let $C(D)$ denote the reduced crossing matrix of D , which is the matrix obtained from $C'(D)$ by removing its last row and last column. We call the arc corresponding to the last column of $C'(D)$ the *base arc*. This matrix describes the group $\text{Col}^{\text{red}}(D)$. The matrix $C^{-1}(D)$ is a matrix with rational entries. However, $n_1 C^{-1}(D)$ is an integral matrix, which we denote by $L_{n_1}(D)$. Observe that the columns of $L_{n_1}(D)$ modulo n_1 represent Fox n_1 -colorings of the diagram D after coloring the base arc by color 0.

The following result also holds for reduced nonalternating diagrams:

Theorem 3.4 *Let D be a reduced diagram of a link with nonzero determinant. Then the columns of $L_{n_1}(D)$ modulo n_1 generate the space of Fox n_1 -colorings of D .*

Proof Let $C(D)$ be the reduced crossing matrix of D and let $\text{Col}^{\text{red}}(D) = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_s}$ with $n_{i+1} | n_i$ for all i . After row and column operations, $C(D)$ can be reduced to its Smith normal form

$$C_{\text{SNF}}(D) = \begin{pmatrix} n_1 & & & & 0 \\ & n_2 & & & \\ & & \ddots & & \\ & & & n_s & \\ 0 & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Its inverse matrix, $C_{\text{SNF}}^{-1}(D)$, with entries in $\mathbb{Q}$ has the form

$$C_{\text{SNF}}^{-1}(D) = \begin{pmatrix} 1/n_1 & & & & 0 \\ & 1/n_2 & & & \\ & & \ddots & & \\ & & & 1/n_s & \\ 0 & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Thus, we obtain the integral matrix

$$L_{n_1}^{\text{SNF}}(D) = n_1 C_{\text{SNF}}^{-1}(D) = \begin{pmatrix} n_1/n_1 & & & & 0 \\ & n_1/n_2 & & & \\ & & \ddots & & \\ & & & n_1/n_s & \\ 0 & & & & n_1 & & \\ & & & & & \ddots & \\ & & & & & & n_1 \end{pmatrix}$$

Now the i^{th} column $(0, 0, \dots, n_1/n_i, \dots, 0)^T$ of $L_{n_1}^{\text{SNF}}(D)$ modulo n_1 with $i \leq s$ generates the subgroup $\mathbb{Z}_{n_i}$ of $\mathbb{Z}_{n_1}$. Since $\text{Col}_{n_1}^{\text{red}}(D) = \text{Hom}(\mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_s}, \mathbb{Z}_{n_1})$, the columns of $L_{n_1}^{\text{SNF}}$ generate $\text{Col}_{n_1}^{\text{red}}(D)$, as desired. $\square$

For alternating diagrams we can prove the following stronger result, which proves part (b) and, equivalently, part (a) of [Conjecture 2.3](#):

Theorem 3.5 *Let D be a reduced alternating diagram of a prime link. For any two arcs a_i and a_j , there exists a column of $L_{n_1}(D)$ which distinguishes them.*

Proof Suppose the arcs (indexing rows) of the coloring matrix $n_1 C^{-1}(D) = L_{n_1}(D)$ are given by $a_1, a_2, \dots, a_{n-1}$ as shown below:

$$L_{n_1}(D) = n_1 C^{-1}(D) = \begin{matrix} a_1 : \\ a_2 : \\ \vdots \\ a_{n-1} : \\ a_n : \end{matrix} \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n-1} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n-1} \\ \vdots & \vdots & \vdots & \vdots \\ c_{n-1,1} & c_{n-1,2} & \cdots & c_{n-1,n-1} \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Recall that the reduced crossing matrix $C(D)$ is obtained from $C'(D)$ by removing its last row and last column. Now each column of $L_{n_1}(D)$ colors the remaining first $n - 1$ arcs of the diagram. For a complete Fox n_1 -coloring of D , we color the last (base) arc a_n by color 0. If any $c_{i,1} = 0 \bmod n_1$ for $i < n$, then column C_1 modulo n_1 cannot distinguish between the arcs a_n and a_i . If all the entries of the rows corresponding to a_i and a_j are not identical modulo n_1 , then they can be automatically distinguished by the column in which they are different.

Step 1 If there is no column C_j of $L_{n_1}(D)$ such that $c_{i,j} \neq 0 \bmod n_1$, then every entry in the i^{th} row is $0 \bmod n_1$. It follows that, in the transpose matrix $L_{n_1}^T(D)$, the column C_i^T is the zero column modulo n_1 . This would result in the existence of a pseudocoloring of $\bar{D}$ (see [Definition 6.1](#) below [\[11\]](#)), which is a contradiction.⁴ Thus, the base arc a_n can be distinguished from any other arc by some column in $L_{n_1}(D)$ modulo n_1 .

Step 2 Furthermore, if there are two arcs a_i and a_j with the same color in every column of $L_{n_1}(D)$, then we choose arc a_j as the base arc, which implies that the colors of a_i are equal to zero. So we are back to Step 1. □

The next theorem proves part (c) of [Conjecture 2.3](#), which is a more general version of [Theorem 3.5](#).

Theorem 3.6 *If $\text{Col}^{\text{red}}(D) = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_s}$, with $n_{i+1} | n_i$, then there are s Fox n_1 -colorings (not necessarily corresponding to the columns of the coloring matrix) which distinguish all arcs. That is, for each pair of arcs, one of these colorings distinguishes them.*

Proof Denote the generators of the group $\text{Col}^{\text{red}}(D)$ by $a_1, a_2, \dots, a_s$. Every generator a_i is a linear combination of some columns of the coloring matrix $L_{n_1}(D)$ modulo n_1 (see [Theorem 3.4](#)). Therefore, they correspond to some coloring of the diagram D . Hence, for every pair of arcs there is a column of $L_{n_1}(D)$ modulo n_1 that distinguishes them. □

Remark 3.7 Let $\text{Col}^{\text{red}}(D)$ be the cyclic group $\mathbb{Z}_{n_1}$.

- (a) There exists a nontrivial Fox n_1 -coloring that distinguishes all arcs. This follows directly from [Theorem 3.6](#) for $s = 1$.
- (b) Additionally, if n_1 is a prime number, then the original Kauffman–Harary conjecture holds. That is, every nontrivial Fox n_1 -coloring distinguishes all arcs. This follows because every nonzero element of $\mathbb{Z}_{n_1}$ is its generator.

4 Nonprime alternating links

Theorems [3.5](#) and [3.6](#) do not hold as stated for the connected sum of alternating links⁵ (see part (a) of [Lemma 4.1](#)). In [Theorem 4.2](#), we present a version of the GKH conjecture which holds for nonprime alternating links.

⁴We work with links but the method of [\[11\]](#) can be extended to this case as well.
⁵The connected sum of alternating links is an alternating link; see for example [\[16\]](#).

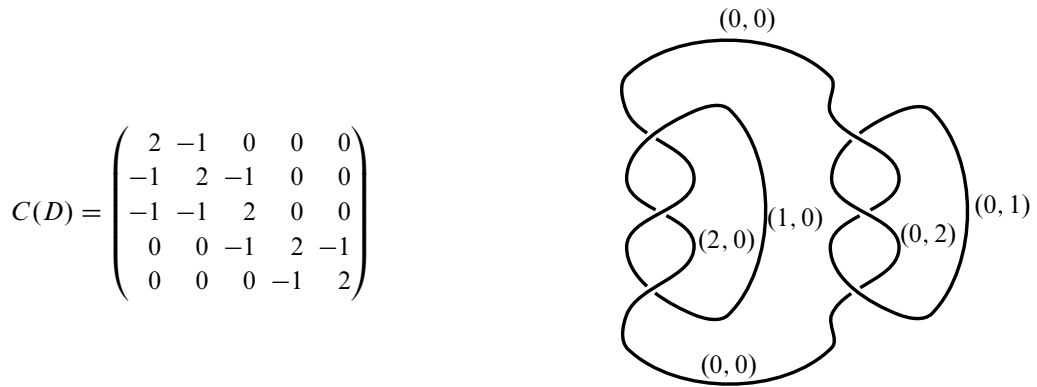


Figure 3: Left: the reduced crossing matrix for the square knot. Right: the square knot $\bar{3}_1 \# 3_1$ with two Fox 3-colorings distinguishing every pair of arcs.

Lemma 4.1 [14] Let $D = D_1 \# D_2$ be the connected sum of two link diagrams. Then:

- (a) The arcs connecting the two components represent the same element in $\text{Col}(D)$.
- (b) $\text{Col}^{\text{red}}(D_1 \# D_2) \cong \text{Col}^{\text{red}}(D_1) \oplus \text{Col}^{\text{red}}(D_2)$.

Theorem 4.2 Let $D = D_1 \# D_2 \# \cdots \# D_n$, where D_i is a reduced alternating diagram of a prime link L_i for $i = 1, 2, \dots, n$. Then:

- (a) For any pair of arcs different from arcs joining D_i with D_{i+1} , there exists a Fox n_1 -coloring which distinguishes them.
- (b) There are t ($t \leq s$) Fox n_1 -colorings such that any pair of arcs different from the ones joining D_i with D_{i+1} is distinguished by one of them.

Proof This result follows from Theorems 3.4 and 3.5 and Lemma 4.1. □

Remark 4.3 Theorem 4.2 was formulated for connected sums of diagrams. However, from William W Menasco’s result (see [12; 8]), it follows that, if an alternating diagram represents the connected sum of alternating links, then it is already a connected sum of diagrams.

Example 4.4 Let D be an alternating diagram of the square knot; that is, $D = \bar{3}_1 \# 3_1$, with reduced crossing matrix $C(D)$ (see Figure 3). Then $\text{Col}^{\text{red}}(\bar{3}_1 \# 3_1) = \mathbb{Z}_3 \oplus \mathbb{Z}_3$. Observe that columns 3 and 5 of $L_3(D)$ modulo 3 (Figure 4) distinguish all pairs of arcs except the ones connecting $\bar{3}_1$ with 3_1 . Also, the

$$L_3(D) = 3C^{-1}(D) = \begin{pmatrix} 3 & 2 & 1 & 0 & 0 \\ 3 & 4 & 2 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 2 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{pmatrix}, \quad L_3(D) \bmod 3 = \begin{pmatrix} 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{pmatrix}$$

Figure 4: Matrices $L_3(D)$ and $L_3(D)$ modulo 3 for the square knot.

$$L(7_7) = \begin{pmatrix} 24 & 20 & 12 & 10 & 16 & 11 \\ 12 & 24 & 6 & 12 & 15 & 9 \\ 15 & 16 & 18 & 8 & 17 & 13 \\ 6 & 5 & 3 & 13 & 4 & 8 \\ 18 & 22 & 9 & 11 & 26 & 10 \\ 12 & 10 & 6 & 5 & 8 & 16 \end{pmatrix}, \quad L(7_7) \bmod 21 = \begin{pmatrix} 3 & 20 & 12 & 10 & 16 & 11 \\ 12 & 3 & 6 & 12 & 15 & 9 \\ 15 & 16 & 18 & 8 & 17 & 13 \\ 6 & 5 & 3 & 13 & 4 & 8 \\ 18 & 1 & 9 & 11 & 5 & 10 \\ 12 & 10 & 6 & 5 & 8 & 16 \end{pmatrix}$$

Figure 5: Matrices $L(7_7)$ and $L(7_7) \bmod 21$. Some nontrivial Fox 21-colorings of 7_7 do not distinguish all arcs, for example columns 1 or 3 of L modulo 21. However, columns 2, 4, 5 and 6 distinguish all arcs.

third row (corresponding to the third crossing in the chosen ordering and, therefore, to the third arc) has all zero entries. That is, the third arc cannot be distinguished from the base arc.

5 Examples of Fox colorings

In this section we study examples of alternating link diagrams and their Fox colorings. For the structure of the group $\text{Col}^{\text{red}}(D) = H_1(M_D^{(2)}, \mathbb{Z})$ for knots up to 10 crossings, see [2, Appendix C].

Example 5.1 Kauffman and Harary showed that the knot 7_7 is a counterexample to their conjecture for a knot with nonprime determinant [7]. We have $\det(7_7) = 21$ and $\text{Col}^{\text{red}}(7_7) = \mathbb{Z}_{21}$.⁶ See Figure 6 for a Fox 21-coloring distinguishing all arcs.

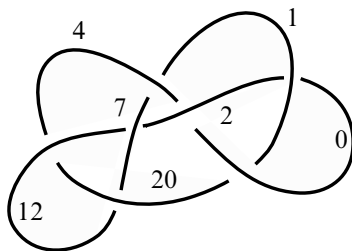


Figure 6: The knot 7_7 with a Fox 21-coloring which distinguishes all arcs.

Example 5.2 Consider the family of links obtained by closing the braids $(\sigma_1 \sigma_2^{-1})^n$. These links are sometimes called Turk's head links and can also be obtained by drawing the Tait diagrams of the wheel graphs W_n . The closed formula for the determinant of the $D(W_n)$ is given in [15]. Examples for $n = 5$ and $n = 6$ are drawn in Figure 7 and their reduced groups of Fox colorings are as follows:

- (a) For $n = 5$, $D(W_5)$ is 10_{123} in Rolfsen's table [17]. $\text{Col}^{\text{red}}(D(W_5)) = \mathbb{Z}_{11} \oplus \mathbb{Z}_{11}$.
- (b) For $n = 6$, $D(W_6)$ is the link 12^3_{474} in Thistlethwaite's tables [15; 18]. $\text{Col}^{\text{red}}(D(W_6)) = \mathbb{Z}_{40} \oplus \mathbb{Z}_8$.

A closed formula for $\text{Col}^{\text{red}}(\sigma_1 \sigma_2^{-1})^n$ is given in [4].

⁶It was noticed in [9] that the Kauffman–Harary conjecture holds for any rational (2-bridge) knot without restrictions on the determinant of the knot. However, as they note, the formulation of the conjecture needs to be changed from “every nontrivial Fox D -coloring” to “there exists a Fox D -coloring”. See Remark 3.7.

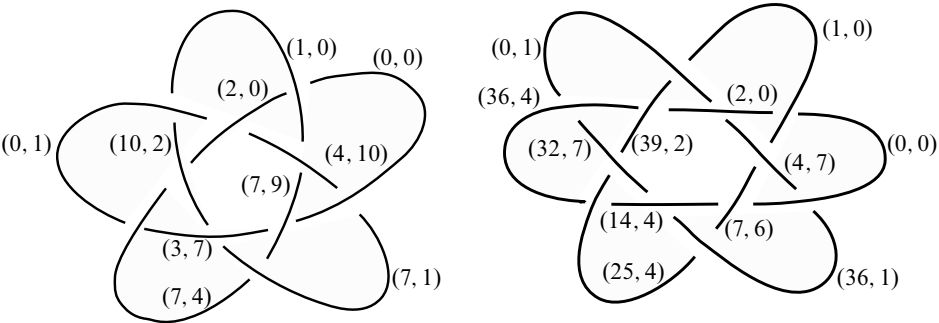


Figure 7: The knot $D(W_5)$ with two Fox colorings distinguishing all arcs (on the left), and the link $D(W_6)$ with two Fox colorings distinguishing all arcs (on the right).

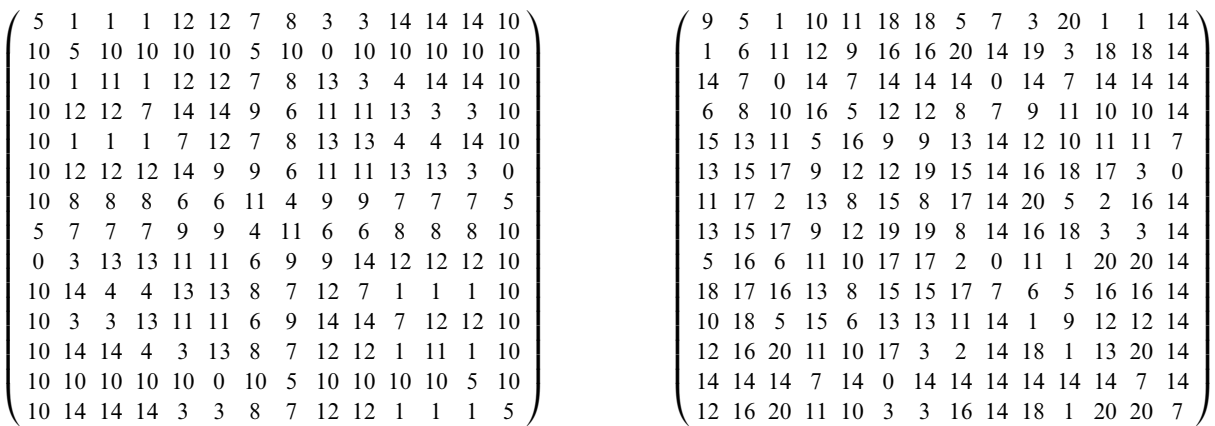


Figure 8: Left: $L(P(3,3,3,3,3))$ modulo 15. The colorings given by columns 3, 4 and 10 distinguish all arcs. Right: $L(P(3,3,3,3,6))$ modulo 21. The colorings given by columns 1 and 6 distinguish all arcs.

Example 5.3 The group Col^{red} for pretzel links is given in [1, Proposition 7] and its generalization to Montesinos links in [1, Proposition 8]. Here are two examples with their coloring matrices modulo n_1 :

- (a) Let $P(3,3,3,3,3)$ be a pretzel knot with 15 crossings. $\text{Col}^{\text{red}} = \mathbb{Z}_{15} \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3$. See its coloring matrix $L(P(3,3,3,3,3))$ modulo 15 in Figure 8, left.
- (b) Let $P(3,3,3,3,6)$ be a pretzel knot with 15 crossings. $\text{Col}^{\text{red}} = \mathbb{Z}_{21} \oplus \mathbb{Z}_3 \oplus \mathbb{Z}_3$. See its coloring matrix $L(P(3,3,3,3,6))$ modulo 21 in Figure 8, right.

6 Odds and ends

6.1 Pseudocolorings

An important tool in our proof of Theorem 3.5 is the idea of pseudocolorings. In [11] and in this paper, it is shown that no pseudocolorings exist for reduced, prime, alternating link diagrams. However, the

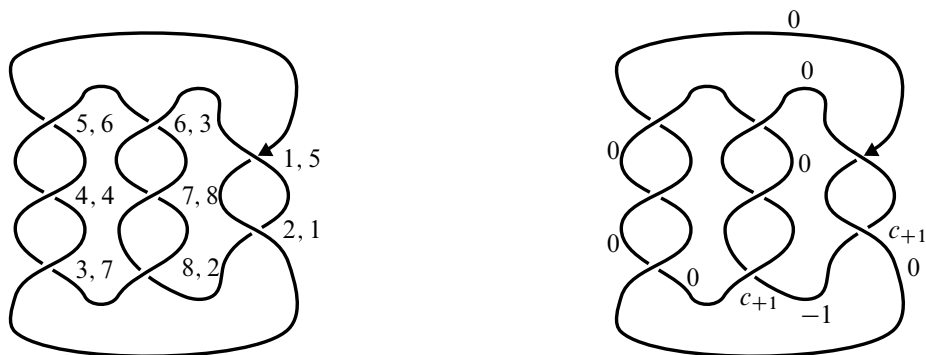


Figure 9: Left: the torus knot $T(3, 4)$ (8_{19} in Rolfsen's table [17]) depicted as the pretzel knot $P(3, 3, -2)$ showing the ordering of crossings and arcs. Right: the pseudocoloring given by the second column; compare Remark 6.5.

existence of pseudocolorings can be used to see how far a diagram is from being an alternating link diagram. In this section, we briefly explore this concept. Proposition 3.2 in [11] depends on the fact that, for reduced alternating diagrams, the rows of the crossing matrix add to zero. This does not hold for nonalternating diagrams, as we illustrate in the following examples.

Definition 6.1 Let D be a link diagram and $\epsilon \in \{-1, +1\}$. Following Mattman and Solis [11], we define an ϵ -pseudocoloring of D as a coloring of the arcs of D such that, at all but two crossings, the Fox coloring convention $2b - a - c = 0$ is satisfied. We denote the other two crossings by c_{+1} and c_{ϵ} , where the coloring conventions are $2b - a - c = +1$ and $2b - a - c = \epsilon$, respectively. To obtain the pseudocolorings as defined in [11], put $\epsilon = -1$.

For an alternating link diagram D , our convention was to order crossings first, and then the set of arcs inherits the order of the set of crossings. Compare Figure 1. The reason for such a choice is that $C'(\bar{D})$ is the same as $C'(D)^T$. This does not work for nonalternating link diagrams.

In general, we can arbitrarily order crossings and arcs. In Figure 9 we give an example of ordering crossings and arcs for the knot 8_{19} . We first choose a basepoint and an orientation (shown by an arrow in Figure 9, left). Starting at this basepoint, we move along the knot and order crossings. Next, arcs can be ordered arbitrarily with the base arc always being the last one. In Figure 9, the first coordinate gives the number of the crossing and the second one gives the number of the arc.

In the following example, we analyze nonsplit, nonprime alternating diagrams:

Example 6.2 Let $D = D_1 \# D_2$ be a nonsplit, nonprime alternating link diagram. D always has a -1 -pseudocoloring using color 1 on D_1 and color 0 on D_2 . We illustrate this idea for the square knot $\bar{3}_1 \# 3_1$ in Figure 10.

On the other hand, nonalternating link diagrams often have -1 -pseudocolorings and $+1$ -pseudocolorings. See Examples 6.3 and 6.4. If the determinant of a knot with diagram D is equal to 1, we have

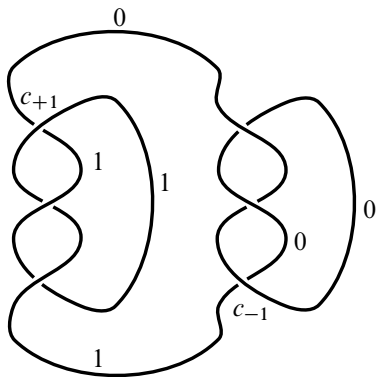


Figure 10: -1 -pseudocoloring of the square knot with the $+1$ -crossing denoted by c_{+1} and the -1 -crossing denoted by c_{-1} .

$$C'(D) = \begin{pmatrix} -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & -1 & -1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 2 \\ -1 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 2 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \end{pmatrix}$$

Figure 11: The crossing matrix of the Conway knot. Notice that the rows of the crossing matrix satisfy the linear equation $R_1 - R_2 - R_3 - R_4 - R_5 - R_6 - R_7 + R_8 + R_9 + R_{10} + R_{11} = 0$.

$L(D) = C^{-1}(D)$ and every column of $C^{-1}(D)$ colors the first $n - 1$ arcs of the diagram. Then, for a complete ϵ -pseudocoloring of D , we color the last (base) arc a_n by color 0.

Example 6.3 Consider the braid word $\sigma_2^3 \sigma_1 \sigma_3^{-1} \sigma_2^{-2} \sigma_1 \sigma_2^{-1} \sigma_1 \sigma_3^{-1}$ whose closure is the Conway knot. The determinant of this knot is 1 and its crossing matrix $C'(D)$ is given in Figure 11. The $+1$ -pseudocoloring given by column 4 and the -1 -pseudocoloring given by column 1 in the matrix shown in Figure 12 are illustrated in Figure 13, left and right, respectively.

Example 6.4 Consider the torus knot $T(3, 4)$ with diagram D and crossings and arcs ordered as illustrated in Figure 9. Its crossing matrix is shown in Figure 15. Three columns of $C^{-1}(D)$ (Figure 16) are integral and they yield ϵ -pseudocolorings. Column 5 gives a -1 -pseudocoloring (Figure 14, right) and columns 1 and 2 give $+1$ -pseudocolorings. The $+1$ -pseudocoloring corresponding to column 1 is shown in Figure 14, left.

Nonalternating link diagrams always have ϵ -pseudocolorings, as we describe in the following remark:

$$L(D) = \begin{pmatrix} 6 & -6 & -2 & 2 & -4 & -10 & -3 & 4 & 8 & 12 \\ -33 & 32 & 12 & -8 & 22 & 52 & 17 & -22 & -44 & -66 \\ -9 & 9 & 4 & -2 & 6 & 14 & 5 & -6 & -12 & -18 \\ 21 & -21 & -8 & 5 & -14 & -34 & -11 & 14 & 28 & 42 \\ -21 & 21 & 8 & -5 & 14 & 33 & 11 & -14 & -28 & -42 \\ 3 & -3 & -1 & 1 & -2 & -5 & -1 & 2 & 4 & 6 \\ 39 & -39 & -15 & 9 & -27 & -63 & -21 & 26 & 52 & 78 \\ 13 & -13 & -5 & 3 & -9 & -21 & -7 & 9 & 18 & 26 \\ -13 & 13 & 5 & -3 & 9 & 21 & 7 & -9 & -18 & -27 \\ 26 & -26 & -10 & 6 & -18 & -42 & -14 & 18 & 35 & 52 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Figure 12: Coloring matrix for the Conway knot. The last row of zeroes correspond to the coloring of the base arc.

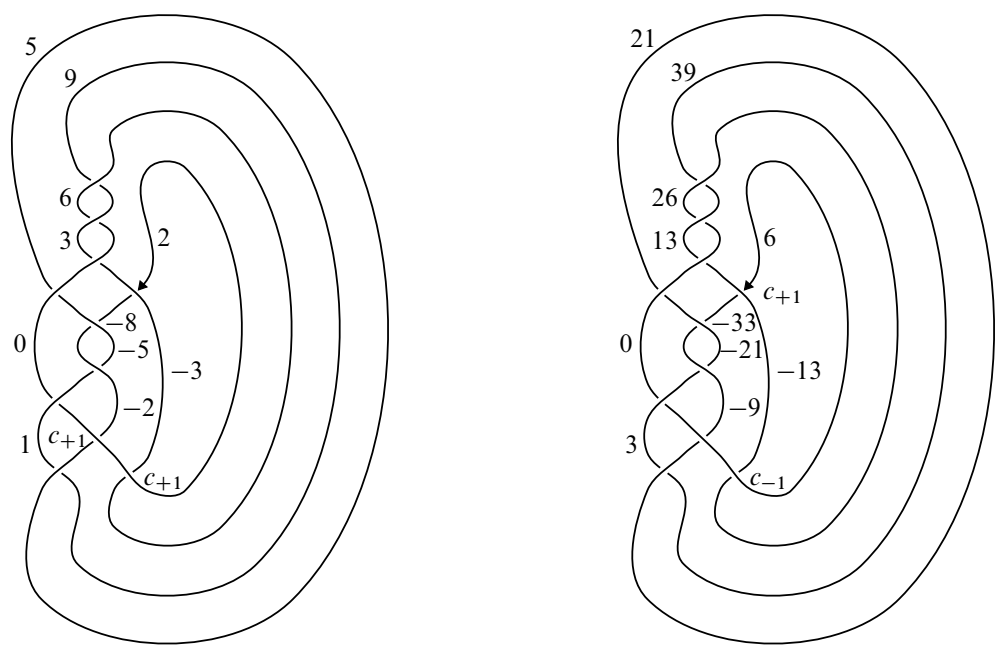


Figure 13: The Conway knot with +1-pseudocoloring (left) and -1 -pseudocoloring (right). The last crossing c_{+1} on the left changes to c_{-1} on the right.

Remark 6.5 Let D be a nonalternating link diagram.

- (1) Every integral column of $C^{-1}(D)$ leads to some ϵ -pseudocoloring.
- (2) D has an ϵ -pseudocoloring. This follows from the fact that every nonalternating diagram has a tunnel of length at least two. Now we can color D by coloring one of the arcs of the tunnel by color -1 and all other arcs by color 0 to get the $+1$ -pseudocoloring. An example of such a coloring is shown in [Figure 9](#), right.

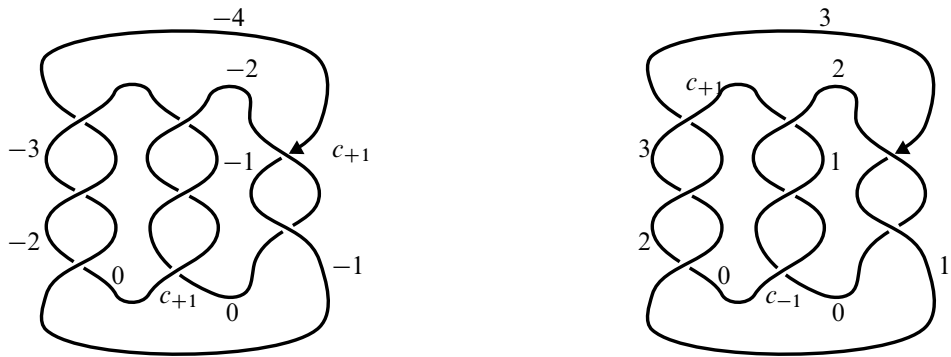


Figure 14: The torus knot $T(3, 4)$ (8_{19} in Rolfsen’s table [17]) depicted as the pretzel knot $P(3, 3, -2)$.

$$C'(D) = \begin{pmatrix} 2 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 & -1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 & 0 & -1 & -1 & 0 \\ 2 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 & 0 & -1 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Figure 15: $C'(P(3, 3, -2))$. The rows satisfy $R_1 + R_2 - R_3 - R_4 - R_5 - R_6 - R_7 - R_8 = 0$.

$$C^{-1}(D) = \begin{pmatrix} -2 & 0 & -\frac{2}{3} & \frac{2}{3} & 2 & \frac{10}{3} & \frac{5}{3} \\ 0 & -1 & \frac{4}{3} & \frac{2}{3} & 0 & -\frac{2}{3} & -\frac{1}{3} \\ -1 & 0 & -\frac{1}{3} & \frac{1}{3} & 1 & \frac{5}{3} & \frac{4}{3} \\ -3 & 0 & -1 & 1 & 3 & 4 & 2 \\ -1 & 0 & \frac{1}{3} & \frac{2}{3} & 1 & \frac{4}{3} & \frac{2}{3} \\ -4 & 0 & -\frac{5}{3} & \frac{2}{3} & 3 & \frac{16}{3} & \frac{8}{3} \\ -2 & 0 & -\frac{1}{3} & \frac{4}{3} & 2 & \frac{8}{3} & \frac{4}{3} \end{pmatrix}$$

Figure 16: $C^{-1}(D)$ corresponding to $T(3, 4)$ with three integral columns.

6.2 Future directions

The Kauffman–Harary conjecture was extended to the case of virtual knots by Mathew Williamson [19] and proved by Zhiyun Cheng [3]. A natural question is to ask whether the conjecture in [1] holds for virtual links whose determinants are not prime. Another path of further research is to look for a natural generalization to nonalternating diagrams using a set-theoretic Yang–Baxter operator or a general Yang–Baxter operator.

An interesting prospect is to approach the generalized Kauffman–Harary conjecture from the perspective of incompressible surfaces in the double branched cover $M_L^{(2)}$ of S^3 branched along L . This was outlined

in [1] with the hope of proving the GKH conjecture. Now that the GKH conjecture is proved, we can proceed in the opposite direction and analyze incompressible surfaces in $M_L^{(2)}$.

Acknowledgements

Bakshi acknowledges the support of Dr Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation. Montoya-Vega acknowledges the support of the National Science Foundation through grant DMS-2212736. Mukherjee was supported by the American Mathematical Society and the Simons Foundation through the AMS–Simons Travel Grant. Przytycki was partially supported by Simons Collaboration Grant 637794.

References

- [1] **MM Asaeda, J H Przytycki, A S Sikora**, *Kauffman–Harary conjecture holds for Montesinos knots*, J. Knot Theory Ramifications 13 (2004) 467–477 [MR](#) [Zbl](#)
- [2] **G Burde, H Zieschang**, *Knots*, 2nd edition, De Gruyter Studies in Mathematics 5, de Gruyter, Berlin (2003) [MR](#) [Zbl](#)
- [3] **Z Cheng**, *Kauffman–Harary conjecture for alternating virtual knots*, J. Knot Theory Ramifications 24 (2015) art. id. 1550046 [MR](#) [Zbl](#)
- [4] **A Christiana, H Guo, J H Przytycki**, *Using Fibonacci numbers and Chebyshev polynomials to express Fox coloring groups and Alexander–Burau–Fox modules of diagrams of wheel graphs*, preprint (2023) [arXiv 2310.19979](#) To appear in Contemp. Math.
- [5] **DB Damiano, EM Sennott**, *Coloring algebraic knots and links*, J. Knot Theory Ramifications 17 (2008) 553–578 [MR](#) [Zbl](#)
- [6] **NE Dowdall, T W Mattman, K Meek, P R Solis**, *On the Harary–Kauffman conjecture and Turk’s head knots*, Kobe J. Math. 27 (2010) 1–20 [MR](#) [Zbl](#)
- [7] **F Harary, L H Kauffman**, *Knots and graphs, I: Arc graphs and colorings*, Adv. in Appl. Math. 22 (1999) 312–337 [MR](#) [Zbl](#)
- [8] **J Hoste**, *The enumeration and classification of knots and links*, from “Handbook of knot theory” (W Menasco, M Thistlethwaite, editors), Elsevier, Amsterdam (2005) 209–232 [MR](#) [Zbl](#)
- [9] **L H Kauffman, S Lambropoulou**, *On the classification of rational tangles*, Adv. in Appl. Math. 33 (2004) 199–237 [MR](#) [Zbl](#)
- [10] **S Lang**, *Algebra*, 3rd edition, Graduate Texts in Math. 211, Springer (2002) [MR](#) [Zbl](#)
- [11] **T W Mattman, P Solis**, *A proof of the Kauffman–Harary conjecture*, Algebr. Geom. Topol. 9 (2009) 2027–2039 [MR](#) [Zbl](#)
- [12] **W Menasco**, *Closed incompressible surfaces in alternating knot and link complements*, Topology 23 (1984) 37–44 [MR](#) [Zbl](#)
- [13] **L Person, M Dunne, J DeNinno, B Guntel, L Smith**, *Colorings of rational, alternating knots and links*, preprint (2002)

- [14] **J H Przytycki**, *3-coloring and other elementary invariants of knots*, from “Knot theory” (V F R Jones, J Kania-Bartoszyńska, J H Przytycki, P Traczyk, V G Turaev, editors), Banach Center Publ. 42, Polish Acad. Sci. Inst. Math., Warsaw (1998) 275–295 [MR](#) [Zbl](#)
- [15] **J H Przytycki**, *From Goeritz matrices to quasi-alternating links*, from “The mathematics of knots” (M Banagl, D Vogel, editors), Contrib. Math. Comput. Sci. 1, Springer (2011) 257–316 [MR](#) [Zbl](#)
- [16] **J H Przytycki**, **R P Bakshi**, **D Ibarra**, **G Montoya-Vega**, **D Weeks**, *Lectures in knot theory: an exploration of contemporary topics*, Springer (2024) [MR](#) [Zbl](#)
- [17] **D Rolfsen**, *Knots and links*, Mathematics Lecture Series 7, Publish or Perish, Berkeley, CA (1976) [MR](#) [Zbl](#)
- [18] **M B Thistlethwaite**, *Knot tabulations and related topics*, from “Aspects of topology” (I M James, E H Kronheimer, editors), London Math. Soc. Lecture Note Ser. 93, Cambridge Univ. Press (1985) 1–76 [MR](#) [Zbl](#)
- [19] **M Williamson**, *Kauffman–Harary conjecture for virtual knots*, master’s thesis, University of South Florida (2007) Available at <https://digitalcommons.usf.edu/etd/3916>

RPB: *Institute for Theoretical Studies, ETH Zurich*
Zurich, Switzerland

Current address: *Department of Mathematics, University of California, Santa Barbara*
Santa Barbara, CA, United States

GM-V: *Department of Mathematics, The Graduate Center CUNY*
New York, NY, United States

Department of Mathematics, University of Puerto Rico at Río Piedras
San Juan, PR, United States

SM: *Department of Mathematics, University of Denver*
Denver, CO, United States

HG, JHP: *Department of Mathematics, The George Washington University*
Washington, DC, United States

JHP: *Department of Mathematics, University of Gdańsk*
Gdańsk, Poland

rheapalak@math.ucsb.edu, hguo30@gwu.edu, gabriel.montoyavega55@gc.cuny.edu,
sujoymukherjee.math@gmail.com, przytyck@gwu.edu

Received: 13 January 2023 Revised: 24 November 2023

Rigidity of elliptic genera for nonspin manifolds

MICHAEL WIEMELER

We discuss the rigidity of elliptic genera for nonspin manifolds M with S^1 -action. We show that if the universal covering of M is spin, then the universal elliptic genus of M is rigid. Moreover, we show that there is no condition which only depends on $\pi_2(M)$ that guarantees the rigidity in the case that the universal covering of M is nonspin.

[53C27](#), [57R15](#), [57R91](#), [57S15](#), [58J26](#)

1 Introduction

A Λ -genus is a ring homomorphism $\varphi: \Omega_*^{\text{SO}} \rightarrow \Lambda$, where Λ is a $\mathbb{C}$ -algebra and Ω_*^{SO} is the oriented bordism ring. For such a homomorphism one denotes by

$$g(u) = \sum_{i \geq 0} \frac{\varphi[\mathbb{C}P^{2i}]}{2i+1} u^{2i+1} \in \Lambda[[u]]$$

the logarithm of φ .

A Λ -genus φ is called *elliptic* if there are $\delta, \epsilon \in \Lambda$ such that its logarithm is given by

$$g(u) = \int_0^u \frac{dz}{\sqrt{1 - 2\delta z^2 + \epsilon z^4}}.$$

Examples of elliptic genera are the signature and the $\hat{A}$ -genus. For background material on elliptic genera see the books by Hirzebruch, Berger and Jung [10] and by Landweber et al. [12].

There is a universal elliptic genus $\phi: \Omega_*^{\text{SO}} \rightarrow \mathbb{C}[[q]]$ which actually takes values in $\mathbb{Q}[[q]]$. For an oriented even-dimensional manifold M the coefficients of $\phi[M]$ are indices of the signature operator on M twisted with some vector bundles (for a precise definition see Bott and Taubes [4, Section 2]). In the following we are mainly concerned with this genus (and its equivariant refinement). Therefore we call ϕ *the* elliptic genus.

If the oriented closed even-dimensional manifold M is acted on by the circle group S^1 , then $\phi[M]$ can be refined to the equivariant elliptic genus $\phi_{S^1}[M] \in R(S^1)[[q]] \otimes \mathbb{Q}$ of M , where $R(S^1) = \mathbb{Z}[\lambda, \lambda^{-1}]$ is the complex representation ring of S^1 . Note that $\phi_{S^1}[M]$ is an equivariant bordism invariant of M .

So $\phi_{S^1}[M]$ is of the form

$$\phi_{S^1}[M] = \sum_{n=0}^{\infty} \sum_{k \in \mathbb{Z}} a_{nk} \lambda^k q^n$$

with $a_{nk} \in \mathbb{Q}$ such that for fixed $n \geq 0$ there are at most finitely many $k \in \mathbb{Z}$ with $a_{nk} \neq 0$.

We call $\phi_{S^1}[M]$ rigid if $a_{nk} = 0$ whenever $k \neq 0$. It has been conjectured by Witten [17] that equivariant elliptic genera of closed, spin manifolds with S^1 -action are rigid. Here an oriented manifold is called spin if its second Stiefel–Whitney class vanishes.

Witten’s conjecture was proved by Bott and Taubes [4]. Alternative proofs have been given by Dessai [6], Liu [14], Taubes [15] and Wiemeler [16]. Herrera and Herrera [9] proposed a proof of the rigidity of elliptic genera of π_2 -finite, oriented manifolds. However, a mistake in the proof was found by Amann and Dessai [1]. They suggested that assuming π_2 - and π_4 -finiteness would be enough to come around this problem and to prove the rigidity (see the correction to [9]).

The aim of this paper is to clarify Lemma 1 of [9]. We find a counterexample to the original statement and obtain the same conclusion under additional hypothesis. We note that Amann and Dessai clarified the proof of Lemma 2 of [9].

Lemma 1 of [9] claims that if M is an even-dimensional orientable manifold acted on by S^1 , then the H -fixed point components in M which contain S^1 -fixed points are orientable for every subgroup $H \subset S^1$. We prove this claim under the extra assumption that the universal covering of M is a spin manifold. This then implies the rigidity of the elliptic genera of closed, oriented S^1 -manifolds M with spin universal cover.

We note that if the oriented, closed, connected manifold M has a finite cover $\tilde{M}$ which is spin then the rigidity of $\phi_{S^1}[M]$ can be deduced directly from the Bott–Taubes theorem. Indeed, by the Lefschetz fixed point formula we can assume that there are S^1 -fixed points in M . This implies that the action of S^1 on M lifts to an action on $\tilde{M}$. Moreover, since $\tilde{M} \rightarrow M$ is an equivariant covering map with $k < \infty$ sheets the equivariant elliptic genera of $\tilde{M}$ and M are related as follows:

$$k\phi_{S^1}[M] = \phi_{S^1}[\tilde{M}].$$

Therefore the rigidity for M follows from the rigidity for $\tilde{M}$.

However, we also note that there are manifolds with spin universal cover which do not admit any finite spin cover (see Ebert’s explanation [7]). Therefore we think it is worth stating and proving the following:

Theorem 1.1 *Let M be a closed, oriented, even-dimensional S^1 -manifold whose universal covering is spin. Then the equivariant elliptic genus of M is rigid.*

Using an argument which goes back to Hirzebruch and Slodowy [11] one sees that this theorem implies the following generalization of a result of Atiyah and Hirzebruch [2].

Corollary 1.2 *Let M be a closed, connected, oriented manifold whose universal covering is spin such that S^1 acts nontrivially on M . Then the $\hat{A}$ -genus of M vanishes.*

In [1] (see also the addendum) examples of oriented, closed, simply connected, effective S^1 -manifolds M with $\pi_2(M) = \mathbb{Z}_2$ and $\hat{A}(M) \neq 0$ were given. In view of the original claim in [9] and Corollary 1.2 one might ask if there is any condition on π_2 of a closed, connected, oriented, effective S^1 -manifold M

with nonspin universal covering which guarantees the vanishing of $\hat{A}(M)$. Using equivariant surgery, we give the following negative answer to this question.

Theorem 1.3 *Let V be a group such that there is a closed, connected, oriented manifold P of dimension $4n \geq 8$ with nonspin universal covering and $\pi_2(P) \cong V$. Then there is a closed, connected, oriented, effective S^1 -manifold M of dimension $4n$ with $\hat{A}(M) \neq 0$ and $\pi_2(M) \cong V$.*

To explain the proof of [Theorem 1.1](#) we have to recall the strategy of proof of Bott and Taubes [\[4\]](#).

Let M be a closed, spin, even-dimensional S^1 -manifold. Then the proof of the rigidity of $\phi_{S^1}[M]$ in [\[4\]](#) has two geometric inputs:

- (1) For every subgroup $H \subset S^1$, the fixed set M^H is orientable.
- (2) Orientations on M^H can be chosen in a specific way.

The first property is established in Lemma 10.1 of [\[4\]](#) and also follows from a theorem of Edmonds [\[8\]](#). How to choose the orientations for M^H is explained on pages 157–158 of [\[4\]](#). That the choice of orientations discussed there is possible is guaranteed by Lemma 8.1 of [\[4\]](#), which is a special case of Lemma 9.3 of [\[4\]](#). These lemmas also give relations between the weights of the isotropy representations at the various S^1 -fixed points in M .

From these two properties Bott and Taubes deduce the rigidity of elliptic genera using the Lefschetz fixed point formula and complex analysis. If M is not spin such that the S^1 -action has these two properties, then one can also prove the rigidity for this S^1 -action along the same lines. As noted by Herrera and Herrera [\[9\]](#), to do so one only needs the orientability of those components of M^H which contain S^1 -fixed points.

Our proof is completed by showing that this holds if the universal covering of M is spin.

This paper is structured as follows. We give a counterexample to Lemma 1 of [\[9\]](#) in [Section 2](#). In [Section 3](#) we then give a proof of this lemma under the extra assumption that the universal covering is spin. With this result we can complete the proof of [Theorem 1.1](#). In [Section 4](#) we then prove a refined version of [Theorem 1.3](#). Then in [Section 5](#) we give another example which shows that there is also a gap in the proof of Lemma 2 in [\[9\]](#). This gap is different from the mistake found by Amann and Dessai.

Acknowledgements

I would like to thank Manuel Amann and Anand Dessai for discussions on the subject of this paper. I would also like to thank Matthias Wink for comments on an earlier version of this paper. I thank the referee for comments which helped to improve the readability of this paper.

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044-390685587, Mathematics Münster: Dynamics–Geometry–Structure and within CRC 1442 Geometry: Deformations and Rigidity.

2 An example of an orientable S^1 -manifold with nonorientable singular strata

Lemma 1 of [9] claims that if M is an orientable S^1 -manifold of even dimension and $F \subset M^H$ is a H -fixed point component for some $H \subset S^1$ with $F^{S^1} \neq \emptyset$ then F is orientable. Here we give an example that shows that this claim is actually false.

For $i = 1, 2$ let V_i be the irreducible S^1 -representation of weight i . Let

$$M_1 = S(V_1^{k_1} \oplus V_2^{k_2} \oplus \mathbb{R})$$

be the unit sphere in the S^1 -representation $V_1^{k_1} \oplus V_2^{k_2} \oplus \mathbb{R}$, where S^1 acts trivially on $\mathbb{R}$ and $k_1 \geq 1$, $k_2 \geq 2$.

Then $F_1 := M_1^{\mathbb{Z}_2} = S(V_2^{k_2} \oplus \mathbb{R})$ is connected and contains both S^1 -fixed points. Moreover, by the slice theorem, there is an orbit $O_1 \subset F_1$ which has an invariant neighborhood in M_1 of the form

$$U_1 = S^1 \times_{\mathbb{Z}_2} D(\mathbb{R}_-^{2k_1} \oplus \mathbb{R}^{2k_2-1}),$$

where $\mathbb{R}_-$ is the nontrivial irreducible $\mathbb{Z}_2$ -representation and $\mathbb{R}$ is the trivial one-dimensional $\mathbb{Z}_2$ -representation. Here $D(V)$ denotes the unit disc in the representation V .

Moreover, let

$$M_2 = S^1 \times_{\mathbb{Z}_2} (P(\mathbb{R}^3 \oplus \mathbb{R}_-) \times S(\mathbb{R}_-^{2k_1-1} \oplus \mathbb{R}^{2k_2-2})).$$

Here $P(V)$ denotes the projectivization of the representation V . Note that M_2 is an orientable S^1 -manifold.

Moreover, the $\mathbb{Z}_2$ -fixed point component $F_2 := (S^1/\mathbb{Z}_2) \times P(\mathbb{R}^3) \times S(\mathbb{R}^{2k_2-2})$ in M_2 is nonorientable and contains an orbit O_2 with an invariant neighborhood in M_2 of the form

$$U_2 = S^1 \times_{\mathbb{Z}_2} D(\mathbb{R}_-^{2k_1} \oplus \mathbb{R}^{2k_2-1}).$$

Hence we can glue $M_1 \setminus \dot{U}_1$ and $M_2 \setminus \dot{U}_2$ along $\partial U_1 = \partial U_2$ to get an orientable S^1 -manifold M . The manifold M contains a $\mathbb{Z}_2$ -fixed point component F which can be constructed from F_1 and F_2 by gluing $F_1 \setminus \dot{U}_1^{\mathbb{Z}_2}$ and $F_2 \setminus \dot{U}_2^{\mathbb{Z}_2}$ along $\partial U_1^{\mathbb{Z}_2} = \partial U_2^{\mathbb{Z}_2}$. Since

$$U_1^{\mathbb{Z}_2} = U_2^{\mathbb{Z}_2} = (S^1/\mathbb{Z}_2) \times D(\mathbb{R}^{2k_2-1}),$$

F is the equivariant connected sum of F_1 and F_2 at the principal orbits O_1 and O_2 . So F is nonorientable and contains all S^1 -fixed points.

The problem in the proof of Lemma 1 of [9] seems to be a false application of Lemma 9.1 of [4]. This lemma gives a formula for the first Chern class $c_1(E) \in H^2(S^2; \mathbb{Z})$ of an equivariant complex vector bundle E over S^2 acted on effectively by S^1 by rotation. It says that

$$(2-1) \quad c_1(E)[S^2] = m_N - m_S,$$

where m_N and m_S are the sums of weights of the S^1 -representations on the fibers of E over the two fixed points $N, S \in S^2$, respectively.

In their proof, Herrera and Herrera construct spheres S^2 with S^1 -action and equivariant maps $f: S^2 \rightarrow M$ and apply the above formula to $E = f^*TM$. This bundle has a compatible complex structure by Lemma 9.2 of [4]. However, the circle actions on these spheres S^2 are in general not effective. They have principal isotropy group some $\mathbb{Z}_k \subset S^1$, $k \geq 1$. In this situation the formula (2-1) has to be corrected by multiplying the left hand side with k .

So if k is even the parity of the right hand side does not tell anything about the parity of the first Chern class $c_1(E)$. But that is needed for the proof.

3 Manifolds with universal covering that is spin

In this section we argue, that the proof of the rigidity of elliptic genera can be carried out under the assumption that the universal cover of the S^1 -manifold is spin. This gives the following theorem.

Theorem 3.1 *Let M be a closed, oriented, even-dimensional S^1 -manifold such that its universal cover admits a spin structure. Then the equivariant elliptic genus of M is rigid.*

As explained in the introduction, to prove the theorem, it suffices to show the following lemma.

Lemma 3.2 *Let M be an oriented, even-dimensional S^1 -manifold whose universal covering is spin. Moreover, let $F \subset M^H$ be a fixed point component of a subgroup $H \subset S^1$ which contains S^1 -fixed points. Then the following holds:*

- F is orientable.
- An orientation for F can be chosen as specified in [4, pages 157–158].

The basic observation which makes the proof of this lemma possible is the following:

Lemma 3.3 *An orientable manifold M has spin universal covering if and only if for every map $f: S^2 \rightarrow M$ we have $f^*w_2(TM) = 0$.*

Here $w_2(E) \in H^2(B; \mathbb{Z}_2)$ denotes the second Stiefel–Whitney class of the vector bundle $E \rightarrow B$. This lemma is well known (see [13, pages 88–89]). But for the sake of completeness we also give a proof here.

Proof We may assume that M is connected. Let $p: \tilde{M} \rightarrow M$ be the universal covering of M . Then by the Hurewicz theorem there is an isomorphism $\pi_2(\tilde{M}) \rightarrow H_2(\tilde{M}; \mathbb{Z})$. Moreover, by the universal coefficient theorem there is an isomorphism

$$H^2(\tilde{M}; \mathbb{Z}_2) \cong \text{hom}(H_2(\tilde{M}; \mathbb{Z}); \mathbb{Z}_2) \cong \text{hom}(\pi_2(\tilde{M}), \mathbb{Z}_2).$$

Therefore $\tilde{M}$ is spin if and only if for every map $f: S^2 \rightarrow \tilde{M}$ we have

$$0 = f^* w_2(T\tilde{M}) = f^* p^* w_2(TM) = (p \circ f)^* w_2(TM).$$

Since p is a covering we also have that $p_*: \pi_2(\tilde{M}) \rightarrow \pi_2(M)$ is an isomorphism. Hence, the claim follows. $\square$

Proof of Lemma 3.2 Using Lemma 3.3 one can modify the proof of Lemma 9.3 of [4] slightly to see that this lemma holds for the tangent bundle of M . Therefore the second claim in Lemma 3.2 follows once the first claim is shown. So we concentrate on the proof of the first claim.

Let M be an oriented, even-dimensional S^1 -manifold whose universal covering is spin. Let $F \subset M^H$ be a fixed point component of $H = \mathbb{Z}_k \subset S^1$ which contains an S^1 -fixed point. We can assume without loss of generality that H is the principal isotropy group of the induced S^1 -action on F . The irreducible real H -representations are given by

$$V_0, V_1, \dots, V_{[k/2]},$$

where V_0 is the trivial one-dimensional representation, for $0 < l < k/2$ we have $V_l = \mathbb{C}$ on which an element $z \in H \subset S^1 \subset \mathbb{C}$ acts by complex multiplication with z^l . If, moreover, k is even then $V_{k/2} = \mathbb{R}$ on which a generator of H acts by multiplication with -1 .

Accordingly there is a equivariant splitting

$$TM|_F = E_0 \oplus E_1 \oplus \dots \oplus E_{[k/2]}$$

into S^1 -invariant subbundles, such that for each $x \in F$, $E_i|_x \cong V_i^{l_i}$ as H -representations for some $l_i \geq 0$ and $i = 0, \dots, [k/2]$. Note that $E_0 = TF$. Moreover, each E_i , $0 < i < k/2$, has an S^1 -invariant complex structure.

So if k is odd, orientability of F follows from the orientability of M . Therefore assume that k is even. As in the proof of Lemma 1 of [9], we construct an S^1 -equivariant map

$$f: S^2 \rightarrow F,$$

such that

- (1) S^1 acts on S^2 by rotation with principal isotropy group H and fixed points N, S ,
- (2) $f(N) = f(S) \in F^{S^1}$,
- (3) $f^* w_2(E_0 \oplus E_{k/2}) \neq 0$ if F is not orientable.

To construct f we start with an equivariant embedding

$$g: S^1 \times (S^1/H) \hookrightarrow F,$$

here S^1 acts on the second factor of $S^1 \times (S^1/H)$ by multiplication. By moving the orbit $g(\{1\} \times (S^1/H))$ into an S^1 -fixed point, we can homotope g to a equivariant map

$$g_1: S^1 \times (S^1/H) \rightarrow F,$$

which is an embedding on $(S^1 \times (S^1/H)) \setminus (\{1\} \times (S^1/H))$ and collapses $\{1\} \times (S^1/H)$ to an S^1 -fixed point. Therefore there is a map $f: S^2 \rightarrow F$ with the properties (1) and (2) and such that $\text{im } f = \text{im } g_1 =: A$. Using Mayer–Vietoris sequences one sees that

$$g_1^*: H^2(A; \mathbb{Z}_2) \rightarrow H^2(S^1 \times (S^1/H); \mathbb{Z}_2), \quad f^*: H^2(A; \mathbb{Z}_2) \rightarrow H^2(S^2; \mathbb{Z}_2)$$

are isomorphisms. Hence it suffices to show that we can choose g such that $g^*w_2(E_0 \oplus E_{k/2}) = g_1^*w_2(E_0 \oplus E_{k/2}) \neq 0$ if F is nonorientable.

Note that, since the S^1 -action on M is orientation preserving, the codimension of all H' -fixed point components for $H' \subset S^1$ is even. Therefore the inclusion $F_0 \hookrightarrow F$ is one-connected. Here F_0 denotes the union of all principal orbits in F . Note that F_0 is an open and dense subset of F . Hence, if F is nonorientable, then F_0 is also nonorientable. Note that $\pi: F_0 \rightarrow F_0/S^1$ is a principal (S^1/H) -bundle.

Therefore we have

$$(3-1) \quad E_0|_{F_0} = TF_0 = \pi^*(T(F_0/S^1)) \oplus \mathbb{R}.$$

Hence, it follows that F_0/S^1 is a nonorientable manifold. Choose an embedding $g': S^1 \hookrightarrow F_0/S^1$ such that $g'^*w_1(T(F_0/S^1)) = a$ is a generator of $H^1(S^1; \mathbb{Z}_2)$.

Then the principal (S^1/H) -bundle $F_0|_{g'(S^1)}$ is trivial. Therefore g' lifts to an S^1 -equivariant embedding

$$g: S^1 \times (S^1/H) \hookrightarrow F_0.$$

It follows from (3-1) that

$$g^*E_0 = g^*TF_0 = g^*\pi^*(T(F_0/S^1)) \oplus \mathbb{R} = \pi^*g'^*(T(F_0/S^1)) \oplus \mathbb{R}.$$

Hence we have

$$g^*w_1(E_0) = a \in H^*(S^1 \times (S^1/H); \mathbb{Z}_2) = \Lambda_{\mathbb{Z}_2}(a, b), \quad g^*w_2(E_0) = 0.$$

Moreover, since S^1 acts transitively on S^1/H , we have $g^*E_{k/2} = E \otimes \gamma$, where E is a vector bundle over the first factor S^1 and $\gamma = S^1 \times_H V_{k/2}$. Note that $E_{k/2}$ and E have even rank since F has even codimension in M . Hence, we have $w_1(E) = g^*w_1(E_{k/2})$ and $g^*w_2(E_{k/2}) = w_1(E)w_1(\gamma) = w_1(E)b$.

Since $E_0 \oplus E_{k/2}$ is orientable, we have $w_1(E) = g^*w_1(E_{k/2}) = g^*w_1(E_0) = a$ and $g^*w_2(E_{k/2}) = ab$. Therefore we get

$$g^*w_2(E_0 \oplus E_{k/2}) = g^*w_2(E_{k/2}) + a^2 = g^*w_2(E_{k/2}) = ab \neq 0.$$

This completes the construction of f .

Since $E' = \bigoplus_{i=1}^{k/2-1} E_i$ has an S^1 -equivariant complex structure, we deduce from (2) and Lemma 9.1 of [4] that $f^*c_1(E') = 0$ and hence $f^*w_2(E') = 0$.

By Lemma 3.3, it follows that

$$0 = f^*w_2(TM) = f^*w_2(E') + f^*w_2(E_0 \oplus E_{k/2}) = f^*w_2(E_0 \oplus E_{k/2}).$$

Hence, by (3), F is orientable and the claim is proved. $\square$

Remark 3.4 The examples of the manifolds M and M_2 in Section 2 show that the first claim in Lemma 3.2 becomes false when one removes one of the assumptions that the universal covering of M is spin or that F contains a fixed point.

Using the arguments of [11, Section 1.5] or [9, Section 1.4] we get:

Corollary 3.5 *Let M be a closed, connected, oriented manifold whose universal covering is spin such that S^1 acts nontrivially on M . Then the $\hat{A}$ -genus of M vanishes.*

For a manifold M , denote by $h_{M_2}: \pi_2(M) \rightarrow H_2(M; \mathbb{Z})$ the Hurewicz map which sends a homotopy class of maps $f: S^2 \rightarrow M$ to the homology class $f_*[S^2]$, where $[S^2]$ denotes the fundamental class of S^2 . Note that Lemma 3.3 can be used to prove the following sufficient conditions for the universal covering being spin.

Lemma 3.6 *A connected, orientable manifold M has spin universal covering if one of the following three conditions holds:*

- (1) $\text{im } h_{M_2}$ is finite and of odd order.
- (2) $\text{im } h_{M_2}$ is finite and M is spin^c .
- (3) M is spin^c and $c_1(M)$ is torsion.

Here a oriented manifold M is called spin^c if $w_2(TM)$ is contained in the image of the natural map $H^2(M; \mathbb{Z}) \rightarrow H^2(M; \mathbb{Z}_2)$. Moreover, if this is the case then for every preimage $b \in H^2(M; \mathbb{Z})$ of $w_2(TM)$ there is a spin^c structure on M with $c_1(M) = b$ (see [13, Appendix D]).

Proof of Lemma 3.6 Let $f: S^2 \rightarrow M$ be a map. By Lemma 3.3, we have to show that under any of the three conditions $f^*w_2(TM) = 0$.

First assume that (3) holds. Then since $H^2(S^2; \mathbb{Z})$ is torsion-free, we have $0 = f^*c_1(M) \equiv f^*w_2(TM) \pmod{2}$.

Next assume that $\text{im } h_{M^2}$ is finite. Let $p: S^2 \rightarrow S^2$ be a map of degree $|\text{im } h_{M^2}|$. Then $f \circ p$ is null homologous. Hence we have

$$0 = p^* f^* w_2(TM) = |\text{im } h_{M^2}| f^* w_2(TM).$$

So, if $|\text{im } h_{M^2}|$ is odd, the claim follows.

If M is spin^c , then we can do the above calculation with $c_1(M)$ instead of $w_2(TM)$ and conclude that $0 = f^* c_1(M) \equiv f^* w_2(TM) \pmod{2}$. $\square$

[Lemma 3.6](#) together with [Theorem 3.1](#) and [Corollary 3.5](#) gives the following corollary.

Corollary 3.7 *Let M be a closed, connected, oriented, even-dimensional S^1 -manifold such that one of the following conditions is satisfied:*

- (1) $\text{im } h_{M^2}$ is finite and of odd order.
- (2) $\text{im } h_{M^2}$ is finite and M is spin^c .
- (3) M is spin^c and $c_1(M)$ is torsion.

Then the equivariant elliptic genus of M is rigid. Moreover, the $\hat{A}$ -genus of M vanishes if the action is nontrivial.

Conditions (1) and (2) in this corollary are quite close to the original condition of Herrera and Herrera for the rigidity of elliptic genera. However, the example of Amann and Dessai [\[1\]](#) and [Corollary 4.3](#) below show that the conditions after the “and” cannot be removed.

4 Equivariant surgery

In this section we prove a refined version of [Theorem 1.3](#) from the introduction. To do so we have to set up some notation.

Let M be a connected manifold with universal covering $p: \tilde{M} \rightarrow M$. Then from the Hurewicz theorem we get an isomorphism

$$\pi_2(M) \cong \pi_2(\tilde{M}) \cong H_2(\tilde{M}; \mathbb{Z}).$$

Since $\pi_1(M)$ is acting on $\tilde{M}$, we get a $\mathbb{Z}[\pi_1(M)]$ -module structure on $\pi_2(M)$. Here $\mathbb{Z}[\pi_1(M)]$ is the group ring of $\pi_1(M)$. In the rest of this section this $\mathbb{Z}[\pi_1(M)]$ -module structure on $\pi_2(M)$ is assumed.

If $\dim M \geq 4$ and M' is constructed from M by taking the connected sum of M with a simply connected manifold N in a point $x \in M$, then the universal covering $\tilde{M}'$ of M' is constructed from $\tilde{M}$ by taking the connected sum with copies of N in every point in $p^{-1}(x)$. In particular,

$$\pi_1(M') \cong \pi_1(M), \quad \pi_2(M') \cong \pi_2(M) \oplus \mathbb{Z}[\pi_1(M')] \otimes \pi_2(N).$$

Similarly, if $\dim M \geq 6$ and M' is constructed from M by surgery on an embedded sphere $S^2 \subset M$, then $p^{-1}(S^2)$ is the disjoint union of embedded spheres in $\tilde{M}$ and $\tilde{M}'$ is constructed from $\tilde{M}$ by surgery on all of these spheres. In particular,

$$\pi_1(M') \cong \pi_1(M), \quad \pi_2(M') \cong \pi_2(M)/\mathbb{Z}[\pi_1(M')][\phi],$$

where $\phi: S^2 \rightarrow M$ is the inclusion of the embedded sphere on which we do surgery.

Now we can state and prove the main result of this section.

Theorem 4.1 *Let G be a finitely presented group and V a $\mathbb{Z}[G]$ -module such that there is a closed, connected, oriented manifold P of dimension $4n \geq 8$ with nonspin universal covering and*

$$(4-1) \quad \pi_1(P) \cong G, \quad \pi_2(P) \cong V.$$

Then there exists a closed, connected, oriented, effective S^1 -manifold M of dimension $4n$ such that:

- $\pi_1(M) \cong G, \pi_2(M) \cong V$.
- *The universal covering of M is nonspin.*
- *M is equivariantly bordant to a linear S^1 -action on $\mathbb{C}P^{2n}$. In particular, $\hat{A}(M) = \hat{A}(\mathbb{C}P^{2n}) \neq 0$.*

Proof Recall first that if M is a manifold and $c \in H^2(M; R)$ for some ring R then c induces a homomorphism

$$\pi_2(M) \rightarrow R, \quad [f] \mapsto f^*c[S^2].$$

We denote this homomorphism also by c . Note that the homomorphism c is invariant under the action of $\pi_1(M)$ on $\pi_2(M)$.

Next we recall some properties of S^1 -manifolds and their orbit spaces.

Let N be a connected effective S^1 -manifold of dimension $m \geq 7$ such that:

- (1) For each nontrivial $H \subset S^1$, $\text{codim } M^H \geq 4$.
- (2) There are S^1 -fixed points in N .

We denote by N_0 the union of principal orbits in N . Then N_0 is an open dense subset of N and by (1) the inclusion $\iota: N_0 \hookrightarrow N$ is 3-connected. Moreover, by (2), the map

$$\mathbb{Z} = \pi_1(S^1) \rightarrow \pi_1(N_0) \cong \pi_1(N)$$

induced by the inclusion of an orbit is trivial.

By the exact homotopy sequence for the fibration $p: N_0 \rightarrow N_0/S^1$ we see that there is a exact sequence

$$0 \rightarrow \pi_2(N_0) \xrightarrow{p_*} \pi_2(N_0/S^1) \xrightarrow{c_1(N_0)} \mathbb{Z} \rightarrow 0.$$

Here we have identified the connecting homomorphism $\pi_2(N_0/S^1) \rightarrow \pi_1(S^1)$ in the exact sequence with the homomorphism induced by the first Chern class $c_1(N_0) \in H^2(N_0/S^1; \mathbb{Z})$ of the S^1 -bundle $N_0 \rightarrow N_0/S^1$. Moreover, we see that $p_*: \pi_1(N_0) \rightarrow \pi_1(N_0/S^1)$ is an isomorphism.

Since $p^*T(N_0/S^1)$ is stably isomorphic to TN_0 we see that for $a \in \pi_2(N_0)$ we have

$$w_2(TN)(\iota_*a) = w_2(TN_0)(a) = w_2(T(N_0/S^1))(p_*a).$$

Now let $\phi: S^2 \times D^{m-3} \hookrightarrow N_0/S^1$, such that $c_1(N_0)([\phi|_{S^2 \times \{0\}}]) = 0$. Then there is an equivariant embedding

$$\tilde{\phi}: S^1 \times S^2 \times D^{m-3} \hookrightarrow N_0 \hookrightarrow N,$$

where S^1 acts on the first factor by multiplication, covering ϕ .

In this situation we can do equivariant surgery on $\tilde{\phi}$, ie construct an S^1 -manifold

$$N' = (N \setminus \tilde{\phi}(S^1 \times S^2 \times \dot{D}^{m-3})) \cup_{S^1 \times S^2 \times S^{m-4}} S^1 \times D^3 \times S^{m-4}$$

by equivariant gluing. On the level of orbit spaces N'_0/S^1 is constructed from N_0/S^1 by surgery on ϕ . Moreover, N' is equivariantly bordant to N and satisfies (1) and (2) from above.

In particular we have $\pi_1(N'_0/S^1) \cong \pi_1(N_0/S^1)$ and

$$\pi_2(N'_0/S^1) \cong \pi_2(N_0/S^1)/\mathbb{Z}[\pi_1(N_0/S^1)][\phi|_{S^2 \times \{0\}}].$$

Moreover, $c_1(N'_0)$ and $w_2(T(N'_0/S^1))$ are the homomorphisms induced from $c_1(N_0)$ and $w_2(T(N_0/S^1))$ on this quotient, respectively.

After these general remarks we describe the construction of M . Let $N = \mathbb{C}P^{2n}$ equipped with a linear effective S^1 -action such that (1) and (2) hold. Moreover, let P' be the $(4n-1)$ -manifold from [Lemma 4.4](#) below. Then (4-1) holds for P' and the universal covering of P' is nonspin.

Let $\phi_1: D^{4n-1} \hookrightarrow N_0/S^1$ and $\phi_2: D^{4n-1} \hookrightarrow P'$ be embeddings. Then form the equivariant connected sum

$$N' = (N \setminus p^{-1}(\phi_1(\dot{D}^{4n-1}))) \cup_{S^{4n-2} \times S^1} (P' \setminus \phi_2(\dot{D}^{4n-1})) \times S^1.$$

Then N' satisfies (1) and (2) and is equivariantly bordant to N since $P' \times S^1$ is an equivariant boundary. On the level of orbit spaces N'_0/S^1 is the connected sum of N_0/S^1 and P' . Therefore we have $\pi_1(N'_0/S^1) = \pi_1(P') = G$ and, moreover

$$\pi_2(N'_0/S^1) \cong \pi_2(P') \oplus \mathbb{Z}[G] \otimes \pi_2(N_0/S^1) = \pi_2(P') \oplus \mathbb{Z}[G]a \oplus \mathbb{Z}[G]b,$$

where a, b is a $\mathbb{Z}$ -basis of $\pi_2(N_0/S^1)$ such that a is a generator of the image of $\pi_2(N_0) \rightarrow \pi_2(N_0/S^1)$ and $c_1(N_0)(b) = 1$. Note that $c_1(N'_0)|_{\pi_2(P')} = 0$ since the restriction of the S^1 -bundle $N'_0 \rightarrow N'_0/S^1$ to $P' \setminus \phi_2(\dot{D}^{4n-1})$ is trivial.

Let $g_1, \dots, g_k$ be generators of G . Then the $(1 - g_i)b$, $i = 1, \dots, k$, are in the kernels of both $c_1(N'_0)$ and $w_2(T(N'_0/S^1))$. So we can represent these homotopy classes by disjointly embedded spheres with trivial normal bundle. We do equivariant surgery on these spheres to get a S^1 -manifold N'' such that (1) and (2) hold for N'' and $\pi_1(N''_0/S^1) \cong G$ and

$$\pi_2(N''_0/S^1) \cong V \oplus \mathbb{Z}[G]a \oplus \mathbb{Z}b.$$

By Lemma 3.3, we can find $c \in \pi_2(P') \cong V$ such that $w_2(TP')(c) = w_2(T(N_0/S^1))(a)$. Then $c_1(N''_0)$ and $w_2(T(N''_0/S^1))$ are trivial on $c + a \in \pi_2(N''_0/S^1)$. So we can represent $c + a$ by an embedding of a sphere with trivial normal bundle and do equivariant surgery on this sphere to get an S^1 -manifold M . For this M , (1) and (2) hold. Moreover, $\pi_1(M_0/S^1) \cong G$, $\pi_2(M_0/S^1) \cong V \oplus \mathbb{Z}b$. Since $c_1(M_0)(b) = 1$, and $c_1(M_0)|_V = 0$ we see from the exact homotopy sequence for the fibration $M_0 \rightarrow M_0/S^1$ that $\pi_1(M_0) = G$ and $\pi_2(M_0) = V$. Since $M_0 \rightarrow M$ is 3-connected the claim follows. $\square$

From Theorem 4.1 we immediately get the following corollary which is Theorem 1.3 from the introduction.

Corollary 4.2 *Let V be a group such that there is a closed, connected, oriented manifold P of dimension $4n \geq 8$ with nonspin universal covering and $\pi_2(P) \cong V$. Then there is a closed, connected, oriented S^1 -manifold M of dimension $4n$ with $\hat{A}(M) \neq 0$ and $\pi_2(M) \cong V$.*

The following corollary shows that the condition (1) of Corollary 3.7 gives the optimal condition which only depends on $\text{im } h_{M_2}$ and guarantees the vanishing of the $\hat{A}$ -genus of a closed, oriented S^1 -manifold M .

Corollary 4.3 *Let V be a finitely generated, abelian group. Assume that V is either infinite or finite and of even order. Then there is a simply connected, closed, oriented manifold M of dimension $4n \geq 8$ such that:*

- $H_2(M; \mathbb{Z}) = \pi_2(M) = V$.
- S^1 acts effectively on M .
- M is equivariantly bordant to a linear S^1 -action on $\mathbb{C}P^{2n}$. In particular, $\hat{A}(M) = \hat{A}(\mathbb{C}P^{2n}) \neq 0$.

Proof By Theorem 4.1 and the Hurewicz theorem it suffices to show that there is a closed oriented $4n$ -manifold P which is simply connected and nonspin such that $H_2(P; \mathbb{Z}) = V$.

To do so first note that by our assumption on V there is a nontrivial homomorphism $\psi: V \rightarrow \mathbb{Z}_2$. Moreover, the group V fits into a exact sequence

$$(4-2) \quad 0 \rightarrow \mathbb{Z}^{k_1} \xrightarrow{\iota} \mathbb{Z}^{k_2} \xrightarrow{\phi} V \rightarrow 0$$

with some $0 \leq k_1 \leq k_2 < \infty$. Let $a_1, \dots, a_{k_1}$ be the standard basis of $\mathbb{Z}^{k_1}$.

Denote by W the total space of the nontrivial S^{4n-2} -bundle with structure group $\text{SO}(4n-1)$ over S^2 . Then by taking the connected sum of several copies of W and $S^2 \times S^{4n-2}$, we construct a simply

connected manifold P_1 with $H_2(P_1; \mathbb{Z}) = \mathbb{Z}^{k_2}$ and such that $w_2(TP_1) \in H^2(P_1; \mathbb{Z}_2)$ corresponds to $\psi \circ \phi$ under the isomorphism $H^2(P_1; \mathbb{Z}_2) \cong \text{hom}(H_2(P_1; \mathbb{Z}), \mathbb{Z}_2)$.

Since (4-2) is exact we can represent the homology classes corresponding to the $\iota(a_i)$ by disjointly embedded spheres in P_1 with trivial normal bundles. By doing surgery on these spheres we get a manifold P as required. $\square$

The following lemma is used in the proof of Theorem 4.1.

Lemma 4.4 *Let P be a closed, connected, oriented manifold of dimension $n \geq 7$. Then there exists a closed, connected, oriented manifold P' such that*

- $\dim P' = n - 1$,
- $\pi_1(P') \cong \pi_1(P)$, $\pi_2(P') \cong \pi_2(P)$,
- the universal covering of P' is spin if and only if the universal covering of P is spin.

Proof Let $f: P \rightarrow]-1, n+1[$ be a Morse-function on P such that $f(x) \in](4i-1)/4, (4i+1)/4[$ for all critical points x of f with index $i = 0, \dots, n$. Let $W = f^{-1}(]-1, \frac{7}{2}])$ and $P' = f^{-1}(\frac{7}{2}) = \partial W$. Then P can be constructed from W by attaching handles of dimension at least 4. Similarly W can be constructed from $f^{-1}([\frac{13}{4}, \frac{7}{2}]) \cong P' \times [\frac{13}{4}, \frac{7}{2}]$ by attaching handles of dimension at least $n-3$. Therefore the inclusion $W \hookrightarrow P$ is 3-connected. Moreover the inclusion $P' \hookrightarrow W$ is $(n-4)$ -connected. Since $n \geq 7$, it follows that $P' \hookrightarrow P$ is 3-connected. So the first two claims of the lemma follow. The last claim follows from Lemma 3.3 because $TP' \oplus \mathbb{R} \cong TP|_{P'}$. $\square$

5 Another example

In this section we give an example which shows that there is also a gap in the proof of Lemma 2 in [9] which persists even after the correction of Amann and Dessai.

This lemma is the analog of Lemma 9.3 in [4]. Its purpose is to guarantee that the second claim of Lemma 3.2 holds for π_2 - and π_4 -finite closed oriented S^1 -manifolds M . In the proof of that lemma the finiteness assumption on the homotopy groups is only used to show that the parity of the sum of weights of the isotropy representation at any S^1 -fixed point in M is independent of the fixed point. Then the lemma is concluded from this condition on the weights.

However, the condition on the weights is not good enough to conclude the lemma as the following example shows. Let $\mathbb{C}P_{\pm}^2$ be $\mathbb{C}P^2$ equipped with the following S^1 -action

$$S^1 \times \mathbb{C}P_{\pm}^2 \rightarrow \mathbb{C}P_{\pm}^2, \quad (g, [z_0 : z_1 : z_2]) \mapsto [z_0 : g^1 z_1 : g^{\pm 1} z_2].$$

Then one can form the equivariant connected sum $\mathbb{C}P_-^2 \# \mathbb{C}P_+^2$ at the fixed points $[1 : 0 : 0] \in \mathbb{C}P_{\pm}^2$. The resulting S^1 -manifold M is diffeomorphic to $\mathbb{C}P^2 \# \mathbb{C}P^2$. Moreover, the sum of weights at each S^1 -fixed point in M is odd and all H -fixed point sets M^H for $H \subset S^1$ are orientable. However, Lemma 2 of [9] does not hold for the $\mathbb{Z}_2$ -fixed point component on which S^1 acts nontrivially. And, indeed, $\hat{A}(M) = -\frac{1}{4} \neq 0$.

We think the problem in the proof of Lemma 2 of [9] is the last equation on page 350. In this equation the constant c is not integral. It is just a rational number with denominator equal to the order of the principal isotropy group of the S^1 -action on the fixed point component in question. Therefore one cannot say anything about the parity of the left hand side of the equation from the parity of the sum on the right hand side unless one only has odd order finite isotropy groups.

Under the assumption that all finite isotropy groups of points in M have odd order, the proofs of both Lemma 1 and Lemma 2 of [9] work so that one can correct the main result of that paper as follows:

Theorem 5.1 *Let M be a connected, oriented, π_2 - and π_4 -finite closed manifold of even dimension such that S^1 acts on M without finite isotropy groups of even order. Then the equivariant elliptic genus of M is rigid. Moreover, the $\hat{A}$ -genus of M vanishes if the action is nontrivial.*

Under the assumption that the S^1 -action on M is semifree, ie all orbits are either free orbits or S^1 -fixed points, we get the following stronger result.

Theorem 5.2 *Let M be a connected, oriented, closed manifold of even dimension such that S^1 acts semifreely on M . Assume that one of the following two conditions holds:*

- *M is π_2 - and π_4 -finite.*
- *The universal covering of M is spin.*

Then a multiple of M is equivariantly bordant to a spin manifold with semifree S^1 -action.

Proof By [5, Theorem VI], π_2 - and π_4 -finiteness imply that the S^1 -action is either even or odd, ie the codimensions of all fixed point components are either congruent to zero modulo four or to two modulo four, respectively. By the ($k=1$)-case of the proof of Lemma 9.3 of [4] and Lemma 3.3, the same conclusion holds if the universal covering of M is spin.

By cutting out an open tubular neighborhood of M^{S^1} from M , one sees that the normal unit sphere bundle to M^{S^1} bounds an oriented free S^1 -manifold. Since the action on M is of even or odd type, this means that the disjoint union of the normal unit sphere bundles to the fixed point components of codimension divisible by four (not divisible by four, respectively) bounds an oriented free S^1 -manifold. So the claim follows from Theorem 2.4 of [3]. $\square$

We conjecture that under the assumptions of Theorem 3.1 or 5.1, a multiple of M is also equivariantly bordant to a spin manifold with circle action. We note in this context that the vanishing of the $\hat{A}$ -genus implies that a multiple of M is nonequivariantly bordant to a spin manifold with nontrivial circle action (see [2, Section 2.3]).

References

- [1] **M Amann, A Dessai**, *The $\hat{A}$ -genus of S^1 -manifolds with finite second homotopy group*, C. R. Math. Acad. Sci. Paris 348 (2010) 283–285 [MR](#) [Zbl](#) Addendum at [arXiv:2302.00051](#) (2023)
- [2] **M Atiyah, F Hirzebruch**, *Spin-manifolds and group actions*, from “Essays on topology and related topics”, Springer (1970) 18–28 [MR](#) [Zbl](#)
- [3] **L D Borsari**, *Bordism of semifree circle actions on Spin manifolds*, Trans. Amer. Math. Soc. 301 (1987) 479–487 [MR](#) [Zbl](#)
- [4] **R Bott, C Taubes**, *On the rigidity theorems of Witten*, J. Amer. Math. Soc. 2 (1989) 137–186 [MR](#) [Zbl](#)
- [5] **G E Bredon**, *Representations at fixed points of smooth actions of compact groups*, Ann. of Math. 89 (1969) 515–532 [MR](#) [Zbl](#)
- [6] **A Dessai**, *Rigidity theorems for $\text{Spin}^{\mathbb{C}}$ -manifolds*, Topology 39 (2000) 239–258 [MR](#) [Zbl](#)
- [7] **J Ebert**, *If the universal cover of a manifold is spin, must it admit a finite cover which is spin?*, MathOverflow answer (2020) Available at <https://mathoverflow.net/a/371214>
- [8] **A L Edmonds**, *Orientability of fixed point sets*, Proc. Amer. Math. Soc. 82 (1981) 120–124 [MR](#) [Zbl](#)
- [9] **H Herrera, R Herrera**, *$\hat{A}$ -genus on non-spin manifolds with S^1 actions and the classification of positive quaternion-Kähler 12-manifolds*, J. Differential Geom. 61 (2002) 341–364 [MR](#) [Zbl](#) Correction in 90 (2012) 521
- [10] **F Hirzebruch, T Berger, R Jung**, *Manifolds and modular forms*, Aspects of Math. E20, Vieweg & Sohn, Braunschweig, Germany (1992) [MR](#) [Zbl](#)
- [11] **F Hirzebruch, P Slodowy**, *Elliptic genera, involutions, and homogeneous spin manifolds*, Geom. Dedicata 35 (1990) 309–343 [MR](#) [Zbl](#)
- [12] **P S Landweber** (editor), *Elliptic curves and modular forms in algebraic topology*, Lecture Notes in Math. 1326, Springer (1988) [MR](#) [Zbl](#)
- [13] **H B Lawson, Jr, M-L Michelsohn**, *Spin geometry*, Princeton Math. Ser. 38, Princeton Univ. Press (1989) [MR](#) [Zbl](#)
- [14] **K Liu**, *On modular invariance and rigidity theorems*, J. Differential Geom. 41 (1995) 343–396 [MR](#) [Zbl](#)
- [15] **C H Taubes**, *S^1 actions and elliptic genera*, Comm. Math. Phys. 122 (1989) 455–526 [MR](#) [Zbl](#)
- [16] **M Wiemeler**, *S^1 -equivariant bordism, invariant metrics of positive scalar curvature, and rigidity of elliptic genera*, J. Topol. Anal. 12 (2020) 1103–1156 [MR](#) [Zbl](#)
- [17] **E Witten**, *Elliptic genera and quantum field theory*, Comm. Math. Phys. 109 (1987) 525–536 [MR](#) [Zbl](#)

Mathematisches Institut, Universität Münster
Münster, Germany

wiemelerm@uni-muenster.de

Received: 20 January 2023 Revised: 18 January 2024

Shrinking without doing much at all

MICHAEL FREEDMAN

MICHAEL STARBIRD

In 1952 Bing astonished the mathematical world with his wild involution on S^3 . It has been among the most seminal examples in topology. The example depends on finding shrinking homeomorphisms of Bing's decomposition of S^3 into points and arcs. If Bing's original homeomorphisms are varied, Bing's original wild involution changes by conjugation, which preserves some analytic properties while altering others. In 1988, Bing published a second paper, *Shrinking without lengthening*, answering a question that one of the present authors posed to him in an effort to understand the geometry of the entire conjugacy class. Here we produce a counterintuitive construction, namely a method to shrink the Bing decomposition doing almost nothing at all: neither lengthening much nor rotating much.

[57K30](#); [57M60](#)

1 Introduction

In 1952, Bing [3] used decomposition space theory (DST) to produce a shrink of what is now called the Bing decomposition as the key step in building a *wild* involution of the three-sphere S^3 . The involution is wild in that it cannot be made *smooth* in any system of coordinates. This single example invigorated several decades of research in DST, beginning with a rich theory in three dimensions, where manifold factors were first discovered by Andrews and Rubin [2]. Then the subject jumped into high dimensions, where J Cannon [6] and R Edwards [9] showed that the double suspensions of homology spheres are homeomorphic to a standard sphere, a project that culminated in fundamental results of F Quinn and R Edwards on manifold recognition (see Quinn [12] and Davis [7] for an introduction). The Bing decomposition and its close relatives were also fundamental to Freedman's proof [10] of the four-dimensional Poincaré conjecture. A Dranishnikov and collaborators constructed remarkable dimension-raising quotient maps; see Dranishnikov and Schchepin [8]. Combining Bing-style DST and Quinn-style surgery, J Bryant, S Ferry, W Mio and S Weinberger constructed the modern theory of ANR homology manifolds [5].

This paper is part of a reconsideration of DST with analytical aspects in mind. Our earlier paper [11] answered a long-standing question about the analytical properties of the Bing involution. In no coordinate system can it ever be made Lipschitz or even quasiconformal. In fact, any topological conjugate of the Bing involution was shown to have, up to a poly-log factor, an exponential modulus of continuity. The

estimate for this intrinsic modulus of continuity (imoc) requires thinking not just about *one* shrink of the Bing decomposition $\mathcal{D}$, but *all* possible shrinks (since the different shrinks can be thought of as conjugates of any single one).

Studying, in this sense, *all* shrinks of $\mathcal{D}$, we found a “surprise shrink”, the subject of this paper. We present it for two reasons: First, because Bing’s decomposition $\mathcal{D}$ is the ur-example of DST. Second, it is the hope of the authors that the method presented here might be combined with [11] to strengthen the main result of that paper and remove the annoying poly-log factor. At first this hope seems odd since this paper provides a novel *shrinking* method and [11] is, in a sense, a *nonshrinking* result: $\mathcal{D}$ can only be shrunk by doing great violence to its $\mathbb{Z}_2$ -reflection symmetry. But, in the proof of [11], the poly-log originates from the possibility that the imagined adversarial shrinker at some point starts to “delay” by making only tiny motions. Our work here gives some insight into what classes of such “tiny motions” indeed result in shrinking — and, when they do, how to quantify the violations of symmetry.

What is the purpose of joining DST to analysis? Our answer is 4-manifolds. By a historical accident the topological theory of 4-manifolds arose simultaneously with Donaldson’s theory of smooth 4-manifolds. Donaldson theory immediately implied that the infinite constructions of DST could not generally yield smoothable results. In a sense, the topologists were given an easy way out: a crisp no-go theorem. In 1982, there was no appetite to dig into shrinking arguments and determine exactly where, and how much, regularity was lost. With [11], the time seems ripe to study the regularity of DST constructions.

2 The Bing decomposition and his two shrinks

Bing’s decomposition is made by intersecting finite stages called “Bing rings”, nested solid tori (see Figure 2). The shrink amounts to figuring out a strategy for stretching, twisting, bending and/or rotating each finer pair of *daughter* solid tori within the previous *mother* stage. Bing found that the rings need be lengthened only infinitesimally during the shrink, whence his title. We find that not only can the lengths of the solid tori be nearly preserved, but the “rotations” of the daughters within the mother can also be made arbitrarily small and can be made to decay towards zero.

The Alexander horned sphere is the first known wild embedding of S^2 in S^3 . One method of creating that wild embedding of S^2 is as follows: Start with a standard S^2 in S^3 . For ease of visualization, think of S^2 as the yz plane in $\mathbb{R}^3$ with a point at infinity that makes $\mathbb{R}^3$ into S^3 and makes the yz plane a standard embedding of S^2 in S^3 . On the positive x side of this S^2 , construct a specific Cantor set’s worth of arcs — one end of each arc will be on S^2 , while the other ends of those arcs will be rather entangled among each other. Specifically, the Cantor set’s worth of arcs are created by taking the components of an infinite intersection of families $\{C_i\}_{i=0}^\infty$ of U-shaped solid cylinders. Each family C_i has 2^i components. C_0 consists of a single U-shaped cylinder as pictured in Figure 1. Every component of C_i contains two U-shaped cylinders of C_{i+1} embedded as shown in Figure 1. With care, each component of $\bigcap_{i=0}^\infty C_i$ is an arc meeting S^2 in a single point, and the totality of those arc endpoints is a Cantor set on S^2 .

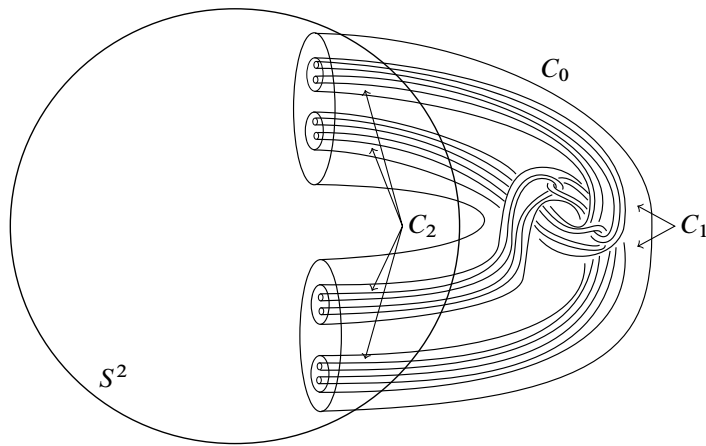


Figure 1: C_2 has four components, continuing the pattern as shown.

Shrinking that Cantor set's worth of arcs is key to creating the Alexander horned sphere embedding of S^2 . The set consisting of those arcs together with the remaining points of S^3 is an upper-semicontinuous decomposition $\mathcal{D}_A$ of S^3 .

The Alexander horned sphere can be created as the image of S^2 under a surjective map $g': S^3 \rightarrow S^3$ whose point preimages are exactly the elements of $\mathcal{D}_A$. The function g' can easily be constructed as the limit of homeomorphisms $\{g'_i: S^3 \rightarrow S^3\}_{i=0}^\infty$ that increasingly compress the arcs of $\mathcal{D}_A$ toward their endpoints off S^2 . Describing this process in terms of quotient spaces, the function g' shows that $S^3/\mathcal{D}_A \cong S^3$, so the 3-sphere has not been changed, but the 2-sphere $g'(S^2)$ is the *wild* Alexander horned sphere. Its wildness is reflected in the fact that the right side component of its complement is no longer simply connected but instead has an infinitely generated fundamental group. The simple closed curve around the center of the cylinder C_0 is an example of a nontrivial element of the fundamental group of the wild component of $S^3 - g'(S^2)$. The image under g' of the arcs in $\mathcal{D}_A$ is a Cantor set that lies on the Alexander horned sphere.

Consider the two *closed* complementary regions of $S^3 - g'(S^2)$. The closed left side region is clearly homeomorphic to the ball B^3 , whereas the closed right side region, the Alexander horned ball, is clearly not, having a nonsimply connected interior.

Now create a new decomposition $\mathcal{D}$ of S^3 by creating nontrivial elements that are symmetric across S^2 . Let I_0 be the standard involution of S^3 with fixed-point set S^2 ; that is, $I_0(x, y, z) = (-x, y, z)$. The nondegenerate elements of $\mathcal{D}$ are the components of the infinite intersection $\bigcap_{i=0}^\infty \tilde{\mathcal{T}}_i$, where $\tilde{\mathcal{T}}_i = C_i \cup I_0(C_i)$. So each $\tilde{\mathcal{T}}_i$ is the union of 2^i tori that each meet S^2 in two meridional disks. We will denote each torus component of $\tilde{\mathcal{T}}_i$ by $\tilde{T}_\sigma$, where σ is a binary string of length i and $\tilde{T}_{\sigma 0}$ and $\tilde{T}_{\sigma 1}$ are the two component tori of $\tilde{\mathcal{T}}_{i+1}$ contained in $\tilde{T}_\sigma$. Figure 2 shows $\tilde{\mathcal{T}}_0$, $\tilde{\mathcal{T}}_1$ and $\tilde{\mathcal{T}}_2$. Notice that $\bigcap_{i=0}^\infty \tilde{\mathcal{T}}_i$ is a Cantor set's worth of arcs, each one piercing S^2 at its center point. Those arcs comprise the nondegenerate elements of $\mathcal{D}$.

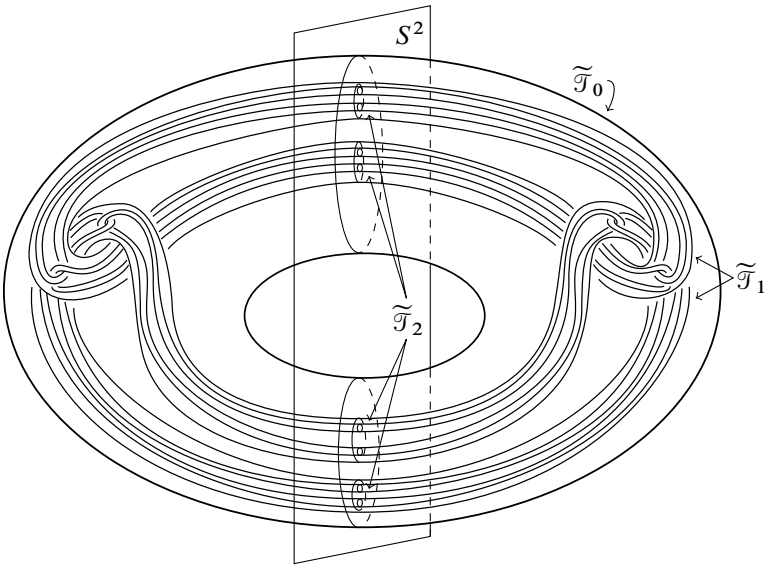


Figure 2

In his 1952 paper, Bing showed that there is a surjective map $g : S^3 \rightarrow S^3$ whose point preimages are precisely the sets in $\mathcal{D}$, thereby producing an involution $I : S^3 \rightarrow S^3$ defined by $I(x) = gI_0g^{-1}(x)$. The involution I is wild since its fixed-point set is the wild 2-sphere $g(S^2)$, and the closure of each component of the complement of $g(S^2)$ is an Alexander horned ball. The involution I swaps sides across a wild sphere.

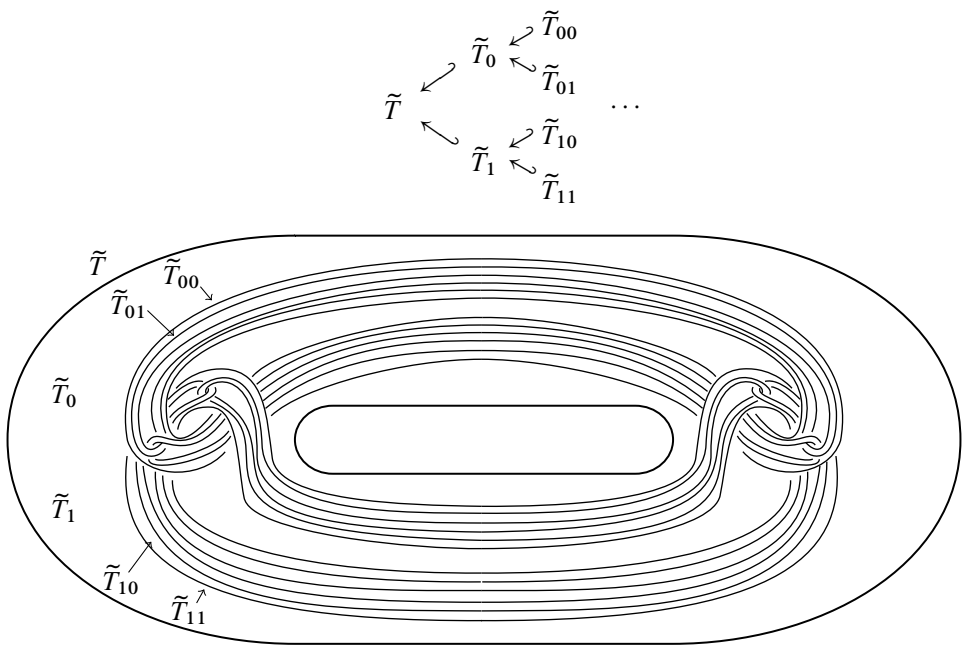


Figure 3

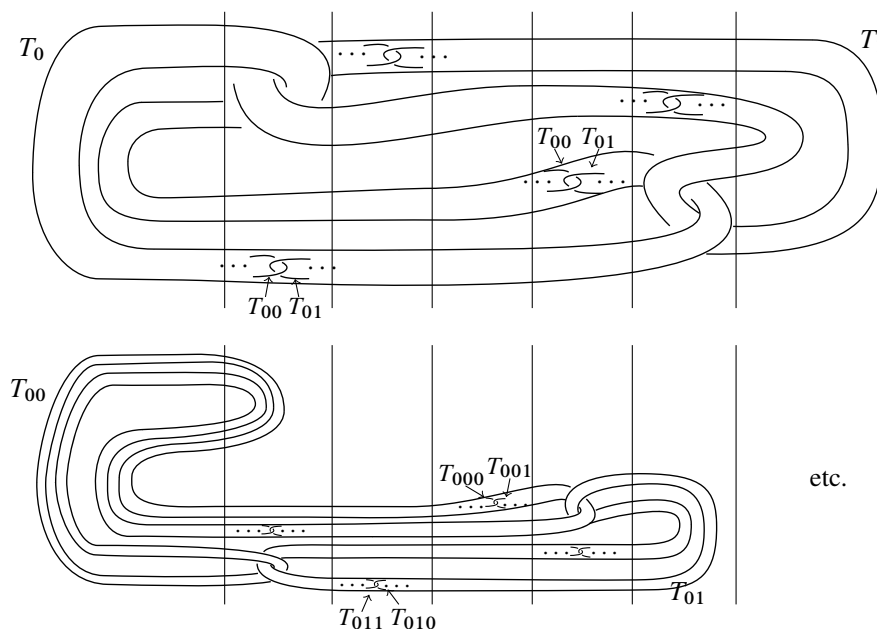


Figure 4

Bing's map $g: S^3 \rightarrow S^3$ is produced as the limit of shrinking homeomorphisms $\{g_i: S^3 \rightarrow S^3\}_{i=0}^\infty$. The g_i homeomorphisms eventually shrink the nondegenerate elements of $\mathcal{D}$ to increasingly smaller diameters as i increases in such a way that the limit g has the property that each nondegenerate element of $\mathcal{D}$ shrinks to a point and $\{g^{-1}(x) \mid x \in S^3\} = \mathcal{D}$. So the challenge is to produce the homeomorphisms g_i that shrink the arcs of $\mathcal{D}$.

Bing's shrinks The astonishing conclusion of [3] is that such shrinking homeomorphisms exist. Let us review two shrinks of $\mathcal{D}$ that Bing published in [3; 4]. In both shrinks, Bing reduces the shrinks to an essentially 1D model where the only important measure of diameter is displacement along the x -axis. Denoting the “Bing tori” dyadically, he lays them out along the x -axis and measures their diameters discretely by choosing a large integer n and erecting parallel planes in intervals of $\frac{1}{n}$. So the original tori are positioned as in Figure 3.

Bing's original shrink [3] was accomplished by describing a sequence of homeomorphisms $\{g_i\}_{i=0}^\infty$. Each homeomorphism g_{i+1} agrees with g_i on $S^3 - \tilde{\mathcal{T}}_i$. The homeomorphisms g_i are defined in sets, meaning we first define the first n_1 g_i 's and then sort of pause while we celebrate a certain amount of shrinking; for example, we could choose our first collection of homeomorphisms so that, for every component torus $\tilde{T}_\sigma$ of $\tilde{\mathcal{T}}_{n_1}$, $\text{diam}(g_{n_1}(\tilde{T}_\sigma))$ meets some diameter goal, say $< \frac{2}{10^1}$. Then we start with a new diameter goal — say $< \frac{2}{10^2}$ — and create the next set of g_i 's, say $i = n_1 + 1, \dots, n_2$. It is only towards the end of each set of homeomorphisms that the diameters shrink. In other words, when we look at the diameters of the components of $\{g_j(\tilde{\mathcal{T}}_j)\}_{j=n_1+1}^{n_2-1}$, generally those images of tori do not have increasingly smaller diameters; however, finally, at stage n_2 , for each torus $\tilde{T}_\sigma \subset \tilde{\mathcal{T}}_{n_2}$, $\text{diam}(g_{n_2}(\tilde{T}_\sigma)) < \frac{2}{10^2}$.

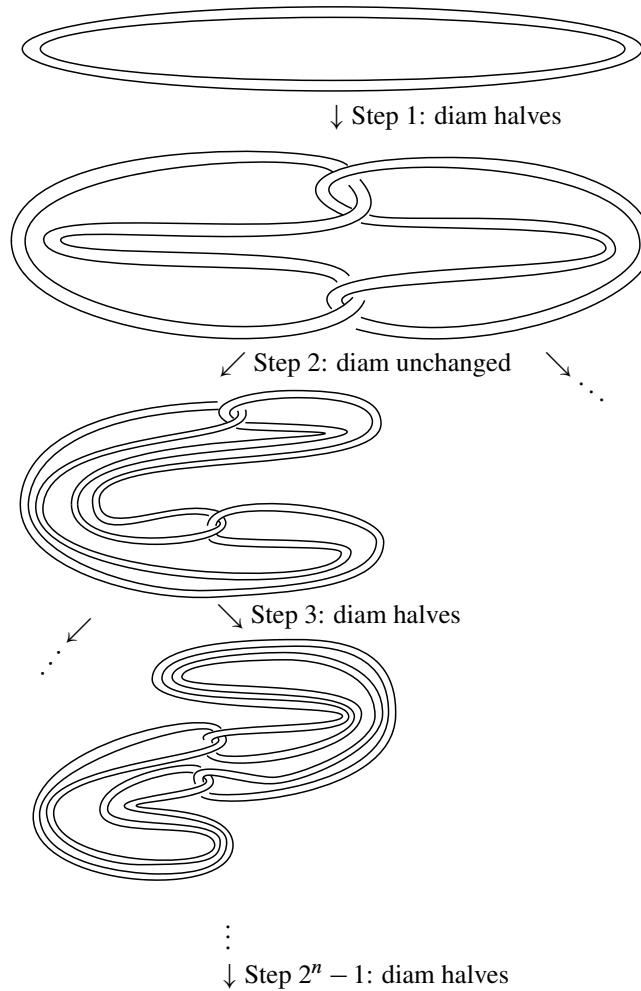


Figure 5

The homeomorphisms $\{g_i\}_{i=1}^{n_1}$ are defined as follows: Let's choose our first goal to be to shrink each torus at some stage to diameter less than, say, $\frac{2}{10^l}$. Imagine the first stage torus as long, thin and flat. Then position sufficiently many parallel planes — say n_1 planes — so that the distance between consecutive parallel planes is less than $\frac{1}{10^l}$ and the torus $\tilde{T}_0$ intersects each of the n_1 planes in a pair of meridional disks. The first homeomorphism g_1 leaves $S^3 - \tilde{\mathcal{T}}_0$ fixed and rotates $\tilde{T}_0 \cup \tilde{T}_1$ in $\tilde{\mathcal{T}}_0$ so that each intersects one less plane (see Figure 4). For simplicity, for each component torus $\tilde{T}_\sigma \subset \tilde{\mathcal{T}}_i$, we will denote $g_i(\tilde{T}_\sigma)$ by T_σ . So after the rotation and using our new notation, T_0 and T_1 each intersects only $n_1 - 1$ planes. In general, each subsequent pair of daughter tori are rotated in their already moved mother so that each of the k -stage daughters meet only $n_1 - k$ planes (Figure 4).

After n_1 generations, no T_σ , where σ is a binary word of length $|\sigma| = n_1$, meets more than one of the planes, so its x -axis extent is $< \frac{2}{10^l}$. Normal to the x -axis, we are free to have chosen a strong

compression, so this procedure produces a homeomorphism g_{n_1} that shrinks each n_1 -stage torus to diameter less than $\frac{2}{10}$. Shrinking the tori of course shrinks the decomposition elements therein.

Next we choose a new diameter goal, say, $\frac{2}{10^2}$. Make many tick marks (actually, meridional disks on parallel planes) along the partially shrunk T_σ 's for $|\sigma| = n_1$ such that the distance between consecutive (around T_σ) meridional disks chosen is less than $\frac{1}{10^2}$. Now start our process over. That is, let $g_{n_1} + 1$ rotate the two daughters in each T_σ in such a way that those daughters each intersect one less meridional disk. Continue defining the g_i 's, each reducing the number of meridional disks intersected by each stage torus, until we reach a number n_2 such that every T_σ , where $|\sigma| = n_2$, meets at most one meridional disk. Again, by compressing dimensions other than the x -extent means every such T_σ has diameter less than $\frac{2}{10^2}$, as desired. Notice that, because of the folded nature of the T_σ 's for $|\sigma| = n_1$, the images under the g_i 's starting with $i = n_1 + 1$ do not decrease the diameters of the T_σ 's for a long time, but, when we reach n_2 , we can again pause to celebrate successful shrinking.

So, after sufficient celebration, we start again and repeat the process with an even more ambitiously small diameter goal. Continue producing such g_i 's. In the limit, the g_i 's converge to a surjective function $g: S^3 \rightarrow S^3$ whose nondegenerate point preimages are precisely the nondegenerate elements of $\mathcal{D}$. This then was Bing's original method of shrinking the decomposition $\mathcal{D}$.

Next we summarize Bing's 1988 shrink, which he produced in answer to questions we asked him at that time. In his 1988 shrink, every other rotation of tori is *greedy*, as it tries (usually in vain) to cut diameters in half by rotating the daughters maximally, that is, rotating the clasp points of the daughters to the midpoints of their mother. The alternate rotations are *patient* — just rotating by $\frac{1}{2^n}$ when going from diameter $\frac{1}{2^{n-1}}$ to diameter $\frac{1}{2^n}$. It turns out greed does not speed the shrinking; it is only at steps indexed by $2^n - 1$ that diameters *actually* are halved. So, again, many steps are taken during which no diameter shrinking is accomplished. Figure 5 shows the idea of Bing's 1988 shrink [4].

3 A small displacement Bing shrink

As described in the introduction, every known shrink of the Bing decomposition consists in starting with a standard torus and describing at each stage how to displace the clasp points of the two daughters relative to the two clasp points of the mother. For known shrinks, the clasps can be imagined as arbitrarily tight, and the reduction of diameter can be studied in a strictly 1D model where the starting torus is configured to tightly surround the unit interval. In this way, the diameter of a torus at a future stage is measured by the x -axis width of the part of $[0, 1]$ covered by that folded solid torus.

Describing the positions of the folded tori at each stage can therefore be captured as a binary tree of functions into $[0, 1]$ as follows. We start with an example.

Example The first, straight torus $T_\emptyset$ is modeled by the identity function $f_\emptyset: [0, 1] \rightarrow [0, 1]$. In the example in Figure 6, the clasp positions of the daughters T_0 and T_1 are displaced by distance $\frac{1}{3}$ at the left

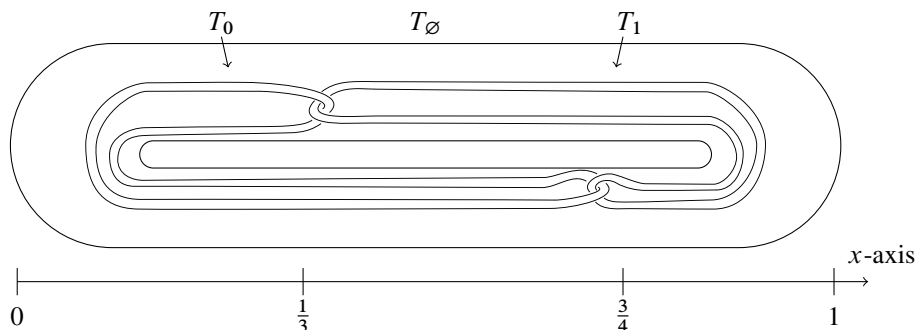


Figure 6

and distance $\frac{1}{4}$ at the right end. So the folded length of T_0 is $\frac{1}{3} + \frac{3}{4} = \frac{13}{12}$ and the folded length of T_1 is $(1 - \frac{1}{3}) + (1 - \frac{3}{4}) = \frac{11}{12}$.

The following two piecewise linear functions model the configurations of T_0 and T_1 . The function $f_0: [-\frac{1}{3}, \frac{3}{4}] \rightarrow [0, 1]$ takes the interval $[-\frac{1}{3}, 0]$ backwards from $\frac{1}{3}$ to 0 and then proceeds forwards from 0 to $\frac{3}{4}$. Specifically,

$$f_0(x) := \begin{cases} f_\emptyset(-x) & \text{if } -\frac{1}{3} \leq x \leq 0, \\ f_\emptyset(x) & \text{if } 0 < x \leq \frac{3}{4}. \end{cases}$$

Thus, the image of f_0 models the shape of T_0 projected into the interval $[0, 1]$.

Likewise, $f_1: [\frac{1}{3}, \frac{5}{4}] \rightarrow [0, 1]$ is defined to model the shape of T_1 . Specifically,

$$f_1(x) := \begin{cases} f_\emptyset(x) & \text{if } \frac{1}{3} \leq x \leq 1, \\ f_\emptyset(1 - (x - 1)) = f_\emptyset(2 - x) & \text{if } 1 < x \leq \frac{5}{4}. \end{cases}$$

Notice that the lengths of the domains of f_0 and f_1 , respectively, roughly equal the lengths of the folded tori T_0 and T_1 , although, since T_0 and T_1 are clapsed, they must be slightly longer than the domains of f_0 and f_1 .

Following the pattern of this example, we next describe how to inductively produce a binary tree of functions $\{f_\sigma\}$ indexed by finite binary strings σ into $[0, 1]$ whose images trace the patterns of the folded tori created by displacing the clasping positions of daughter pairs relative to clasping positions of their mothers when shrinking the Bing decomposition.

Definition 3.1 We define $f_\emptyset$ as the identity function on $[0, 1]$ and define $[c_\emptyset, d_\emptyset] := [0, 1]$. Suppose σ is a finite binary string and the function $f_\sigma: [c_\sigma, d_\sigma] \rightarrow [0, 1]$ has been produced. The string σ has two daughters, $\sigma 0$ and $\sigma 1$, and f_σ will have two daughters $f_{\sigma 0}$ and $f_{\sigma 1}$. The definitions of the daughter functions of f_σ are determined by the choice of a pair of points a_σ, b_σ in $[c_\sigma, d_\sigma]$ with $c_\sigma \leq a_\sigma \leq b_\sigma \leq d_\sigma$. We call a_σ and b_σ the daughter *clasp points*. As in the example, the domain of $f_{\sigma 0}$ extends the domain of f_σ to the left so that $f_{\sigma 0}$ begins by retracing f_σ backwards from a_σ to c_σ , then forwards from c_σ to b_σ . So the domain of $f_{\sigma 0}$ is $[c_\sigma - (a_\sigma - c_\sigma), b_\sigma] = [2c_\sigma - a_\sigma, b_\sigma]$. Likewise, $f_{\sigma 1}$ retraces f_σ forward from a_σ to d_σ and then backwards from d_σ to b_σ (see Figure 7).

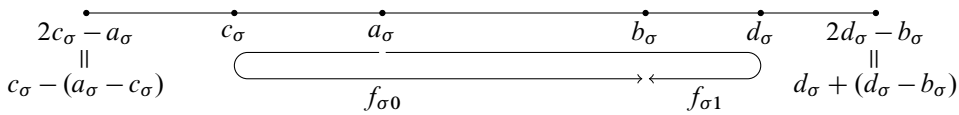


Figure 7

Here are the formal definitions of $f_{\sigma 0}$, $f_{\sigma 1}$, $c_{\sigma 0}$, $d_{\sigma 0}$, $c_{\sigma 1}$ and $d_{\sigma 1}$:

$$\begin{aligned}
 f_{\sigma 0}: [2c_{\sigma} - a_{\sigma}, b_{\sigma}] &\rightarrow [0, 1], & f_{\sigma 0}(x) &:= \begin{cases} f_{\sigma}(-x + 2c_{\sigma}) & \text{if } 2c_{\sigma} - a_{\sigma} \leq x \leq c_{\sigma}, \\ f_{\sigma}(x) & \text{if } c_{\sigma} \leq x \leq b_{\sigma}, \end{cases} \\
 [c_{\sigma 0}, d_{\sigma 0}] &:= [2c_{\sigma} - a_{\sigma}, b_{\sigma}], \\
 f_{\sigma 1}: [a_{\sigma}, 2d_{\sigma} - b_{\sigma}] &\rightarrow [0, 1], & f_{\sigma 1}(x) &:= \begin{cases} f_{\sigma}(x) & \text{if } a_{\sigma} \leq x \leq d_{\sigma}, \\ f_{\sigma}(-x + 2d_{\sigma}) & \text{if } d_{\sigma} \leq x \leq 2d_{\sigma} - b_{\sigma}, \end{cases} \\
 [c_{\sigma 1}, d_{\sigma 1}] &:= [a_{\sigma}, 2d_{\sigma} - b_{\sigma}].
 \end{aligned}$$

□

Notice that each f_{σ} defined above is piecewise linear with each piece having slope ± 1 . There are many interesting choices for the tree of clasp points $\{a_{\sigma}, b_{\sigma}\}$ and, hence, the tree of functions f_{σ} associated with potential shrinks of Bing's decomposition.

Definition 3.2 (1) A binary tree of functions $\{f_{\sigma}\}$ starting with $f_{\emptyset}$ being the identity on $[0, 1]$ and defined as above will be called a *Bing tree of functions*.

(2) A Bing tree of functions $\{f_{\sigma}\}$ *shrinks* if and only if $\text{length}(f_{\sigma}([c_{\sigma}, d_{\sigma}])) \rightarrow 0$ wherever the string length $|\sigma| \rightarrow \infty$.

(3) The values $|a_{\sigma} - c_{\sigma}|$ and $|d_{\sigma} - b_{\sigma}|$ are *displacements*.

□

A Bing tree of functions that shrinks indicates a method of shrinking the Bing decomposition; however, such a tree of functions is just a model because the clasping positions of daughter tori cannot occur at single points. So the values a_{σ} and b_{σ} just indicate where the clasps occur, although in reality both $T_{\sigma 0}$ and $T_{\sigma 1}$ will be slightly longer than the domains $[c_{\sigma 0}, d_{\sigma 0}]$ and $[c_{\sigma 1}, d_{\sigma 1}]$ of $f_{\sigma 0}$ and $f_{\sigma 1}$, respectively. Nevertheless, a Bing tree of functions that shrinks can guide a shrink of the Bing decomposition simply by making the clasps sufficiently tight.

Previously known shrinks of the Bing decomposition included some large displacements, that is, instances where $|a_{\sigma} - c_{\sigma}|$ and $|d_{\sigma} - b_{\sigma}|$ were relatively large compared to $|d_{\sigma} - c_{\sigma}|$. However, we show in this paper that it is possible to construct a Bing shrink—or, actually, a Bing tree of functions $\{f_{\sigma}\}$ that shrinks—even though all the displacements are small. The shrinks we will produce have the additional property that the length of each torus, or equivalently the domain of each f_{σ} , grows by less than any desired quantity.

Before constructing our small displacement shrink, let's make some observations about the functions f_{σ} in a Bing tree of functions. When visualizing the following, it might be useful to imagine the displacements—that is, the $|a_{\sigma} - c_{\sigma}|$'s and $|d_{\sigma} - b_{\sigma}|$'s—as very small compared to the length of $|d_{\sigma} - c_{\sigma}|$.

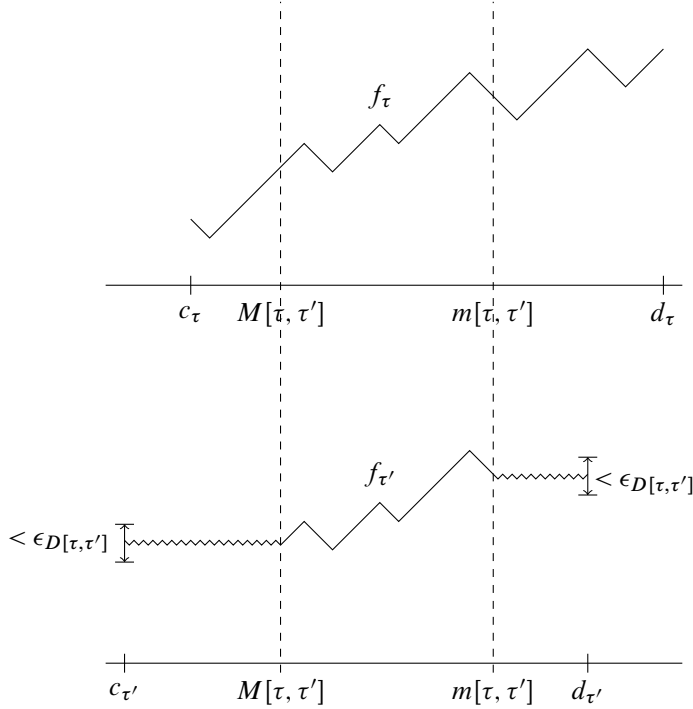


Figure 8: Note that $c_{\tau'}$ may be greater than, or less than, or equal to c_τ , and similarly for $d_{\tau'}$. The salient observation is that f_τ and $f_{\tau'}$ are identical on the interval $[M[\tau, \tau'], m[\tau, \tau']]$.

Consider the daughter $f_{\sigma 0}$ of f_σ . The domain of f_σ , namely $[c_\sigma, d_\sigma]$, shifts downward to create the domain of $f_{\sigma 0}$, namely $[c_\sigma - (a_\sigma - c_\sigma), b_\sigma] = [2c_\sigma - a_\sigma, b_\sigma]$. Let's think about the relationship between the images of f_σ and $f_{\sigma 0}$.

First notice that the domains of f_σ and $f_{\sigma 0}$ share the interval $[c_\sigma, b_\sigma]$, and by definition agree there. The function $f_{\sigma 0}$ is defined on the additional interval $[2c_\sigma - a_\sigma, c_\sigma]$, but $f_{\sigma 0}([2c_\sigma - a_\sigma, c_\sigma]) = f_\sigma([c_\sigma, a_\sigma])$ as sets, so no new points are added to the image of $f_{\sigma 0}$ compared to the image of f_σ ; that is, $f_{\sigma 0}([c_{\sigma 0}, d_{\sigma 0}]) \subset f_\sigma([c_\sigma, d_\sigma])$. Similarly, the domain of $f_{\sigma 1}$ shifts up compared to the domain of f_σ , namely $[c_{\sigma 1}, d_{\sigma 1}] = [a_\sigma, d_\sigma + (d_\sigma - b_\sigma)] = [a_\sigma, 2d_\sigma - b_\sigma]$. These observations allow us to record some relationships between an f_τ and its descendants in the lemma and corollary below.

Definition 3.3 Let τ' be a binary string descendant of τ (that is, τ is the beginning of the string τ') and let $\{f_\sigma\}$ be a Bing tree of functions. Then $M[\tau, \tau'] := \max\{a_\mu \mid \tau \leq \mu \leq \tau'\}$, $m[\tau, \tau'] := \min\{b_\mu \mid \tau \leq \mu \leq \tau'\}$, and $\epsilon_{D[\tau, \tau']} := \max\{|a_\mu - c_\mu|, |d_\mu - b_\mu| \mid \tau \leq \mu \leq \tau'\}$. $\square$

Lemma 3.4 Let $\{f_\sigma\}$ be a Bing tree of functions and let τ' be a descendant of the binary string τ . Then:

- (1) $f_{\tau'}$ restricted to $[M[\tau, \tau'], m[\tau, \tau']]$ equals f_τ restricted to $[M[\tau, \tau'], m[\tau, \tau']]$.
- (2) $\text{length}(f_{\tau'}([c_{\tau'}, M[\tau, \tau']])) \leq \epsilon_{D[\tau, \tau']}$.
- (3) $\text{length}(f_{\tau'}[m[\tau, \tau'], d_{\tau'}]) \leq \epsilon_{D[\tau, \tau']}$.

Corollary 3.5 Let $\{f_\sigma\}$ be a Bing tree of functions and let τ' be a descendant of the binary string τ . Then

$$\text{length}(f_{\tau'}([c_{\tau'}, d_{\tau'}])) \leq |m[\tau, \tau'] - M[\tau, \tau']| + 2\epsilon_{D[\tau, \tau']}$$

We are now ready to construct a shrink that uses only small displacements.

Theorem 3.6 For any $\epsilon_L > 0$ and $\{\epsilon_i \geq 0 \mid \sum_{i=0}^{\infty} \epsilon_i^2 = \infty\}$, there exists a Bing tree of functions $\{f_\sigma\}$ that shrinks such that, for every σ ,

- (1) $|d_\sigma - c_\sigma| < 1 + \epsilon_L$, and
- (2) for $|\sigma| = i$, $|a_\sigma - c_\sigma| \leq \epsilon_i$ and $|d_\sigma - b_\sigma| \leq \epsilon_i$.

Proof We will describe the functions f_σ by describing the domain intervals $[c_\sigma, d_\sigma]$. Notice that knowing the domain intervals of f_σ and of $f_{\sigma 0}$ automatically implies what a_σ and b_σ are and what $f_{\sigma 1}$ is. Specifically, if the domain of f_σ is $[c_\sigma, d_\sigma]$ and the domain of $f_{\sigma 0}$ is $[c_{\sigma 0}, d_{\sigma 0}]$, then $a_\sigma = c_\sigma + (c_\sigma - c_{\sigma 0}) = 2c_\sigma - c_{\sigma 0}$ and $b_\sigma = d_{\sigma 0}$. So the domain and definition of $f_{\sigma 1}$ are also determined.

It will be convenient to associate each domain interval $[c_\sigma, d_\sigma]$ with the point in the plane (c_σ, d_σ) . Notice the following:

- (1) Since during our construction c_σ will always be less than d_σ , each point (c_σ, d_σ) lies above the main diagonal Δ ; specifically, Δ is the graph of $y = x$.
- (2) The horizontal (or vertical) distance from Δ to (c_σ, d_σ) equals the length of $[c_\sigma, d_\sigma]$.
- (3) Interval inclusion $[u, v] \subset [c_\sigma, d_\sigma]$ corresponds to the point (c_σ, d_σ) lying in the NW quadrant with respect to the point (u, v) .
- (4) The vector from the point in the plane (c_σ, d_σ) to $(c_{\sigma 0}, d_{\sigma 0})$, namely $(c_\sigma - a_\sigma, b_\sigma - d_\sigma)$, is the negative of the vector from (c_σ, d_σ) to $(c_{\sigma 1}, d_{\sigma 1})$.

These observations allow us to construct the desired Bing tree of functions $\{f_\sigma\}$ that shrinks by describing an associated binary tree of points in the plane (c_σ, d_σ) with $\text{length}([c_\sigma, d_\sigma]) < 1 + \epsilon_L$ and $\text{length}(f_\sigma([c_\sigma, d_\sigma]))$ eventually getting short.

Definition 3.7 Let τ be a finite binary string. Then a finite set of finite binary strings $\{\tau_i\}_{i=1, \dots, n}$ is a *complete* set of descendants of τ if and only if

- (1) for each i , τ_i is a descendant of τ , and
- (2) for every infinite binary string $\hat{\tau}$ that starts with τ , there exists exactly one τ_i such that $\hat{\tau}$ starts with τ_i .

Similarly, we will refer to finite, complete sets of descendants of f_τ and (c_τ, d_τ) if their subscripts are a finite, complete set of descendants of τ . □

The proof of the following lemma contains the heart of the construction.

Lemma 3.8 *Let $\epsilon_L > 0$, $\{\epsilon_i \geq 0 \mid \sum_{i=0}^{\infty} \epsilon_i^2 = \infty\}$, $\delta > 0$, and τ' be a finite binary string descendant of τ . Suppose the beginning of a Bing tree of functions has been defined starting at $f_{\emptyset}$ and including all descendants of $f_{\emptyset}$ that are ancestors of $f_{\tau'}$. Suppose:*

- (1) $|d_{\tau'} - c_{\tau'}| < 1 + \epsilon_L$.
- (2) $\text{length}(f_{\tau'}([c_{\tau'}, M[\tau, \tau']])) < \delta$.
- (3) $\text{length}(f_{\tau'}([m[\tau, \tau'], d_{\tau'}])) < \delta$.
- (4) $|m[\tau, \tau'] - M[\tau, \tau']| \geq 3\delta$.

Then there exists a finite complete set A of descendants of τ' such that, for each $f_{\tau''}$ in A ,

- (1) $\max\{|a_{\mu} - c_{\mu}|, |d_{\mu} - b_{\mu}|\} \leq \min\{\epsilon_{|\mu|}, \delta\}$ for each μ with $\tau' \leq \mu \leq \tau''$, and
- (2) $|m[\tau, \tau''] - M[\tau, \tau'']| \leq \frac{2}{3}|m[\tau, \tau'] - M[\tau, \tau']|$.

Corollary 3.9 *Given the hypotheses of Lemma 3.8, there exists a finite complete set A of descendants of τ' such that, for each $f_{\tau''} \in A$,*

$$\text{length}(f_{\tau''}([c_{\tau''}, d_{\tau''}])) \leq \frac{2}{3}|m[\tau, \tau'] - M[\tau, \tau']| + 2\delta.$$

Proof of Lemma 3.8 We will describe descendants of $f_{\tau'}$ by looking at points in the plane. The interval $[M[\tau, \tau'], m[\tau, \tau']]$ is contained in $[c_{\tau'}, d_{\tau'}]$. And, of course, the middle third of $[M[\tau, \tau'], m[\tau, \tau']]$, namely $I := [M[\tau, \tau'] + \frac{1}{3}(m[\tau, \tau'] - M[\tau, \tau']), M[\tau, \tau'] + \frac{2}{3}(m[\tau, \tau'] - M[\tau, \tau'])]$, is also contained in $[c_{\tau'}, d_{\tau'}]$. So the point $(c_{\tau'}, d_{\tau'})$ lies in the NW quadrant above I . Call that NW quadrant Q . Figure 9 shows that relationship.

Figure 9 also suggests a large number of concentric circles centered at some distant point p in that NW quadrant Q on the slope -1 ray from I heading up and left. Recall that $|d_{\tau'} - c_{\tau'}| < 1 + \epsilon_L$. Let $C_{\tau'}$ be the circle centered at p and containing the point $(c_{\tau'}, d_{\tau'})$. The point p is so distant that every point $(x, y) \in C_{\tau'} \cap Q$ has $y - x < 1 + \epsilon_L$. This choice of circle center will guarantee that no point (c_{μ}, d_{μ}) that is created during our process has corresponding interval $[c_{\mu}, d_{\mu}]$ with length greater than $1 + \epsilon_L$.

Now we are ready to construct our binary tree of points (c_{μ}, d_{μ}) . We begin at $(c_{\tau'}, d_{\tau'})$ and proceed along the tangent of the circle $C_{\tau'}$ in both directions until we hit the next circle. We choose the size of the next circle so that the distance along the tangent of $C_{\tau'}$ to the next concentric circle equals $\min\{\epsilon_{|\tau'|}, \delta\}$. Those two points will be $(c_{\tau'0}, d_{\tau'0})$ and $(c_{\tau'1}, d_{\tau'1})$. From each of those points we do the same thing — that is, from $(c_{\tau'0}, d_{\tau'0})$ we move along the tangent of its circle $C_{\tau'0}$ to find the points $(c_{\tau'00}, d_{\tau'00})$ and $(c_{\tau'01}, d_{\tau'01})$ on the next circle $C_{\tau'00}$. At each step, when going from (c_{μ}, d_{μ}) to points $(c_{\mu0}, d_{\mu0})$ and $(c_{\mu1}, d_{\mu1})$, we choose the size of circle $C_{\mu0}$ so that the distance along the tangent of circle C_{μ} to a point on $C_{\mu0}$ equals $\min\{\epsilon_{|\mu|}, \delta\}$.

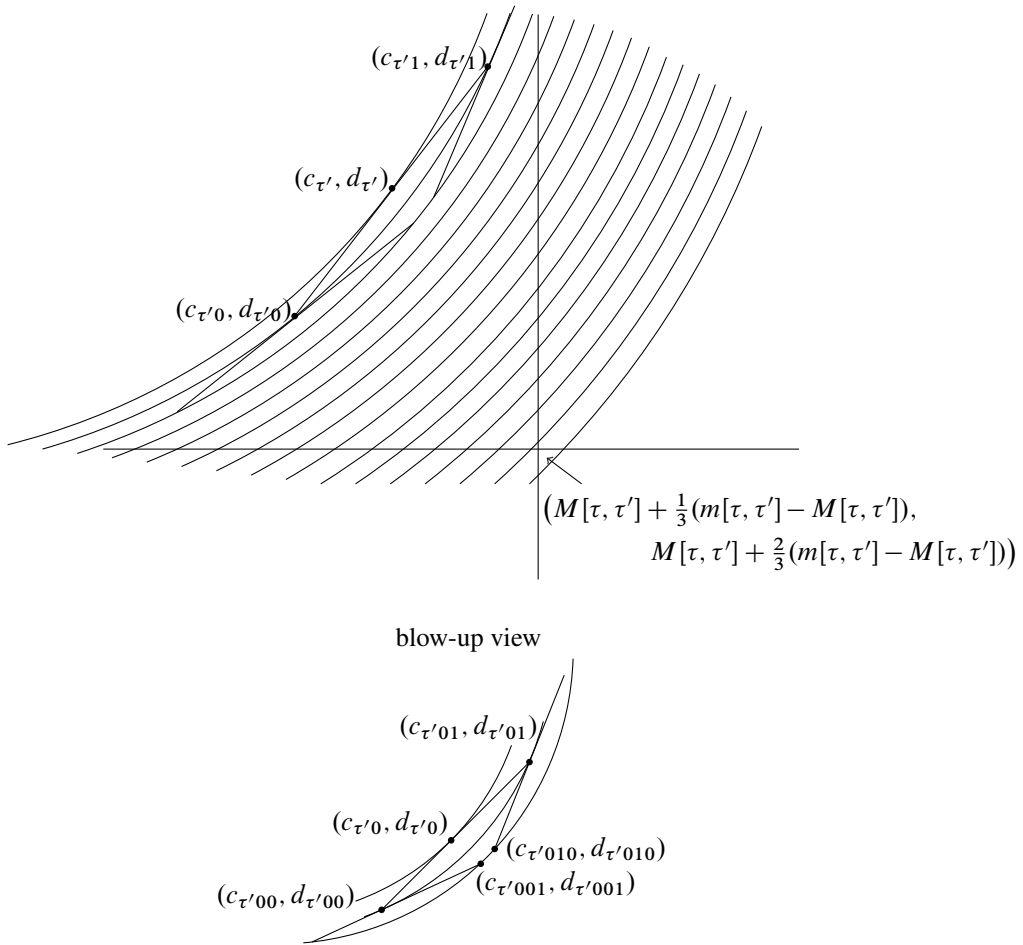


Figure 9

We continue creating this binary tree of points until we arrive at a point $(c_{\tau''}, d_{\tau''})$ where either

- (1) $c_{\tau''} > M[\tau, \tau'] + \frac{1}{3}(m[\tau, \tau'] - M[\tau, \tau'])$, or
- (2) $d_{\tau''} < M[\tau, \tau'] + \frac{2}{3}(m[\tau, \tau'] - M[\tau, \tau'])$,

that is, $(c_{\tau''}, d_{\tau''}) \notin Q$.

The following lemma implies that every branch of descendants does eventually leave Q .

Lemma 3.10 Let $\{\epsilon_i \geq 0 \mid \sum_{i=0}^{\infty} \epsilon_i^2 = \infty\}$, $\delta > 0$, and let $\{C_i\}_{i=1}^{\infty}$ be a nested sequence of concentric circles in $\mathbb{R}^2$ centered at point p such that, for every i , the radius of C_i is r_i and the distance from a point on C_i along the tangent to C_i to a point on C_{i+1} is $\min\{\epsilon_i, \delta\}$. Then the sequence of radii $(r_i)_{i=1}^{\infty}$ is unbounded.

Proof Let $M \in \mathbb{R}^+$. Given the hypotheses, we will show that there is a k such that $r_k \geq M$. If not, then, for every i , $r_i < M$ (see Figure 10).

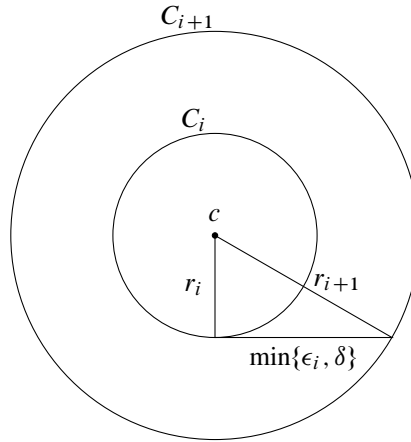


Figure 10

By the Pythagorean theorem,

$$r_{i+1}^2 - r_i^2 = \epsilon_i^2, \text{ so } r_{i+1} - r_i = \frac{\min\{\epsilon_i, \delta\}^2}{r_{i+1} + r_i} > \frac{\min\{\epsilon_i, \delta\}^2}{2M}.$$

But $\sum_{i=0}^{\infty} \min\{\epsilon_i, \delta\}^2 / 2M$ diverges, meaning some $r_k > M$. □

Every path branch of the binary tree of points that we construct stops by a predictable stage since our hypothesis about the divergence of the sum of the squares of the ϵ_i 's guarantees that there are only a finite number of circles before the point $(M[\tau, \tau'] + \frac{1}{3}(m[\tau, \tau'] - M[\tau, \tau']), M[\tau, \tau'] + \frac{2}{3}(m[\tau, \tau'] - M[\tau, \tau']))$ is itself inside a circle C_μ . Any point constructed with our procedure on such a circle C_μ must be outside Q . When we have arrived at a point $(c_{\tau''}, d_{\tau''})$ that is not in Q , let us define $a_{\tau''} := c_{\tau''}$ and $b_{\tau''} := d_{\tau''}$. The function $f_{\tau''}$ defined using these values $c_{\tau''}$, $a_{\tau''}$, $b_{\tau''}$ and $d_{\tau''}$ then satisfies the conditions for being an element of the finite, complete set of descendants of $f_{\tau'}$ that we seek, thereby concluding the proof of [Lemma 3.8](#). □

Note Our proof of [Lemma 3.8](#) has actually proved a stronger result. Since we always stop at the first time that a $(c_{\tau''}, d_{\tau''})$ is not in Q —that is, either $c_{\tau''} > M[\tau, \tau'] + \frac{1}{3}(m[\tau, \tau'] - M[\tau, \tau'])$ or $d_{\tau''} < M[\tau, \tau'] + \frac{2}{3}(m[\tau, \tau'] - M[\tau, \tau'])$ —then, in the former case, $c_{\tau''} = M[\tau, \tau'']$, while, in the latter case, $d_{\tau''} = m[\tau, \tau'']$. So [Corollary 3.9](#) could actually be improved by replacing the 2δ with δ . However, this subtlety is not necessary when applying [Lemma 3.8](#) and [Corollary 3.9](#) to prove [Theorem 3.6](#).

We can now conclude the proof of how to construct the desired Bing tree of functions that shrinks as follows. We begin with $f_\emptyset: [0, 1] = [c_\emptyset, d_\emptyset] \rightarrow [0, 1]$ being the identity and $a_\emptyset := 0$ and $b_\emptyset := 1$.

We apply [Lemma 3.8](#), starting with $\tau = \emptyset$ and $\tau' = \emptyset$, and with $\delta = \frac{1}{10}$ repeatedly until we have created a finite, complete set of descendants $A_\emptyset$ where the small displacement conditions are satisfied at each stage and such that, for each $\tau \in A_\emptyset$, $f_\tau: [c_\tau, d_\tau] \rightarrow [0, 1]$ has the property that $\text{length}(f_\tau([c_\tau, d_\tau])) < 3\delta + 2\delta = \frac{5}{10} = \frac{1}{2}$.

We then start with each such f_τ and choose $a_\tau := c_\tau$ and $b_\tau := d_\tau$, and use it and itself as its descendants, with $\delta = \frac{1}{10^2}$ as the hypotheses of [Lemma 3.8](#). We use [Lemma 3.8](#) repeatedly to create a finite, complete set A_τ of descendants of f_τ such that the construction respects the small displacement and nonlengthening conditions and, for each $\sigma \in A_\tau$, $\text{length}(f_\sigma([c_\sigma, d_\sigma])) < \frac{3}{10^2} + \frac{2}{10^2} = \frac{5}{10^2}$.

Continuing in this fashion produces the desired Bing tree of functions, that shrinks while respecting the small displacement conditions, thus proving [Theorem 3.6](#). $\square$

4 Conclusion and questions

Bing's 1952 construction of a wild involution of S^3 opened the door to many further insights — and many further questions. In [\[11\]](#), we proved that all conjugates of Bing's involution *must* share certain analytic features with the involution derived from Bing's original shrink. Understanding such shared features turned up a surprise, the subject of this paper, whose exact relation to the previous paper is still to be worked out. Bing always told us not to form fixed beliefs about what you do not know. Following his advice, let us state some questions without presuming to guess their answers.

Question The first question involves the rate at which small displacements can shrink. Bing's original method of shrinking and his shrinking without lengthening method shrink stage tori to size about $\frac{1}{n}$ their original diameter at stage order n . The analysis of [\[1\]](#) shows that there is no faster way to shrink the Bing decomposition $\mathcal{D}$. It appears that if one imposes constraints on displacement and lengthening, as we have here, shrinking must be even slower. For example, if lengthening and displacement are restricted to 0.1%, our algorithm takes about 10^{12} stages to get from diameter = 1 to diameter = 0.001. What is the actual functional form for shrinking using our algorithm? And are there more efficient algorithms respecting the same constraints? In both cases, what are the analytical properties, the modulus of continuity (moc), of the corresponding involutions?

Question Suppose we use the 1D model for shrinking as in this paper. We showed that it is possible to shrink the Bing decomposition with small displacement shrinks. Bing's shrinks and the shrink in this paper seem to require some insight or even cleverness, but might that apparent cleverness be an illusion? Suppose the a_σ 's and b_σ 's were simply chosen randomly in the intervals $[c_\sigma, d_\sigma]$. Would such a random selection lead to a shrink of the Bing decomposition with probability 1? If so, wouldn't we feel silly. Shrinking (or not) is a *tail event*, meaning independent of any initial segment of choices, so Kolmogorov's 0-1 law tells us that, within a probabilistic model, shrinking will occur with probability either 0 or 1, which depends on the model. If the model is artificially concentrated near our explicit shrink, the probability will be 1, but if $\{(a_\sigma, b_\sigma)\}$ are independent and uniformly distributed, we do not know.

We were slow to accept that Bing's decomposition could be shrunk using only tiny jiggles. Bing understood that the unknown is actually *unknown*. He told us that he would alternately work toward

proving a conjecture was true and toward producing a counterexample with the idea that both attempts would help to expose the truth, whatever that truth turns out to be.

Acknowledgement

We would like to thank the referee for many helpful suggestions that have greatly improved the exposition of the construction.

References

- [1] **FD Ancel, MP Starbird**, *The shrinkability of Bing–Whitehead decompositions*, *Topology* 28 (1989) 291–304 [MR](#)
- [2] **J J Andrews, L Rubin**, *Some spaces whose product with E^1 is E^4* , *Bull. Amer. Math. Soc.* 71 (1965) 675–677 [MR](#)
- [3] **R H Bing**, *A homeomorphism between the 3-sphere and the sum of two solid horned spheres*, *Ann. of Math.* 56 (1952) 354–362 [MR](#)
- [4] **R H Bing**, *Shrinking without lengthening*, *Topology* 27 (1988) 487–493 [MR](#)
- [5] **J Bryant, S Ferry, W Mio, S Weinberger**, *Topology of homology manifolds*, *Ann. of Math.* 143 (1996) 435–467 [MR](#)
- [6] **J W Cannon**, $\Sigma^2 H^3 = S^5/G$, *Rocky Mountain J. Math.* 8 (1978) 527–532 [MR](#)
- [7] **R J Daverman**, *Decompositions of manifolds*, *Pure and Applied Math.* 124, Academic, Orlando, FL (1986) [MR](#) [Zbl](#)
- [8] **A N Dranishnikov, E V Shchepin**, *Cell-like mappings: the problem of the increase of dimension*, *Uspekhi Mat. Nauk* 41 (1986) 49–90 [MR](#) In Russian; translated in *Russ. Math. Surv.* 41 (1986) 59–111
- [9] **R D Edwards**, *The topology of manifolds and cell-like maps*, from “Proceedings of the International Congress of Mathematicians” (O Lehto, editor), *Acad. Sci. Fennica, Helsinki* (1980) 111–127 [MR](#)
- [10] **M H Freedman**, *The topology of four-dimensional manifolds*, *J. Differential Geometry* 17 (1982) 357–453 [MR](#)
- [11] **M Freedman, M Starbird**, *The geometry of the Bing involution*, preprint (2022) [arXiv 2209.07597](#)
- [12] **F Quinn**, *Ends of maps, II*, *Invent. Math.* 68 (1982) 353–424 [MR](#)

MF: *Station Q, Microsoft*
Santa Barbara, CA, United States

Department of Mathematics, University of California, Santa Barbara
Santa Barbara, CA, United States

MS: *Department of Mathematics, The University of Texas at Austin*
Austin, TX, United States

mikehartleyfreedman@outlook.com, starbird@math.utexas.edu

Received: 24 January 2023 Revised: 4 October 2023

Action of the Mazur pattern up to topological concordance

ALEX MANCHESTER

In the 1980s, Freedman showed that the Whitehead doubling operator acts trivially up to topological concordance. On the other hand, Akbulut showed that the Whitehead doubling operator acts nontrivially up to smooth concordance. The Mazur pattern is a natural candidate for a satellite operator which acts by the identity up to topological concordance but not up to smooth concordance. Recently there has been a resurgence of study of the action of the Mazur pattern up to concordance in the smooth and topological categories. Examples showing that the Mazur pattern does not act by the identity up to smooth concordance have been given by Cochran, Franklin, Hedden and Horn and by Collins. We give evidence that the Mazur pattern acts by the identity up to topological concordance.

In particular, we show that two satellite operators P_{K_0, η_0} and P_{K_1, η_1} with η_0 and η_1 freely homotopic have the same action on the topological concordance group modulo the subgroup of (1)-solvable knots, which gives evidence that they act in the same way up to topological concordance. In particular, the Mazur pattern and the identity operator are related in this way, and so this is evidence for the topological side of the analogy to the Whitehead doubling operator. We give additional evidence that they have the same action on the full topological concordance group by showing that, up to topological concordance, they cannot be distinguished by Casson–Gordon invariants or metabelian ρ -invariants.

57K10; 57N70

1 Introduction

The Whitehead doubling operator, shown in [Figure 1](#), is a fundamental example of a satellite operator. The strands going through the box should be tied into the knot J . A general satellite operator acts on a knot J by grabbing a collection of strands and tying them into J . We will give a precise definition in [Section 2](#).

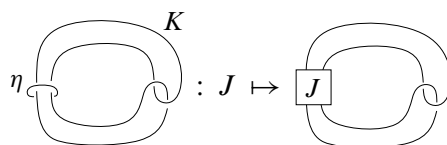


Figure 1: The Whitehead doubling operator.

Another important satellite operator is the *Mazur pattern*, which is shown in [Figure 2](#), left, along with the identity operator on the right, which takes every knot to itself. These two operators are closely related in a way which we will make precise in [Definition 1.2](#); for now we will simply note that

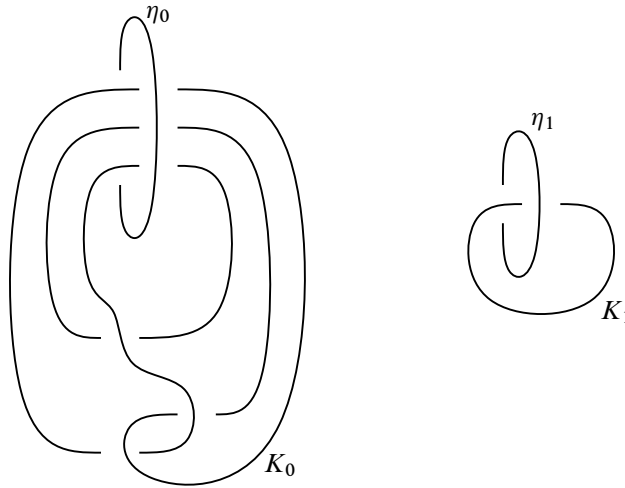


Figure 2: The Mazur pattern and the identity operator are homotopically related.

$\text{lk}(K_0, \eta_0) = \text{lk}(K_1, \eta_1)$ and that both operators take the unknot to itself. The main theorem of this paper, [Theorem 1.4](#), essentially says that satellite operators which are related in a certain way induce the same function on knots up to some equivalence relation.

All knots are assumed to be oriented, and, given a knot K , the knot $-K$ is the mirror image of K with the reverse orientation. Unless otherwise specified, manifolds and submanifolds (in particular, slice disks and concordances) are smooth, though we will often use the words “smooth” and “smoothly” for emphasis.

Satellite operators give well-defined maps on the concordance groups $\mathcal{C}_{\text{top}}$ and $\mathcal{C}_{\text{sm}}$ which are of significant interest. For example, their kernels and images have been studied, and of particular note is the result of Cochran, Davis and Ray that satellite operators with strong winding number ± 1 are injective up to topological concordance (see [\[3\]](#), especially Theorem 5.1, as well as Hedden and Pinzón-Caicedo [\[23\]](#) and Levine [\[24\]](#)). Also, satellite operators have been used to give evidence for a “fractal nature” of the concordance groups (see Cochran and Harvey [\[6\]](#), Cochran, Harvey and Leidy [\[10\]](#), Cochran, Harvey and Powell [\[11\]](#) and Ray [\[33\]](#)). Furthermore, they are typically not homomorphisms, and it is conjectured that only three quite trivial satellite operators are homomorphisms (see Lidman, Miller and Pinzón-Caicedo [\[25\]](#) and Miller [\[30\]](#)). Instead, satellite operators can be interpreted as group actions (see Davis and Ray [\[16\]](#)). Satellite operators have also been used to give evidence for and against the slice–ribbon conjecture (see Gompf and Miyazaki [\[21\]](#), Miller and Piccirillo [\[31\]](#) and Yasui [\[35\]](#)).

The following theorem is a special case of [\[5, Theorem 1.5\]](#) of Cochran, Friedl and Teichner, and is the motivation for this paper, particularly the main result, [Theorem 1.4](#):

Theorem 1.1 (see [\[5, Theorem 1.5\]](#)) *If K is topologically (or smoothly) slice and η is nullhomotopic in some topological slice disk complement for K , then $P(J)$ is topologically slice for any J (in other words, $P: \mathcal{C}_{\text{top}} \rightarrow \mathcal{C}_{\text{top}}$ is the zero map).*

Note that the Whitehead doubling operator shown in [Figure 1](#) satisfies the condition of [Theorem 1.1](#) (since K is the unknot and so has the trivial slice disk, whose complement has fundamental group $\mathbb{Z}$, and η has linking number 0 with K and so must be nullhomotopic). So, this theorem gives a proof that all Whitehead doubles are topologically slice. The fact that all Whitehead doubles are topologically slice was first proved by Freedman [17], and [Theorem 1.1](#) is a large generalization of Freedman's result. It is also worth noting that all Whitehead doubles have Alexander polynomial 1, and Freedman and Quinn [18, Theorem 11.7B] showed that all knots with Alexander polynomial 1 are topologically slice.

We think of the following definition as being a relative version of the condition in [Theorem 1.1](#):

Definition 1.2 Two satellite operators P_0 and P_1 defined by the links $K_0 \cup \eta_0$ and $K_1 \cup \eta_1$, respectively, are *homotopically related* if K_0 and K_1 are concordant and η_0 and η_1 cobound an immersed annulus in the complement of the concordance (that is, if $\eta_0 \times \{0\}$ and $\eta_1 \times \{1\}$ are freely homotopic in the concordance complement). Note that this definition makes sense in both the smooth and topological settings; however, we will primarily make use of it in the smooth setting.

Satellite operators which satisfy the condition of [Theorem 1.1](#) are exactly those which are topologically homotopically related to the satellite operator given by the unlink $K \sqcup \eta$.

Two (smoothly) homotopically related satellite operators which do not satisfy the condition of [Theorem 1.1](#) are the aforementioned Mazur pattern and identity operator, shown in [Figure 2](#). These two can be seen to be homotopically related since K_0 is the unknot, and so its complement has fundamental group $\mathbb{Z}$, as does the complement of the concordance from the unknot to itself. And so, η_0 is freely homotopic to a meridian since it has linking number 1 with K_0 , which gives the immersed annulus.

[Theorem 1.1](#) motivates the following question, asked by Ray at the 2019 AIM workshop *Smooth concordance classes of topologically slice knots*:

Question 1.3 For homotopically related satellite operators P_0 and P_1 , is $P_0(J)$ topologically concordant to $P_1(J)$ for each J ? In other words, is $P_0(J) \# -P_1(J)$ topologically slice?

Note that if the answer to [Question 1.3](#) is “yes”, this would imply that the actions of satellite operators P which take the unknot to itself (for example the Whitehead doubling operator, the Mazur pattern and the identity operator) are completely determined up to topological concordance by their winding number via an argument similar to the one above.

Here we give some evidence that the answer to this question is “yes”, in particular the following theorem. The condition of (1)-solvability lies between algebraic sliceness and sliceness, and will be defined later.

Theorem 1.4 Given (smoothly) homotopically related satellite operators P_ϵ for $\epsilon \in \{0, 1\}$, for any J the knot $P_0(J) \# -P_1(J)$ is (1)-solvable. In other words, $P_0 \equiv P_1$ on $\mathcal{C}/\mathcal{F}_{(1)}$, where $\mathcal{F}_{(1)}$ is the subgroup of (1)-solvable knots.

In fact, at the end of [Section 3](#), we will show that $P_0(J) \# -P_1(J)$ is *homotopy ribbon (1)-solvable* when $K_0 \# -K_1$ is homotopy ribbon, which is slightly stronger than being (1)-solvable.

In the particular case of the Mazur pattern, denoted by Q , the 0-surgeries M_J and $M_{Q(J)}$ are (smoothly) homology cobordant rel meridian. In fact this is true for any satellite operator with winding number 1 which takes the unknot to itself, as shown by Cochran, Franklin, Hedden and Horn [[4](#), Corollary 2.2]. It is an open question whether any two knots whose 0-surgeries are smoothly or topologically homology cobordant rel meridian are topologically concordant.

In the smooth setting, [Theorem 1.1](#) is false, and the answer to [Question 1.3](#) is “no”. In unpublished work in 1983, Akbulut used gauge theory to show that the Whitehead double of the right-handed trefoil is not smoothly slice. Gauge-theoretic techniques are very sensitive to matters of handedness, and it is still unknown whether the Whitehead double of the left-handed trefoil is smoothly slice. The question of smooth sliceness of the Whitehead double of the left-handed trefoil is one of the most important questions in low-dimensional topology. Furthermore, one can use Heegaard Floer theory to show that, for many knots J (in particular the figure-eight knot), the knots J and $Q(J)$ are not necessarily smoothly concordant (see Cochran, Franklin, Hedden and Horn [[4](#), Theorem 3.1] and Collins [[14](#), Proposition 1.2]).

In [Section 2](#), we will give some basic definitions and background. In [Section 3](#), we will construct a (1)-solution for $P_0(J) \# -P_1(J)$ (which is moreover a homotopy ribbon (1)-solution when $K_0 \# K_1$ is homotopy ribbon). In [Section 4](#), we will discuss some additional conditions which would guarantee that these knots are topologically slice. In [Sections 5](#) and [6](#), we will show that $P_0(J)$ and $P_1(J)$ cannot be distinguished up to topological concordance using certain obstructions coming from Casson–Gordon invariants and metabelian ρ invariants, respectively, which gives further evidence that these two knots are topologically concordant.

Acknowledgements

I would like to thank Shelly Harvey, Miriam Kuzbary and Allison Miller for their many valuable conversations and useful suggestions. I was partially supported by the RTG award NSF DMS-1745670.

2 Background

We begin with a precise definition of satellite operators:

Definition 2.1 A *satellite operator* $P = P_{K,\eta}$ is given by a two-component link $K \cup \eta$, where η is an unknot, and acts on the set of knots in S^3 as follows: Notice that the exterior of η is a solid torus E_η which contains the knot K and which has as a preferred longitude a meridian μ_η of η . The result of applying P to J is the knot given by deleting a tubular neighborhood of $J \subset S^3$ and gluing in E_η in such a way as to identify the meridian of η with the preferred longitude of J and vice versa (which

guarantees that the ambient manifold is still S^3). Then $P(J)$ is the image of K in the resulting S^3 . In other words, a satellite operator acts by gluing together the exterior E_η of η and the exterior of J along their torus boundary via the homeomorphism $T^2 \rightarrow T^2$ which identifies the meridian of η with the preferred longitude of J and vice versa, and, while doing this, keeps track of the knot $K \subset E_\eta$.

Notice that $P(U) = K$ since, when $J = U$, the exterior of J is just a solid torus, which, when glued to E_η simply replaces the solid torus neighborhood of η in the trivial way.

A knot in S^3 is called *smoothly/topologically slice* if it is the boundary of a smooth/locally flat, properly embedded disk in B^4 . Two knots K and J are *smoothly/topologically concordant* if there is a smooth/locally flat, properly embedded annulus in $S^3 \times I$ with boundary $(K \times \{0\}) \sqcup (-J \times \{1\})$. The set of equivalence classes of knots up to smooth/topological concordance is denoted by $\mathcal{C}_{\text{sm}}$ or $\mathcal{C}_{\text{top}}$, respectively.

The sets $\mathcal{C}_{\text{sm}}$ and $\mathcal{C}_{\text{top}}$ both form groups under the connect sum operation $\#$, with $[K]^{-1} = [-K]$. We will abuse notation and write K for its concordance class. The fact that $-K = K^{-1}$ can be seen by taking the pair $(S^3 \times I, K \times I)$ and removing a tubular neighborhood of an arc running from a point on $K \times \{0\}$ to a point on $-K \times \{1\}$ (note that taking S^3 with the standard orientation means that the boundary on the “1” side is $(S^3, -K)$), after which the remainder of $S^3 \times I$ is a 3-ball and the remainder of the annulus $K \times I$ is a slice disk for $K \# -K \subset S^3 \# S^3 = S^3$. (This construction works in either the smooth or the topological setting.)

Note that K and J are (smoothly or topologically) concordant if and only if the connect sum $K \# -J$ is (smoothly or topologically) slice, by an argument similar to that in the previous paragraph showing that $[K]^{-1} = [-K]$.

Satellite operators descend to maps $\mathcal{C}_{\text{sm}} \rightarrow \mathcal{C}_{\text{sm}}$ and $\mathcal{C}_{\text{top}} \rightarrow \mathcal{C}_{\text{top}}$ (but, as noted before, are typically not homomorphisms, and can instead be thought of as group actions).

We will need the following well-known characterization of topological sliceness:

Proposition 2.2 (see eg [5, Proposition 2.1]) *A knot K is topologically slice if and only if its 0-surgery M_K bounds a topological 4-manifold W such that*

- (1) $\pi_1(W)$ is normally generated by the meridian μ of K ;
- (2) $H_1(W) \cong \mathbb{Z}$ (in other words, the inclusion map $H_1(M_K) \rightarrow H_1(W)$ is an isomorphism); and
- (3) $H_2(W) = 0$.

It is often difficult to check whether M_K bounds such a topological 4-manifold, but sometimes it is easier to show that M_K bounds a smooth 4-manifold satisfying the following weaker condition, due to Cochran, Orr and Teichner [12]. Recall that the derived series of a group G is defined inductively as $G^{(0)} := G$ and $G^{(n+1)} := [G^{(n)} : G^{(n)}]$; that is, each term is the commutator subgroup of the previous term.

Definition 2.3 A knot K is called (n) -solvable if M_K bounds a smooth 4-manifold W such that

- (1) the inclusion map $H_1(M_K) \rightarrow H_1(W)$ is an isomorphism;
- (2) $H_2(W)$ is freely generated by embedded surfaces $\{S_i, T_i\}$ with trivial normal bundle such that S_i and T_i intersect once transversely, and otherwise the surfaces are disjoint; and
- (3) the inclusion maps $\pi_1(S_i), \pi_1(T_i) \rightarrow \pi_1(W)$ land in $\pi_1(W)^{(n)}$.

If, additionally, the inclusion maps $\pi_1(S_i) \rightarrow \pi_1(W)$ land in $\pi_1(W)^{(n+1)}$, then K is said to be $(n.5)$ -solvable. If K is (h) -solvable for some $h \in \frac{1}{2}\mathbb{N}$, the manifold W is called an (h) -solution for K .

Remark 2.4 Smooth concordance preserves (h) -solvability since, if K and J are concordant and J is (h) -solvable, we can construct an (h) -solution for K in the following way: Do 0-surgery $\times I$ along a concordance from K to J , and glue an (h) -solution for J to the M_J boundary component. By a Seifert–Van Kampen and Mayer–Vietoris argument, this gives an (h) -solution for K with the surfaces representing generators for the second homology simply the images of the surfaces representing the second homology in the (h) -solution for J .

Also, if K and J are both (h) -solvable, with (h) -solutions W_K and W_J , respectively, one can construct an (h) -solution W for $K \# J$ by gluing together W_K and W_J along the solid tori in their boundaries given by tubular neighborhoods of meridians of K and J , respectively, identifying the meridians and preferred longitudes in the torus boundaries. The boundary of W is then $M_{K \# J}$, which can be seen by viewing $K \# \cdot$ as a satellite operator given by $K \cup \mu_K$: the tubular neighborhoods of meridians of K and J are now buried in the interior of W , and the boundary is given by gluing together the exteriors of μ_K and μ_J (note that μ_J can be freely homotoped to be the core of the surgery solid torus in M_J , and so the exterior of μ_J is just the exterior of J). Of course, we could just as well have interchanged the roles of K and J . Again by Seifert–Van Kampen and Mayer–Vietoris, this W is an (h) -solution for $K \# -J$.

Lastly, if K is (h) -solvable, then so is $-K$, by taking the mirror image of an (h) -solution for K . So, the (h) -solvable filtration is a filtration by subgroups.

In the above remark, we used the fact that the 0-surgery on a connect sum $K \# J$ is homeomorphic to the union of the exteriors of K and J , glued along their torus boundary in such a way as to identify the meridians and longitudes. We will use this fact at a few points throughout this paper.

Definition 2.5 The set of (h) -solvable knots up to smooth concordance is denoted by $\mathcal{F}_{(h)}$.

This defines a filtration of the concordance group $\mathcal{C}_{\text{sm}}$,

$$\cdots \subseteq \mathcal{F}_{(1.5)} \subseteq \mathcal{F}_{(1)} \subseteq \mathcal{F}_{(0.5)} \subseteq \mathcal{F}_{(0)} \subseteq \mathcal{C}_{\text{sm}}.$$

This filtration was first defined by Cochran, Orr and Teichner [12]. Note that, if the surfaces S_i and T_i were spheres, condition (3) of Definition 2.3 would be trivially satisfied for all h . In this case, we could do surgery on one sphere from each pair, replacing a tubular neighborhood $S^2 \times D^2$ with $B^3 \times S^1$ (which

have the same boundary), in order to get a manifold which satisfies the conditions of [Proposition 2.2](#). All topologically slice knots are (h) -solvable for every $h \in \frac{1}{2}\mathbb{N}$, as noted in [\[8, Section 1\]](#), and it is conjectured that $\bigcap_{h \in \frac{1}{2}\mathbb{N}} \mathcal{F}(h)$ contains precisely the topologically slice knots. It is known that, for all $n \in \mathbb{N}$, the quotient $\mathcal{F}(n)/\mathcal{F}(n.5)$ has infinite rank (see [\[8\]](#)). However, it is unknown whether $\mathcal{F}(n.5)/\mathcal{F}(n+1)$ is even nontrivial for any n .

3 Construction of a (1)-solution

In this section we will prove [Theorem 1.4](#) by constructing a (1)-solution for $P_0(J) \# -P_1(J)$.

Fix the following notation: Let P_ϵ be homotopically related satellite operators coming from the links $K_\epsilon \cup \eta_\epsilon$ for $\epsilon \in \{0, 1\}$. Let C be the (smooth) concordance between K_0 and K_1 , and let $W = (S^3 \times I) \setminus \nu(C)$. Note that W is a manifold satisfying the conditions of [Proposition 2.2](#) for $K_0 \# -K_1$. Let J be a knot, and let $P_\epsilon(J)$ be the image of J under the satellite operator P_ϵ . Let μ be a fixed meridian of J .

Let N be the manifold obtained by gluing two copies of $M_J \times I$ to W in the following way (see [Figure 3](#) for a schematic picture): Recall that $M_J = (S^3 \setminus J) \cup_f (S^1 \times D^2)$, where $f: \partial(S^1 \times D^2) \rightarrow \partial(S^3 \setminus J)$ identifies the meridian of $S^1 \times D^2$ with the preferred longitude of J and the longitude of $S^1 \times D^2$ with the meridian of J . Then N is obtained by gluing two copies of $M_J \times I$, which we will denote by $(M_J \times I)_\epsilon$ for $\epsilon \in \{0, 1\}$, to W , by identifying the surgery solid torus $S^1 \times D^2 \subset (M_J \times \{0\})_\epsilon$ with $\nu(\eta_\epsilon)$ in such a way as to identify the longitude of the surgery solid torus (that is, a meridian of J) with the 0-framed longitude of η . This gluing “buries” the two solid tori that get glued together inside the interior of the new 4-manifold. The effect of these identifications on the boundary amounts to gluing the exteriors of J and η together via the homeomorphism $T^2 \rightarrow T^2$ which interchanges meridians and longitudes, so that $\partial N = M_{P_0(J) \# -P_1(J)} \sqcup -M_J \sqcup -M_J$. Denote the two copies of $\mu \times \{1\}$ by μ_ϵ , and let A be the immersed annulus in W joining the η_ϵ , extended through $(M_J \times I)_\epsilon$ by crossing with I , so that its boundary components are the two μ_ϵ . Finally, let $\hat{\Sigma}_\epsilon$ denote Seifert surfaces for J capped off by longitudinal disks in the copies of $(M_J \times \{1\})_\epsilon$.

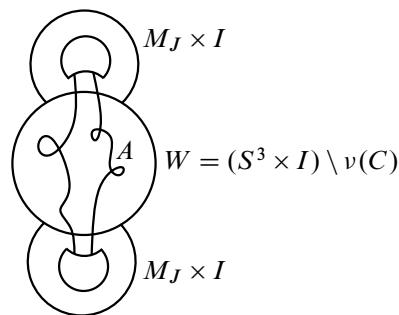


Figure 3: Construction of N : Two copies of $M_J \times I$ are glued to $W = (S^3 \times I) \setminus \nu(C)$ along solid torus portions of their respective boundaries. This “buries” the solid tori in the interior of N , and the new boundary is $M_{P_0(J) \# -P_1(J)} \sqcup -M_J \sqcup -M_J$.

We will now construct a (1)-solution for $P_0(J) \# -P_1(J)$. This will involve cutting out a neighborhood of A , giving a manifold with two boundary components, one of which is $M_{P_0(J) \# -P_1(J)}$. We will then find an appropriate manifold with which to “cap off” the other boundary component, which will give us our (1)-solution.

This will involve a series of lemmas keeping track of the fundamental group and homology of the intermediate steps in the construction. Throughout, we will only care about normal generating sets for fundamental groups and subgroups, and so we will only care about elements of the normal generating sets up to conjugacy, and in turn will only care about their representatives up to free homotopy. Given group elements $g_1, \dots, g_n \in G$, we denote the subgroup they normally generate by $\langle\langle g_1, \dots, g_n \rangle\rangle$.

- Lemma 3.1** (1) $\pi_1(N)$ is normally generated by the meridian of $P_0(J)$ (or equivalently of $P_1(J)$, K_0 or K_1 since these meridians are all freely homotopic inside N).
- (2) $H_1(S^3 \setminus \nu(P_\epsilon(J))) \rightarrow H_1(N)$ is an isomorphism. That is, $H_1(N) \cong \mathbb{Z}$ and is generated by the meridian of $P_0(J)$, or equivalently of $P_1(J)$, K_0 or K_1 .
- (3) $H_2(M_J) \oplus H_2(M_J) \rightarrow H_2(N)$ is an isomorphism. That is, $H_2(N) \cong \mathbb{Z}^2$ and is generated by the $\hat{\Sigma}_\epsilon$.

Proof (1) $\pi_1(W)$ is normally generated by the meridian of K_0 (or equivalently of K_1 since these are freely homotopic and so conjugate in $\pi_1(W)$). In N , a meridian of K_ϵ is freely homotopic to a meridian of $P_\epsilon(J)$. Moreover, $\pi_1((M_J \times I)_\epsilon)$ is normally generated by μ_ϵ , which is freely homotopic to the curve $\mu \times \{0\} \subset (M_J \times I)_\epsilon$ identified with η_ϵ under the gluing used to construct N , and so is in the subgroup normally generated by a meridian of K_0 . Therefore, by Seifert–Van Kampen, $\pi_1(N) = \pi_1(W) *_{\langle\langle \eta_0 \mu_0^{-1} \rangle\rangle} \pi_1(M_J \times I) *_{\langle\langle \eta_1 \mu_1^{-1} \rangle\rangle} \pi_1(M_J \times I)$ is normally generated by a meridian of K_0 , or equivalently of K_1 , $P_0(J)$ or $P_1(J)$ since these are freely homotopic in N .

(2)–(3) If we attach one copy of $M_J \times I$, say the one glued to η_0 , we have the Mayer–Vietoris sequence

$$\begin{array}{ccccccc} 0 & \longrightarrow & 0 \oplus H_2(M_J \times I) & \longrightarrow & H_2(W \cup (M_J \times I)) & \longrightarrow & \\ & & & & \searrow & & \\ & & & & H_1(\nu(\eta_0)) & \longrightarrow & H_1(W) \oplus H_1(M_J \times I) \longrightarrow H_1(W \cup (M_J \times I)) \longrightarrow 0. \end{array}$$

Since η_0 is identified with $\mu \times \{0\} \subset (M_J \times \{0\})_0$, which is freely homotopic to μ_0 , the map $H_1(\nu(\eta_0)) \rightarrow H_1(M_J \times I)$ is an isomorphism. This implies that $H_1(W) \rightarrow H_1(W \cup (M_J \times I))$ and $H_2(M_J \times I) \rightarrow H_2(W \cup (M_J \times I))$ are isomorphisms.¹ Noting that $H_1(S^3 \setminus \nu(P_\epsilon(J))) \cong H_1(S^3 \setminus \nu(K_\epsilon)) \rightarrow H_1(W)$ is an isomorphism, we see that $H_1(S^3 \setminus \nu(P_\epsilon(J))) \rightarrow H_1(W)$ is an isomorphism.

¹To see that the map $H_1(W) \rightarrow H_1(W \cup (M_J \times I))$ is an isomorphism, choose a basis for $H_1(W) \oplus H_1(M_J \times I)$ consisting of the image of a generator of $H_1(\nu(\eta_0))$ (which is $\{\text{some element of } H_1(W)\} \oplus \{\text{a generator of } H_1(M_J \times I)\}$) and $\{\text{a generator of } H_1(W)\} \oplus \{0\}$. This is a basis since $H_1(\nu(\eta_0)) \rightarrow H_1(M_J \times I)$ is an isomorphism. Then one can see that $H_1(W \cup (M_J \times I))$ is generated by the image of a generator of $H_1(W)$. We apply a similar argument later in this proof, and also in the proof of Lemma 3.5.

$$\begin{array}{ccccccc} 0 & \longrightarrow & H_2(W \cup (M_J \times I)) \oplus H_2(M_J \times I) & \longrightarrow & H_2(N) & & \\ & & & & \downarrow & & \\ & & & & H_1(v(\eta_1)) & \longrightarrow & H_1(W \cup (M_J \times I)) \oplus H_1(M_J \times I) \longrightarrow H_1(N) \longrightarrow 0. \end{array}$$

From now on, let λ denote the normal generator of $\pi_1(N)$ from the previous lemma. Recall that μ_ϵ are the meridians of J in $M_J \times \{1\} \subset (M_J \times I)_\epsilon$, and A is the immersed annulus cobounded by the η_ϵ , extended through the $(M_J \times I)_\epsilon$ to the boundary $-M_J$'s via the product structure so that it is cobounded by μ_ϵ .

(3) $H_2(N \setminus \nu(A)) \cong \mathbb{Z} \oplus G$ (where $\hat{\Sigma}_0 - \hat{\Sigma}_1$ generates the $\mathbb{Z}$ summand, and the G summand has one generator for each self-intersection of A , with possibly some relations)

$$\begin{aligned} H_n(N, N \setminus v(A)) &= H_n(vA, \partial' v(A)) && \text{(by excision)} \\ &= H^{4-n}(vA, v(\partial A)) && \text{(by Poincaré–Lefschetz duality)} \\ &= H^{4-n}(A, \partial A) && \text{(by a deformation retract)} \\ &= \begin{cases} \mathbb{Z} & \text{if } n = 2, \\ \mathbb{Z}^{c+1} & \text{if } n = 3, \\ 0 & \text{otherwise.} \end{cases} && \text{(where } c \text{ is the number of self-intersections of } A), \end{aligned}$$

Algebraic & Geometric Topology, Volume 25 (2025)

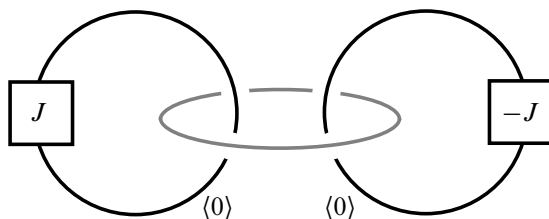


Figure 4: $\partial^-(N \setminus \nu(\text{arc})) = M_J \# M_{-J}$ with the gray curve bounding an immersed disk obtained by cutting A open along the arc.

These homology groups then fit in to the long exact sequence for the pair $(N, N \setminus \nu(A))$,

$$\begin{array}{ccccccc} H_3(N, N \setminus \nu(A)) & \longrightarrow & H_2(N \setminus \nu(A)) & \longrightarrow & H_2(N) & & \\ & & & & \pi & & \\ & \longleftarrow & H_2(N, N \setminus \nu(A)) & \longrightarrow & H_1(N \setminus \nu(A)) & \longrightarrow & H_1(N) \longrightarrow 0. \end{array}$$

Note that the map $\pi: H_2(N) \rightarrow H_2(N, N \setminus \nu(A))$ is given by the algebraic intersection number with A since the generator of $H_2(N, N \setminus \nu(A))$ is the Poincaré–Lefschetz dual of A . Since the boundary components of A are the meridians μ_ϵ of J in $(M_J \times I)_\epsilon$, and the capped-off Seifert surfaces $\hat{\Sigma}_\epsilon$ intersect these meridians once, the images of the classes $\hat{\Sigma}_\epsilon \in H_2(N)$ under π are $\pm 1 \in \mathbb{Z} \cong H_2(N, N \setminus \nu(A))$. This shows that π is surjective and has kernel $\langle \hat{\Sigma}_0 - \hat{\Sigma}_1 \rangle$. This gives (2), and the fact that $H_2(N \setminus \nu(A)) \cong \mathbb{Z} \oplus G$, where G is the image of the map $H_3(N, N \setminus \nu(A)) \rightarrow H_2(N \setminus \nu(A))$.

A priori, G is some quotient of $\mathbb{Z}^{c+1}$. However, recall that the submanifold representing the extra generator described above (the Poincaré–Lefschetz dual of an arc running from one boundary component of A to the other) has boundary a torus which can be taken to be isotopic to the boundary of a tubular neighborhood of either μ_ϵ . This is then nullhomologous in $N \setminus \nu(A)$ since it bounds a copy of $M_J \setminus \nu(\mu_J)$. Therefore, G in fact just has one generator for each self-intersection of A , with possibly some relations. $\square$

Notice that $N \setminus \nu(A)$ has two boundary components, with $\partial^+(N \setminus \nu(A)) = M_{P_0(J) \# -P_1(J)}$ unchanged from $\partial^+ N$. We wish to “plug the hole” by gluing in some 4-manifold with boundary equal to $\partial^-(N \setminus \nu(A))$. To that end, we will find a surgery description of this 3-manifold.

To start, take an arc in N running from μ_0 to μ_1 along A (recall that the μ_ϵ are the copies of the meridian of J living in $M_J \times \{1\} \subset (M_J \times I)_\epsilon$), missing all self-intersections of A . If we cut out a neighborhood of this arc from N , the new boundary is $M_J \# M_{-J}$, and the remainder of A is an immersed disk bounded by the gray curve in Figure 4. When we remove the remainder of A , if there were no self-intersections, this would amount to doing 0-surgery on this gray curve. We can account for the self-intersections as in [19], to get the surgery description shown in Figure 5. Each self-intersection corresponds to one of the Whitehead links which is banded on to the diagram. The sign of the clasp should match the sign of the self-intersection, but won’t affect our discussion so we draw the clasps with indeterminate crossings, following the convention in [19]. A priori, the immersed disk bounded by the gray curve in Figure 4

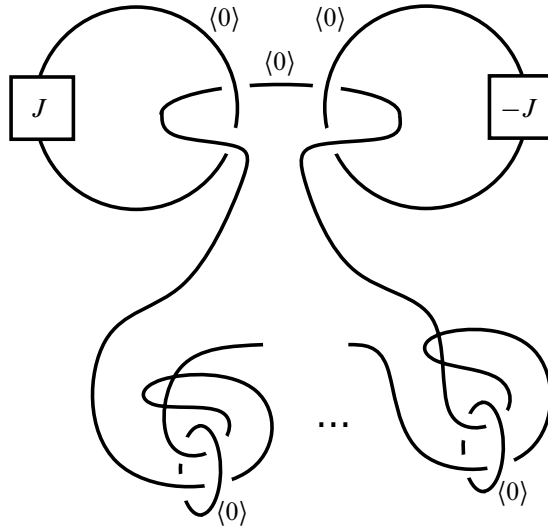


Figure 5: $\partial^-(N \setminus \nu(A))$ accounting for self-intersections of A ; the sign of each clasp matches the sign of the corresponding self-intersection.

might induce a nontrivial framing. However, as A is an immersed annulus in B^4 , the framing difference between the two ends must be twice the algebraic self-intersection number of A , similar to how adding a local kink changes the framing of an immersed disk by ± 2 . We can thus add local kinks to undo this framing difference, which makes the immersed disk bounded by the gray curve in Figure 4 0-framed.

Notice that doing 0-surgery on the gray curve in Figure 4 produces $M_{J\#-J}$ (as in Figure 10). Notice that the 3-manifold doesn't change if we "blow up" the clasps as in Figure 6 (for now ignore the different labels on the components and instead treat everything as 0-framed). Again, the signs of the clasps match

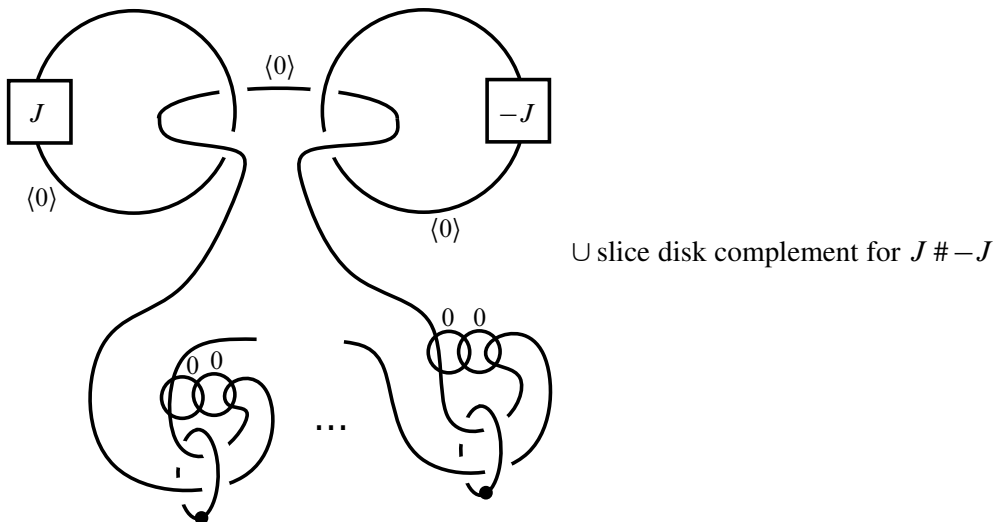


Figure 6: Construction of Z . See [22] for an introduction to the Kirby diagram notation.

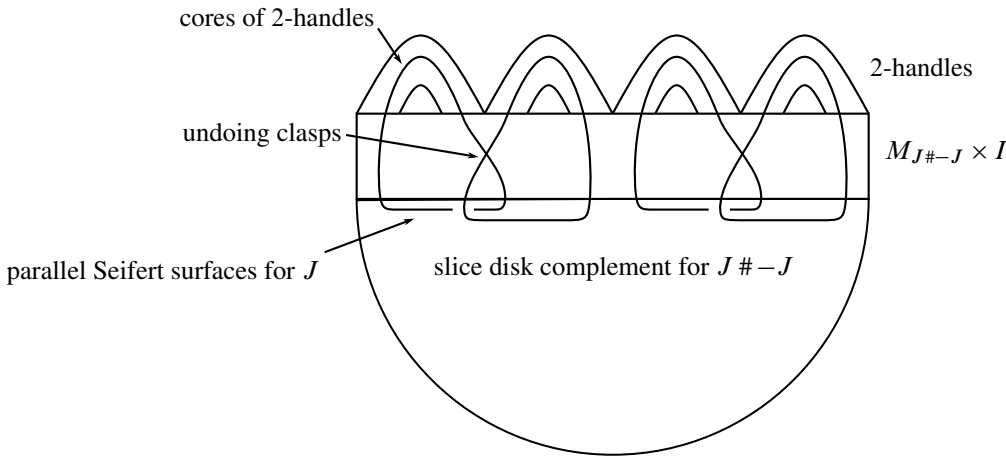


Figure 7: Schematic picture of Z with embedded surfaces generating $H_2(Z)$ (1-handles are not pictured).

the signs of the corresponding self-intersections of A . Now, paying attention to the labels in Figure 6, we see that $\partial^-(N \setminus \nu(A))$ bounds a manifold Z (framings written inside angle brackets denote a surgery description of a 3-manifold, which in this case is $M_{J\#-J}$), with schematic picture shown in Figure 7.

It is straightforward to compute the homology of Z and $\partial Z = \partial^-(N \setminus \nu(A))$, with results as in the following two lemmas.

- Lemma 3.3** (1) $H_1(Z) = \mathbb{Z}^{c+1}$, where c is the number of self-intersections of A , and the extra generator is the meridian of the slice disk for $J \# -J$.
- (2) $H_2(Z) = \mathbb{Z}^{2c}$, and is represented by surfaces which are the unions of the cores of the 2-handles in Z , the homotopies through $M_{J\#-J} \times I$ which consist of undoing the clasps in the attaching circles of the 2-handles in the construction of Z , and parallel Seifert surfaces for J . See Figure 7.

- Lemma 3.4** (1) $H_1(\partial Z) = \mathbb{Z}^{c+1}$ (and $H_1(\partial Z) \rightarrow H_1(Z)$ is an isomorphism).
- (2) $H_2(\partial Z) = \mathbb{Z}^{c+1}$ (and $H_2(\partial Z) \rightarrow H_2(Z)$ is the zero map). Moreover, $H_2(\partial Z) \rightarrow H_2(N \setminus \nu(A))$ takes generators to corresponding generators.

Now let $N' = (N \setminus \nu(A)) \cup Z$.

- Lemma 3.5** (1) $H_1(N') \cong \mathbb{Z}$ and is generated by λ .
- (2) $H_2(Z) \rightarrow H_2(N')$ is an isomorphism.

Proof Consider the Mayer–Vietoris sequence

$$\begin{array}{ccccccc} H_2(\partial Z) & \longrightarrow & H_2(Z) \oplus H_2(N \setminus \nu(A)) & \longrightarrow & H_2(N') & & \\ & & & & \downarrow & & \\ & & & & H_1(\partial Z) & \longrightarrow & H_1(Z) \oplus H_1(N \setminus \nu(A)) \longrightarrow H_1(N') \longrightarrow 0. \end{array}$$

Since $H_1(\partial Z) \rightarrow H_1(Z)$ is an isomorphism, $H_1(N \setminus \nu(A)) \rightarrow H_1(N')$ is an isomorphism. Composing with isomorphisms from before, we get (1). We also see that $H_2(N') \rightarrow H_1(\partial Z)$ is the zero map.

Since $H_2(\partial Z) \rightarrow H_2(Z)$ is also the zero map, and $H_2(\partial Z) \rightarrow H_2(N \setminus \nu(A))$ takes generators to generators, $H_2(Z) \rightarrow H_2(N')$ is an isomorphism. $\square$

We can now prove our main result:

Proof of Theorem 1.4 By Lemma 3.5, $H_2(N')$ is generated by the same surfaces that generate $H_2(Z)$. These surfaces consist of Seifert surfaces which are capped off by 0-framed 2-handles (along with the homotopies that preserve the framing), and so have trivial normal bundle. They intersect geometrically once in pairs, and the generators of their fundamental groups lie on a Seifert surface for J , and thus are in the commutator subgroup of $\pi_1(Z)$, which has a natural homomorphism into $\pi_1(N')$. Therefore N' is a (1)-solution for $P_0(J) \# -P_1(J)$. $\square$

Remark 3.6 If the map on fundamental groups induced by the inclusion from the Seifert surface for J to the slice disk complement for $J \# -J$ has image in the n^{th} term of the derived series, then N' will be an (n) -solution for $P_0(J) \# -P_1(J)$. We did not have to use the usual slice disk for $J \# -J$, and one might be able to get a better result by choosing a different slice disk for $J \# -J$. In Section 4, we will give even stronger conditions that would guarantee that $P_0(J) \# -P_1(J)$ is topologically slice.

Given a slice knot, one might wonder if it is ribbon:

Definition 3.7 A (smoothly) slice knot K is *ribbon* if it has a (smooth) slice disk Δ such that the restriction of the radial Morse function on B^4 to Δ has only index 0 and 1 critical points.

This gives rise to the following famous conjecture:

Conjecture 3.8 (slice–ribbon conjecture) Every slice knot is ribbon.

One can weaken ribbonness to the following algebraic topological condition, which makes sense in both the smooth and topological settings (unlike ribbonness, which relies on a Morse function):

Definition 3.9 A (smoothly/topologically) slice knot is *(smoothly/topologically) homotopy ribbon* if it bounds a (smooth/topological) slice disk Δ such that the inclusion-induced map $\pi_1(M_K) \rightarrow \pi_1(B^4 \setminus \nu(\Delta))$ is surjective.

Every ribbon knot is homotopy ribbon. This gives rise to the topological analogue of the slice–ribbon conjecture:

Conjecture 3.10 Every topologically slice knot is topologically homotopy ribbon.

Smooth homotopy ribbonness can be extended to (n) -solutions:

Definition 3.11 An (h) -solvable knot for $h \in \frac{1}{2}\mathbb{N}$ is *homotopy ribbon (h) -solvable* if it bounds an (h) -solution W such that the inclusion map $\pi_1(M_K) \rightarrow \pi_1(W)$ is surjective. The manifold W is then called a *homotopy ribbon (h) -solution*.

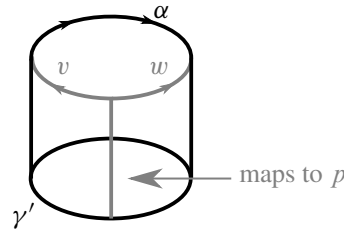


Figure 8: The arc α is homotopic fixing endpoints to the arc traveling along one whisker, then traversing γ' , and finally traveling back along the other whisker.

Note that the homotopy ribbon (h)-solution must be a smooth manifold.

The (1)-solution we have constructed for $P_0(J) \# -P_1(J)$ is in fact a homotopy ribbon (1)-solution when $K_0 \# -K_1$ is smoothly homotopy ribbon:

Proposition 3.12 *In the notation of Definition 1.2 and Theorem 1.4, if $K_0 \# -K_1$ is smoothly homotopy ribbon via the disk $C \setminus$ (an arc running from K_0 to K_1) (where C is the concordance between K_0 and K_1), then the (1)-solution N' for $P_0(J) \# -P_1(J)$ constructed in Theorem 1.4 is in fact a homotopy ribbon (1)-solution.*

Before proving this proposition, we will need the following lemma:

Lemma 3.13 *Given a 4-manifold W with connected boundary M such that the inclusion map $\pi_1(M) \rightarrow \pi_1(W)$ is surjective, any arc α in W with both ends in M may be homotoped to an arc in M fixing the endpoints of the arc.*

Proof Choose the basepoint of both $\pi_1(M)$ and $\pi_1(W)$ to be a point $p \in M$. Choose whiskers v and w from the basepoint to each of the ends of α inside M . This gives a based loop $\gamma = v\alpha w^{-1}$. The loop γ is then homotopic (rel basepoint) to a loop γ' lying entirely in $\partial W = M$, ie there is a continuous map from an annulus to W which on the boundary is $\gamma \cup \gamma'$. Therefore, the arc $v^{-1}\gamma'w$ is homotopic rel boundary to α , by homotoping along the map from an annulus (see Figure 8). $\square$

Proof of Proposition 3.12 Recall that $N' = (N \setminus \nu(A)) \cup Z$, where N has schematic as depicted in Figure 3 and Z has schematic as depicted in Figure 7 and partial Kirby diagram as shown in Figure 6. Take any based loop $\gamma \in \pi_1(N')$. We will homotope this curve to lie entirely in $\pi_1(M_{P_0(J) \# -P_1(J)})$, and, abusing notation, will also denote each step along the way by γ . Without loss of generality, the intersection of this curve with Z is a collection of arcs.

Since Z consists of the complement of a ribbon disk for $J \# -J$, with some 1- and 2-handles attached, $\pi_1(\partial Z) \rightarrow \pi_1(Z)$ is surjective since an inclusion-induced map from the boundary of a ribbon disk complement is already surjective, and a standard argument shows that attaching 1- and 2-handles will preserve surjectivity. Therefore, using Lemma 3.13, each of the arcs may be homotoped fixing their endpoints to lie in ∂Z , and then pushed off to lie entirely in $N \setminus \nu(A)$. The intersections of γ with the

$(M_J \times I)_\epsilon$ can be pushed straight outward to $(M_J \times \{0\})_\epsilon$. Since the attaching regions between W and the $(M_J \times I)_\epsilon$ are neighborhoods of curves, we can perturb γ slightly so that it does not intersect them. At this point, γ lies entirely in $(W \setminus \nu(A)) \cup \partial N'$.

Now, the intersections of γ with $W \setminus \nu(A)$ without loss of generality again consist of a collection of arcs. Since $K_0 \# -K_1$ is homotopy ribbon, these arcs could be homotoped to ∂W , again using [Lemma 3.13](#) (since $(S^3 \times I) \setminus \nu(C)$ is homeomorphic to the exterior of a homotopy ribbon disk for $K_0 \# -K_1$), but might hit A along the way. However, this is no problem, as a meridian of A can be homotoped to a curve in ∂Z which is isotopic through Z to the attaching curve of any of the 2-handles (it is a meridian of an unknotted curve with framing $\langle 0 \rangle$), and so in particular bounds a disk. Therefore, we can replace neighborhoods of potential intersections with A with these disks, thus achieving the desired homotopy.

So now γ lies entirely in $\partial(N') = M_{P_0(J) \# -P_1(J)}$, and we are done. $\square$

Remark 3.14 This proof holds as long as we use a smooth homotopy ribbon disk for $J \# -J$ in the proof of [Theorem 1.4](#). In particular, as long as we use a (smooth) homotopy ribbon disk for $J \# -J$, [Remark 3.6](#) will still apply. Moreover, [Proposition 4.1](#) below will still apply even if we only use a topological homotopy ribbon disk for $J \# -J$, though this involves an application of the sphere embedding theorem up to s -cobordism with a π_1 -null condition and topological input. See [Theorem 6.1](#) of [\[18\]](#) as well as the discussion at the beginning of Chapter 5 and Sections 5.2 and 5.3.

4 Some conditions that would guarantee topological sliceness

Proposition 4.1 *If there is a smooth or topological slice disk complement for $J \# -J$ in which any number of parallel longitudes of J bound disjoint framed π_1 -null immersed disks with algebraically trivial self-intersections, then $P_0(J) \# -P_1(J)$ is topologically slice.*

Before proving this proposition, we will review some definitions and techniques from the theory of 4-manifolds. These make sense in both the smooth and topological settings.

Definition 4.2 An h -cobordism rel boundary between two manifolds M^n and N^n with homeomorphic (possibly empty) boundary P is a manifold W^{n+1} with boundary $M \cup (P \times I) \cup N$, where M and N are glued to $P \times I$ along their boundaries, such that W deformation retracts to M and to N . Given homeomorphic properly immersed submanifolds $A \looparrowright M$, $B \looparrowright N$, an h -cobordism rel boundary of the pairs (M, A) and (N, B) is an h -cobordism rel boundary W of M and N together with a proper immersion $A \times I \looparrowright W$ which is a product on $P \times I$ and such that the restriction to $A \times \{0\}$ gives $A \looparrowright M$ and the restriction to $A \times \{1\}$ gives $B \looparrowright N$. See [Figure 9](#).

Definition 4.3 An s -cobordism (rel boundary of pairs) is an h -cobordism (rel boundary of pairs) such that the Whitehead torsion $\tau(W, M)$ vanishes. For a definition and discussion of the Whitehead torsion in the topological setting, see [\[13, Section IV\]](#). For a definition and discussion in the smooth setting, see [\[32\]](#).

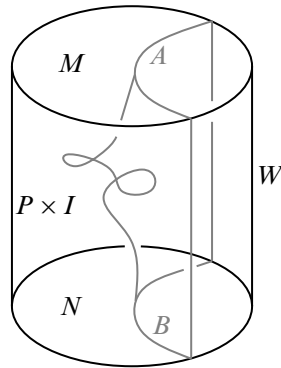


Figure 9: A schematic picture of an h -cobordism rel boundary of pairs. Note that A and B may also be immersed, but must be homeomorphic.

Now we will describe a 4-dimensional analogue of 0-surgery. Given a sphere embedded in a 4-manifold $S^2 \hookrightarrow W^4$ with trivial normal bundle (that is, the given embedding extends to an embedding of $S^2 \times D^2$), we can cut out a neighborhood of the sphere. The new boundary is naturally homeomorphic to $S^2 \times S^1$. We then replace the piece we cut out with $D^3 \times S^1$, gluing along the identity homeomorphism $S^2 \times S^1 \rightarrow S^2 \times S^1$. If we are given two such spheres which intersect once transversely, each of which represents a generator of a $\mathbb{Z}$ summand of $H_2(W)$, then, by Seifert–Van Kampen and Mayer–Vietoris, doing surgery on one of the two spheres exactly kills the $\mathbb{Z}^2$ summand in $H_2(W)$ and leaves $\pi_1(W)$ and the rest of the homology groups unchanged.

Proof of Proposition 4.1 In the construction of Z depicted in Figures 6 and 7, use the slice disk complement for $J \# -J$ given by the conditions of Proposition 4.1. (For the purposes of Theorem 1.4, we could have used any slice disk complement for $J \# -J$.) The meridian of η , together with the longitudes of J which are freely homotopic to the attaching circles of the 2-handles, all bound disjoint framed π_1 -null immersed disks in the slice disk complement (these are all parallel longitudes of J); see Figure 10. Since the meridian of η is then freely nullhomotopic in Z , we see that $\pi_1(N')$ is normally generated by λ . By the topological

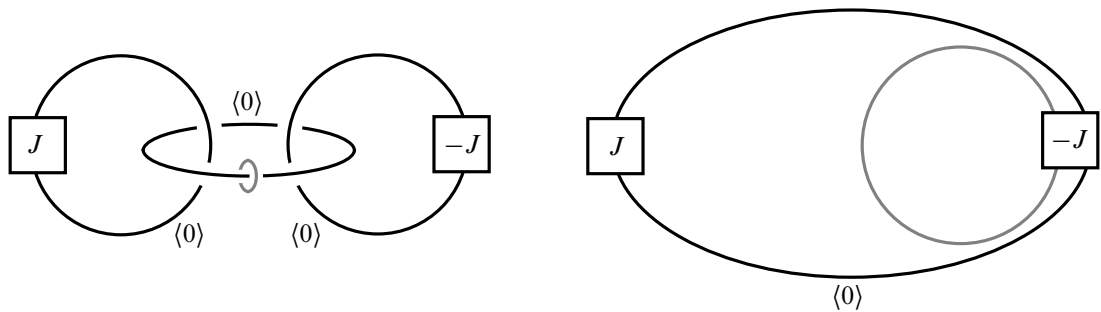


Figure 10: Two equivalent diagrams of the boundary of a slice disk complement for $J \# -J$. Any number of parallel copies of the gray curve must bound disjoint immersed π_1 -null disks in order to prove $P_0(J) \# -P_1(J)$ is slice.

version of [18, Theorem 6.1], there is an s -cobordism rel boundary of the pair $(N', \text{immersed spheres})$ to a pair $(N'', \text{embedded spheres})$, where the embedded spheres still come in pairs which intersect once transversely. (See the discussion at the beginning of [18, Chapter 5] for how to extend Theorem 6.1 to the topological setting.) Moreover, since the original disks were framed, these embedded spheres have trivial normal bundles. We can then do surgery on one sphere from each pair as described above to get a manifold that satisfies the conditions in Proposition 2.2, and so $P_0(J) \# -P_1(J)$ is topologically slice. $\square$

Remark 4.4 In the usual slice disk complement for $J \# -J$, the parallel longitude J is not nullhomotopic, as the slice disk complement is homeomorphic to $(S^3 \setminus J) \times I$ and the longitude of a knot is not nullhomotopic in its complement. So, in order for the conditions of Proposition 4.1 to be satisfied, we would need to find a different (potentially topological) slice disk. A promising approach might be the work of Friedl and Teichner [20] and Conway [15]. The first gives an algebraic condition on a surjection $\pi_1(M_K) \twoheadrightarrow G$ which guarantees that K has a topological homotopy ribbon disk with fundamental group G (this is a generalization of the theorem of Freedman that any knot with Alexander polynomial 1 is topologically $\mathbb{Z}$ -slice, as their condition reduces to the Alexander polynomial 1 condition for the abelianization map $\pi_1(M_K) \rightarrow \mathbb{Z}$). Conway expands on this, giving an algebraic classification of topological homotopy ribbon disks corresponding to a given surjection $\pi_1(M_K) \twoheadrightarrow G$. So, if we could find a surjection $\pi_1(M_{J \# -J}) \twoheadrightarrow G$ for some group G which satisfies the conditions in the paper and which kills the longitude of J , this would be a large step towards producing a slice disk which satisfies the conditions of Proposition 4.1. However, it is usually quite difficult to tell whether a particular surjection $\pi_1(M_{J \# -J}) \twoheadrightarrow G$ satisfies the relevant algebraic conditions.

5 Casson–Gordon invariants

We will now review the setup for Casson–Gordon invariants, first defined in [1]. For any knot K , let $L_{K,n}$ be the n -fold cyclic branched cover of S^3 over K . Let $\chi: H_1(L_{K,n}) \rightarrow \mathbb{Z}_m$ be a character. We will also use χ to denote the postcomposition of this map with the map $\mathbb{Z}_m \rightarrow \mathbb{C}^*$ sending $1 \mapsto e^{2\pi i/m}$. To each such character, Casson and Gordon associate a *Witt class* $\tau(K, \chi) \in L_0(\mathbb{C}(t)) \otimes \mathbb{Q}$.

The *Witt group* $L_0(k)$ for a field k is the set of equivalence classes of nondegenerate symmetric bilinear forms, where two forms are equivalent if one can direct sum a metabolic form to each and get isometric forms. A metabolic form is a form which has a half-dimensional subspace on which it vanishes. Such a subspace is called a *metabolizer*. The set of nondegenerate symmetric bilinear forms under this equivalence relation forms a group under the direct sum. For more information on Witt groups, see eg [29, Chapter 5].

The following is a theorem of Casson and Gordon:

Theorem 5.1 [1, Theorem 2] *Fix n a prime power. If K is smoothly or topologically slice, then the linking form on $H_1(L_{K,n})$ has a metabolizer R , and, for any character χ with prime-power order m which vanishes on R , we have $\tau(K, \chi) = 0$.*

The sliceness condition on K can be weakened to (1.5)-solvability:

Theorem 5.2 (see [7, Theorem B.1; 12, Theorem 9.11]) *The conclusion in Theorem 5.1 holds even if K is only (1.5)-solvable.*

So, Casson–Gordon invariants give us an obstruction to (1.5)-solvability, which we will show fails for $P_0(J) \# -P_1(J)$.

Proposition 5.3 *The conclusion of Theorems 5.1 and 5.2 hold for $P_0(J) \# -P_1(J)$. In other words, Theorem 5.2 fails to obstruct the (1.5)-solvability of the knot $P_0(J) \# -P_1(J)$.*

Before proving this proposition, we will review some results of Litherland about the Casson–Gordon invariants of satellite knots; see [26; 27]. For a satellite operator P given by $K \cup \eta$, let $w = \text{lk}(K, \eta)$ be the winding number of P . Let $h = \gcd(n, w)$ and let $k = n/h$. Write $\text{Ch}_n(K) = \text{Hom}(H_1(L_{K,n}), \mathbb{C}^*)$ for the group of characters. There are canonical isomorphisms

$$H_1(L_{P(J),n}) \cong H_1(L_{K,n}) \oplus hH_1(L_{J,k}) \quad \text{and} \quad \text{Ch}_n(P(J)) \cong \text{Ch}_n(K) \oplus h\text{Ch}_k(J),$$

where hA is the direct sum of h copies of the group A , and the linking form on $H_1(L_{P(J),n})$ is given by the direct sum of the linking forms on the summands. (When $k = 1$, the 1-fold branched cover $L_{J,1}$ is just S^3 , and so $\text{Ch}_1(J)$ is the trivial group.)

Then, given a character $\chi \in \text{Ch}_n(P_\epsilon(J))$ which is identified under this isomorphism with $(\chi_K, \chi_1, \dots, \chi_h)$, we have

$$\tau(P(J), \chi)[t] = \tau(K, \chi_K)[t] + \sum_{i=1}^h \tau(J, \chi_i)[\chi_K(\tilde{\eta}^i) t^{w/h}],$$

where the $\tilde{\eta}^i$ are the h lifts of η to $L_{K,n}$ (or, more precisely, the h lifts of $k\eta$).

Moreover, for any knot G , the map $\cdot \# G$ is the satellite operator given by $G \cup \mu_G$, where μ_G is the meridian of G . In this case, $w = h = 1$ and $k = n$, so we get

$$\tau(F \# G, \chi) = \tau(F, \chi_F) + \tau(G, \chi_G)$$

(note that $\tilde{\mu}_G$ bounds a disk and in particular is nullhomologous in any cyclic branched cover over G , and so $\chi_G(\tilde{\mu}_G) = 1 \in \mathbb{C}^*$).

Proof of Proposition 5.3 From the discussion above, we get the canonical isomorphisms

$$\begin{aligned} H_1(L_{P_0(J) \# -P_1(J),n}) &\cong H_1(L_{P_0(J),n}) \oplus H_1(L_{-P_1(J),n}) \\ &\cong H_1(L_{K_0,n}) \oplus hH_1(L_{J,k}) \oplus H_1(L_{-K_1,n}) \oplus hH_1(L_{-J,k}) \\ &\cong H_1(L_{K_0 \# -K_1,n}) \oplus hH_1(L_{J \# -J,k}). \end{aligned}$$

Similarly, we have canonical isomorphisms

$$\begin{aligned} \text{Ch}_n(P_0(J) \# -P_1(J)) &\cong \text{Ch}_n(K_0) \oplus h\text{Ch}_k(J) \oplus \text{Ch}_n(-K_1) \oplus h\text{Ch}_k(-J) \\ &\cong \text{Ch}_n(K_0 \# -K_1) \oplus h\text{Ch}_k(J \# -J). \end{aligned}$$

By Theorem 5.1, the linking forms on branched cyclic covers of $K_0 \# -K_1$ and $J \# -J$ have metabolizers since these knots are topologically slice. Denote the metabolizer of $H_1(L_{K_0 \# -K_1, n})$ by R and the metabolizer of $H_1(L_{J \# -J, l})$ by S . The image of the submodule $R \oplus hS$ under the canonical isomorphism is thus a metabolizer for the linking form on $H_1(L_{P_0(J) \# -P_1(J), n})$. A character $\chi \in \text{Ch}_n(P_0(J) \# -P_1(J))$ vanishes on $R \oplus hS$ if and only if φ vanishes on R and each ψ^i vanishes on S , where $\varphi \oplus \bigoplus_{i=1}^h \psi^i \in \text{Ch}_1(L_{K_0 \# -K_1, n}) \oplus h \text{Ch}_1(L_{J \# -J, k})$ corresponds to χ .

Now let χ be a character which vanishes on $R \oplus hS$. Using the canonical isomorphisms above, decompose χ as

$$\chi_{P_0(J)} \oplus \chi_{-P_1(J)} = \chi_{K_0} \oplus \bigoplus_{i=1}^h \chi_J^i \oplus \chi_{-K_1} \oplus \bigoplus_{i=1}^h \chi_{-J}^i,$$

where the subscript denotes the knot whose character group contains the character.

Now we have

$$\begin{aligned} \tau(P_0(J) \# -P_1(J), \chi)[t] &= \tau(P_0(J), \chi_{P_0(J)}[t] + \tau(-P_1(J), \chi_{-P_1(J)}[t]) \\ &= \tau(K_0, \chi_{K_0})[t] + \sum_{i=1}^h \tau(J, \chi_J^i)[\chi_{K_0}(\tilde{\eta}_0^i) t^{w/h}] + \tau(-K_1, \chi_{-K_1})[t] \\ &\quad + \sum_{i=1}^h \tau(-J, \chi_{-J}^i)[\chi_{-K_1}(-\tilde{\eta}_1^i) t^{w/h}]. \end{aligned}$$

By $\tilde{\eta}_\epsilon^i$ we are denoting the lifts of η_ϵ for $\epsilon \in \{0, 1\}$.

Now we wish to show that $\chi_{K_0}(\tilde{\eta}_0^i) = \chi_{-K_1}(-\tilde{\eta}_1^i)$, after possibly reindexing the lifts of the η_ϵ . First consider the n -fold cyclic branched cover of B^4 over the natural slice disk for $K_0 \# -K_1$. The boundary of this branched cover is the n -fold cyclic branched cover of $K_0 \# -K_1$. Now, by the homotopy lifting property, the h lifts of η_0 are each joined to one of the h lifts of $-\eta_1$ through an immersed annulus in the branched cover of B^4 which is a lift of A (these lifts are all k -fold covers). Therefore, after possibly reindexing, $\tilde{\eta}_0^i \oplus -(-\tilde{\eta}_1^i)$ is nullhomologous in the branched cover of B^4 , so is in the kernel of the inclusion-induced map, and therefore lies in the metabolizer R of the linking form on $H_1(L_{K_0 \# -K_1, n})$. (We write two minus signs since the $-\tilde{\eta}_1^i$ are lifts of $-\eta_1$, and $-(-\tilde{\eta}_1^i)$ are the opposites of the homology classes of these lifts.)

Now, since φ vanishes on the metabolizer R of $H_1(L_{K_0 \# -K_1})$,

$$0 = \varphi(\tilde{\eta}_0^i \oplus -(-\tilde{\eta}_1^i)) = \chi_{K_0}(\tilde{\eta}_0^i) + \chi_{-K_1}(-(-\tilde{\eta}_1^i)) = \chi_{K_0}(\tilde{\eta}_0^i) - \chi_{-K_1}(-\tilde{\eta}_1^i),$$

and so indeed $\chi_{K_0}(\tilde{\eta}_0^i) = \chi_{-K_1}(-\tilde{\eta}_1^i)$.

Therefore,

$$\tau(P_0(J) \# -P_1(J), \chi) = \tau(K_0 \# -K_1, \varphi)[t] + \sum_{i=1}^h \tau(J \# -J, \psi)[\chi_{K_0}(\tilde{\eta}_0^i) t^{w/h}] = 0 + \sum_{i=1}^h 0 = 0$$

since $K_0 \# -K_1$ and $J \# -J$ are both topologically slice and the corresponding characters vanish on the metabolizer by tracing through the isomorphism on H_1 (we could have equivalently written $\chi_{-K_1}(-\tilde{\eta}_1^i)$ instead of $\chi_{K_0}(\tilde{\eta}_1^i)$ in the final equation). $\square$

6 Metabelian ρ invariants

Recall that the rational Alexander module $\mathcal{A}_0(K) = H_1(M_K, \mathbb{Q}[t^{\pm 1}])$ comes equipped with the Blanchfield form

$$\text{Bl}_0: \mathcal{A}_0(K) \times \mathcal{A}_0(K) \rightarrow \mathbb{Q}(t)/\mathbb{Q}[t^{\pm 1}]$$

defined in the following way: since $\mathcal{A}_0(K)$ is $\mathbb{Q}[t^{\pm 1}]$ -torsion, given homology classes $[x], [y] \in \mathcal{A}_0(K)$, there is a Laurent polynomial $p \in \mathbb{Q}[t^{\pm 1}]$ such that py is nullhomologous. That is, there is some 2-chain $w \in C_2(M_K, \mathbb{Q}[t^{\pm 1}])$ such that $py = \partial w$. Then

$$\text{Bl}_0([x], [y]) = \frac{1}{p} \sum_{n \in \mathbb{Z}} \langle t^n x, w \rangle t^{-n}.$$

This can be thought of as a linking form which is equivariant with respect to the deck action.

For a 3-manifold M (for our purposes, this will always be 0-surgery on a knot) and a representation $\varphi: \pi_1(M) \rightarrow \Gamma$, Cheeger and Gromov [2] defined *von Neumann ρ -invariants* $\rho(M, \varphi)$. Typically, we will assume that Γ is *poly-torsion-free-abelian* (PTFA), which means that it admits a normal series

$$\{1\} \triangleleft \Gamma_1 \triangleleft \cdots \triangleleft \Gamma_{n-1} \triangleleft \Gamma_n = \Gamma$$

such that each quotient Γ_{i+1}/Γ_i is torsion-free abelian. If M bounds a 4-manifold W , and ψ extends to $\pi_1(W)$, then $\rho(M, \varphi)$ is given by the difference $\sigma_{\Gamma}^{(2)}(W, \psi) - \sigma(W)$, where $\sigma(W)$ is the ordinary signature of W and $\sigma_{\Gamma}^{(2)}(W, \psi)$ is a certain kind of twisted signature. In other words, the ρ -invariants can be thought of as “signature defects”. We will not define them rigorously here, and instead will use a few key properties:

- Given an injection $\Gamma \hookrightarrow \Gamma'$, we get a map $\varphi': \pi_1(M) \rightarrow \Gamma \hookrightarrow \Gamma'$. In this situation, $\rho(M, \varphi) = \rho(M, \varphi')$.
- If φ is the zero map, then $\rho(M, \varphi) = 0$.
- If K is slice, Γ is a PTFA group and $\varphi: \pi_1(M_K) \rightarrow \Gamma$ extends to the fundamental group of the exterior of a slice disk, then $\rho(M_K, \varphi) = 0$.

We will also abuse notation slightly, writing $\rho(K, \varphi) := \rho(M_K, \varphi)$.

Given a submodule $P \subset \mathcal{A}_0(K)$, we get a normal subgroup of $\pi_1(M_K)$,

$$\tilde{P} = \ker(\pi_1(M_K)^{(1)} \twoheadrightarrow \pi_1(M_K)^{(1)}/\pi_1(M_K)^{(2)} = \mathcal{A}(K) \rightarrow \mathcal{A}_0(K) \twoheadrightarrow \mathcal{A}_0(K)/P),$$

where $\mathcal{A}(K) = H_1(M_K; \mathbb{Z}[t^{\pm 1}])$ is the ordinary Alexander module. A priori, $\tilde{P}$ is only normal in $\pi_1(M_K)^{(1)}$. But, since $\pi_1(M_K) = \pi_1(M_K)^{(1)} \rtimes \mathbb{Z}$, we can write any element of $\pi_1(M_K)$ as $t^k a$, where

t is the generator of $\mathbb{Z}$ and $a \in \pi_1(M_K)^{(1)}$. Then, given an element $p \in \tilde{P}$, we see that $(t^k a)p(t^k a)^{-1} = t^k (apa^{-1})t^{-k}$. Thus, since $\tilde{P}$ is normal in $\pi_1(M_K)^{(1)}$, we see that $apa^{-1} \in \tilde{P}$, and therefore its image in $\mathcal{A}_0(K)$ is in P . But then, since P is a submodule of $\mathcal{A}_0(K)$, the element $t^k (apa^{-1})t^{-k} = t^k \cdot apa^{-1}$ must also be in P , and therefore $(t^k a)p(t^k a)^{-1} \in \tilde{P}$. Thus, $\tilde{P}$ is normal in the larger group $\pi_1(M_K)$.

The normal subgroup $\tilde{P}$ gives us a quotient map $\varphi_{\tilde{P}}: \pi_1(M_K) \twoheadrightarrow \pi_1(M_K)/\tilde{P}$. Since $\pi_1(M_K)^{(2)} \subseteq \tilde{P}$, this quotient is metabelian. As in [9], the groups $\pi_1(M_K)/\tilde{P}$ are all PTFA, and so we will not need to worry about this condition in the following discussion.

Recall from Section 5 that, given a form on some space (for example the Blanchfield form Bl_0 on the rational Alexander module $\mathcal{A}_0(K)$ of a knot K), a *metabolizer* is a subspace which is its own orthogonal complement. Metabolizers of the Blanchfield form are intimately related with sliceness of knots: if K is slice with a slice disk Δ , then

$$P_{\Delta} = \ker(\mathcal{A}_0(K)) = H_1(M_K; \mathbb{Q}[t^{\pm 1}]) \rightarrow H_1(B^4 \setminus \Delta; \mathbb{Q}[t^{\pm 1}])$$

is a metabolizer for the Blanchfield form on $\mathcal{A}_0(K)$.

The following is a theorem of Cochran, Harvey and Leidy:

Theorem 6.1 [9, Theorem 4.2] *If K is slice, then, for some metabolizer $P_{\Delta} \subset \mathcal{A}_0(K)$ corresponding to a slice disk Δ , we have $\rho(K, \varphi_{\tilde{P}_{\Delta}}) = 0$. In particular, the set*

$$\{\rho(K, \varphi_{\tilde{P}}) \mid P \text{ is a metabolizer for } \mathcal{A}_0(K)\}$$

contains 0.

Cochran, Harvey and Leidy showed that, just as with Casson–Gordon invariants, the condition of sliceness in the previous theorem can be weakened to (1.5)-solvability:

Theorem 6.2 [9, Theorem 10.1] *If K is (1.5)-solvable, then the set*

$$\{\rho(K, \varphi_{\tilde{P}}) \mid P \text{ is a metabolizer for } \mathcal{A}_0(K)\}$$

contains 0.

Similarly to Casson–Gordon invariants, we can write the metabelian ρ invariants of a satellite knot $P(J)$ in terms of the ρ invariants of $K = P(U)$ and J . For this we will need to choose a basepoint common to $M_{P(J)}$, $M_K \setminus \nu(\eta)$ and $S^3 \setminus \nu(J)$, as well as meridians and longitude for the boundaries of $M_K \setminus \nu(\eta)$ and $S^3 \setminus \nu(J)$. Our choices are shown in Figure 11 for the particular case of the Mazur pattern. We will implicitly carry this choice of basepoint with us throughout the following discussion.

Let Γ be a metabelian group. Given a homomorphism $\psi: \pi_1(M_{P(J)}) \rightarrow \Gamma$, we get homomorphisms $\psi_K: \pi_1(M_K \setminus \nu(\eta)) \rightarrow \Gamma$ and $\psi_J: \pi_1(S^3 \setminus \nu(J)) \rightarrow \Gamma$ since $M_K \setminus \nu(\eta)$ and $S^3 \setminus \nu(J)$ are both submanifolds of $M_{P(J)}$. When gluing together $M_K \setminus \nu(\eta)$ and $S^3 \setminus \nu(J)$ to get $M_{P(J)}$ a meridian μ_{η} of η and a longitude λ_J of J are identified. And, since λ_J lies in $\pi_1(S^3 \setminus \nu(J))^{(2)}$, which includes into $\pi_1(M_{P(J)})^{(2)}$, we must have that ψ sends the image of λ_J (which is also the image of μ_{η}) to $0 \in \Gamma$.

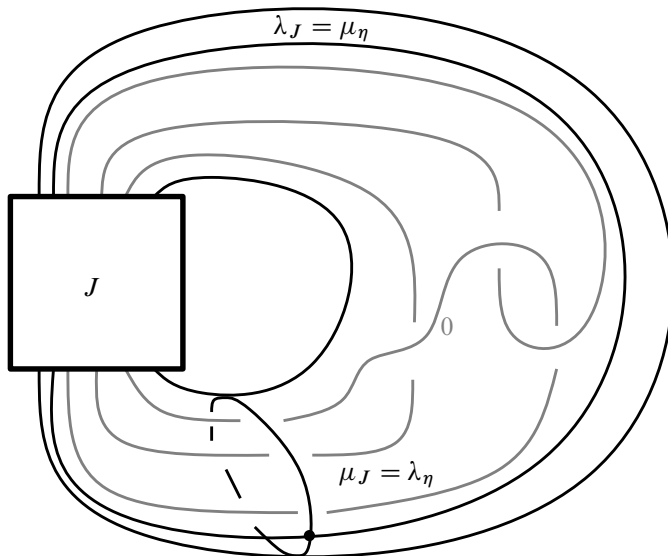


Figure 11: Decomposition of M_{P_J} as $M_K \setminus \nu(\eta)$ and $S^3 \setminus J$ for the Mazur pattern. The torus, with K sitting inside, should be tied into J , and so it bounds the two submanifolds on either side. The basepoint is the intersection of the two labeled curves.

Therefore ψ_K and ψ_J extend uniquely to homomorphisms $\pi_1(M_K) \rightarrow \Gamma$ and $\pi_1(M_J) \rightarrow \Gamma$, which we will also denote by ψ_K and ψ_J , respectively. Similarly, $\psi_K(\eta) = \psi_J(\mu_J)$.

Conversely, given homomorphisms $\psi_K: \pi_1(M_K) \rightarrow \Gamma$ and $\psi_J: \pi_1(M_J) \rightarrow \Gamma$ such that $\psi_K(\eta) = \psi_J(\mu_J)$, we can restrict them to $M_K \setminus \nu(\eta)$ and $S^3 \setminus \nu(J)$, and they will be compatible on the torus boundary when we glue these two manifolds together to get $M_{P(J)}$ (notice that the meridian of η is nullhomotopic in M_K and the longitude of J is in $\pi_1(M_J)^{(2)}$ and so both must be sent to $1 \in \Gamma$). Thus, ψ_K and ψ_J give a homomorphism $\psi: \pi_1(M_{P(J)}) \rightarrow \Gamma$ whose restrictions to $M_K \setminus \nu(\eta)$ and $S^3 \setminus \nu(J)$ are the homomorphisms coming from ψ_K and ψ_J , respectively.

In summary, there is a bijection

$$\text{Hom}(\pi_1(M_{P(J)}), \Gamma) \leftrightarrow \{(\psi_K: \pi_1(M_K) \rightarrow \Gamma, \psi_J: \pi_1(M_J) \rightarrow \Gamma) \mid \psi_K(\eta) = \psi_J(\mu_J)\}.$$

The following is [9, Lemma 2.1]:

Lemma 6.3 *In the notation in the previous paragraph,*

$$\rho(P(J), \varphi) = \rho(K, \varphi_K) + \rho(J, \varphi_J).$$

Recall the setup from [Section 1](#): we have satellite operators P_0 and P_1 given by $K_0 \sqcup \eta_0$ and $K_1 \sqcup \eta_1$, respectively, which are homotopically related. That is, K_0 and K_1 are concordant, and $\eta_0 \times \{0\}$ and $\eta_1 \times \{1\}$ are homotopic through the complement of a concordance between K_0 and K_1 . We showed in [Theorem 1.4](#) that, for any J , the difference $P_0(J) \# -P_1(J)$ is (1)-solvable.

Proposition 6.4 *The set*

$$\{\rho(P_0(J) \# -P_1(J), \varphi_P) \mid P \text{ is a metabolizer for } \mathcal{A}_0(P_0(J) \# -P_1(J))\}$$

contains 0. In other words, Theorem 6.2 fails to obstruct the (1.5)-solvability of the knot $P_0(J) \# -P_1(J)$.

Before beginning the proof of Proposition 6.4, we will fix some notation which will make it easier to work with metabelian groups.

Definition 6.5 The *metabelianization* of a group G is defined to be $G/G^{(2)}$ and will be denoted by G^{mab} .

Since homomorphisms to a metabelian group factor through the metabelianization of the domain, we can work with homomorphisms from metabelianizations. Since the longitude of a knot K lies in the second commutator subgroup of the knot group, $\pi_1(K)^{\text{mab}} \cong \pi_1(M_K)^{\text{mab}}$, and so we can replace $\pi_1(M_K)$ with $\pi_1(K)$ or $\pi_1(K)^{\text{mab}}$ in the discussion above. It will be easier to think in terms of knot groups than in terms of fundamental groups of 0-surgeries, and this is what we will do for the remainder of this section.

We have the following group-theoretic lemma:

Lemma 6.6 *There is a canonical isomorphism $(G/G^{(2)})^{(1)} \cong G^{(1)}/G^{(2)}$.*

This lemma gives us the following corollary, since the Alexander module of a knot group is the abelianization of its commutator subgroup:

Corollary 6.7 *The Alexander module of a knot is the commutator subgroup of the metabelianization of the knot group.*

Proof of Lemma 6.6 Throughout, we will denote by $\bar{x}$ the equivalence class in $G/G^{(2)}$ or $G^{(1)}/G^{(2)}$ of an element x of G or $G^{(1)}$.

Define a map $(G/G^{(2)})^{(1)} \rightarrow G^{(1)}/G^{(2)}$ by taking a commutator $[\bar{a}, \bar{b}]$ of elements $\bar{a}, \bar{b} \in G/G^{(2)}$ with representatives $a, b \in G$ to $\overline{[a, b]} \in G^{(1)}/G^{(2)}$. Given different representatives $ag, bg \in G$ for $\bar{a}$ and $\bar{b}$ (where $g \in G^{(2)}$), we have $\overline{[ag, bg]} = \overline{[a, b]}$ since we are working mod $G^{(2)}$, and so this map is well defined.

Conversely, we can define a map $G^{(1)}/G^{(2)} \rightarrow (G/G^{(2)})^{(1)}$ by taking an element $\overline{[a, b]} \in G^{(1)}/G^{(2)}$ where $a, b \in G$ to the commutator $[\bar{a}, \bar{b}]$. Given a different representative $[a, b]g$ for $\overline{[a, b]}$ (where $g \in G^{(2)}$), we see that its image is given by $[\bar{a}, \bar{b}] \times [\bar{x}_i, \bar{y}_i]$, where $x_i, y_i \in G^{(1)}$ since $g \in G^{(2)}$. But, in $G/G^{(2)}$, commutators commute, and so this is equal to $[\bar{a}, \bar{b}]$, and thus the map is well defined.

These two maps are inverses, and so $(G/G^{(2)})^{(1)} \cong G^{(1)}/G^{(2)}$. □

We will also frequently use the fact that any group homomorphism $\varphi: H \rightarrow G$ induces a homomorphism on the metabelianizations $\varphi^{\text{mab}}: H^{\text{mab}} \rightarrow G^{\text{mab}}$ (since $\varphi(H^{(2)}) \subseteq G^{(2)}$). This appears most frequently when H is a subgroup of G and φ is the inclusion map.

We will now recall some fundamental facts about the Alexander modules and Blanchfield forms of satellite knots following [28] (see also [34]).

Theorem 6.8 [28, Theorem 2] *The Alexander matrix (ie a matrix giving a Λ -module presentation for the Alexander module) of a satellite knot corresponding to a Seifert surface obtained from Seifert surfaces for K and J is given by*

$$A_{P(J)}(t) = A_K(t) \oplus A_J(t^w),$$

where $\oplus$ denotes the block sum, and $w = \text{lk}(K, \eta)$ as in Section 5. In particular, the Alexander module is given by

$$\mathcal{A}(P(J)) \cong \mathcal{A}(K) \oplus w\mathcal{A}(J),$$

where, as an abelian group, $w\mathcal{A}(J)$ is the direct sum of w copies of $\mathcal{A}(J)$, but carries the Λ -module structure defined by

$$t \cdot (a_1, a_2, \dots, a_w) = (t \cdot a_w, a_1, \dots, a_{w-1}).$$

Moreover, after choosing suitable bases, the matrix for the Blanchfield form has a similar decomposition:

$$B_{P(J)}(t) = B_K(t) \oplus B_J(t^w).$$

Note that the Λ -module $w\mathcal{A}(J)$ can be seen as the first homology of the cover corresponding to the multiple of the abelianization map $w \cdot \text{ab}: \pi_1(J) \rightarrow \mathbb{Z}$ (taken with $\mathbb{Z}$ coefficients, and coming with a $\mathbb{Z}$ -action giving a Λ -module structure). Note also that, if we restrict to the $\mathbb{Z}[t^w]$ -module structure, then $w\mathcal{A}(J)$ is the direct sum of w copies of $\mathcal{A}(J)$, where t^w acts as the original t -action. Also, the map $w\mathcal{A}(J) \rightarrow \pi_1(J)^{\text{mab}}$ given by the inclusion map on each ($\mathbb{Z}$ -module) $\mathcal{A}(J)$ summand makes the following diagram commute (all maps represented by solid arrows are inclusions or induced by inclusions):

$$\begin{array}{ccc} \mathcal{A}(P(J)) & \longrightarrow & \pi_1(P(J))^{\text{mab}} \\ \uparrow & & \uparrow \\ w\mathcal{A}(J) & \dashrightarrow & \pi_1(J)^{\text{mab}} \end{array}$$

Implicit in the proof in [28] of Theorem 6.8 is that the diagram

$$(*) \quad \begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{A}(K) & \xrightarrow{i} & \pi_1(K)^{\text{mab}} & \xrightarrow{\text{ab}} & \mathbb{Z} \longrightarrow 0 \\ & & \downarrow i & & \downarrow & & \downarrow \text{id} \\ 0 & \longrightarrow & \mathcal{A}(P(J)) & \xrightarrow{i} & \pi_1(P(J))^{\text{mab}} & \xrightarrow{\text{ab}} & \mathbb{Z} \longrightarrow 0 \\ & & \uparrow i & & \uparrow & & \uparrow \text{id} \\ & & w\mathcal{A}(J) & \longrightarrow & \pi_1(J)^{\text{mab}} & \xrightarrow{w \cdot \text{ab}} & \mathbb{Z} \end{array}$$

commutes with exact rows, where the maps labeled with “ i ” are inclusion maps, the maps labeled with “ ab ” are abelianization maps, the vertical maps to $\pi_1(P(J))^{\text{mab}}$ are induced by the inclusion of the appropriate submanifolds, and the map $w\mathcal{A}(J) \rightarrow \pi_1(J)^{\text{mab}}$ is the map described above.

In the special case of a connect sum, we have

$$\mathcal{A}(K \# J) = \mathcal{A}(K) \oplus \mathcal{A}(J),$$

and similarly for the Blanchfield form and the commutative diagram, setting $w = 1$.

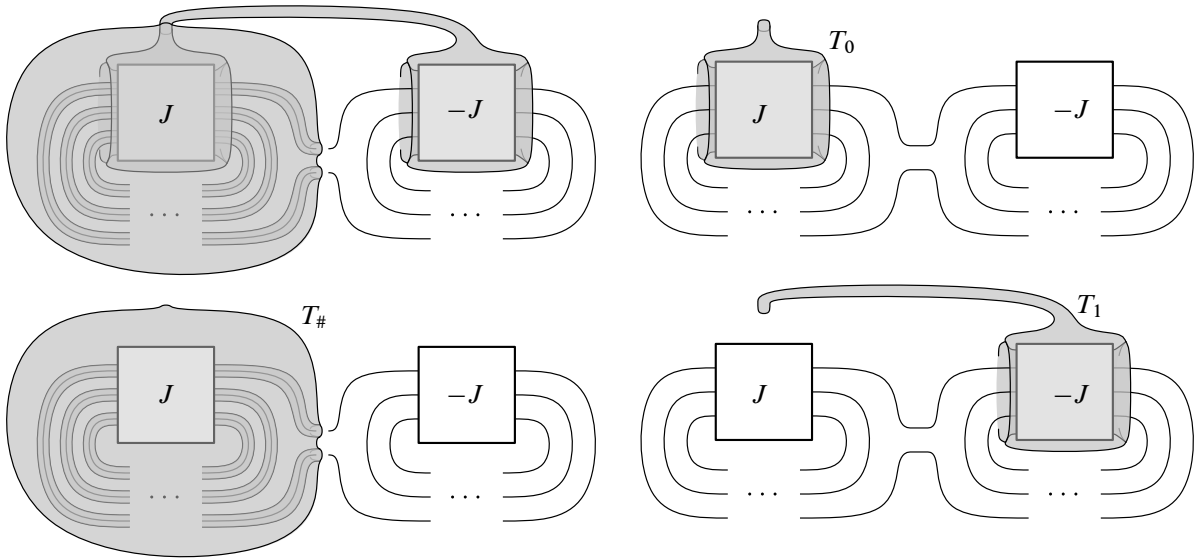


Figure 12: Top left is $P_0(J) \# -P_1(J)$ with the three satellite tori $T_\#$, T_0 and T_1 . The other three diagrams show the tori individually. For the metabelian ρ invariant calculations, take the basepoint to be in the spot common to all three tori.

Proof of Proposition 6.4 We have isomorphisms of Alexander modules

$$\begin{aligned} \mathcal{A}(P_0(J) \# -P_1(J)) &\cong \mathcal{A}(K_0) \oplus w\mathcal{A}(J) \oplus \mathcal{A}(-K_1) \oplus w\mathcal{A}(-J) \\ &\cong \mathcal{A}(K_0 \# -K_1) \oplus w\mathcal{A}(J \# -J) \end{aligned}$$

and the corresponding decompositions of the Blanchfield forms. There are metabolizers of $\mathcal{A}(K_0 \# -K_1)$ and $w\mathcal{A}(J \# -J)$ coming from slice disks, whose direct sum is then a metabolizer for $\mathcal{A}(P_0(J) \# -P_1(J))$. Let $\psi: \pi_1(P_0(J) \# -P_1(J))^{\text{stab}} \rightarrow \Gamma$ be the quotient map killing the metabolizer as above.

We will now use several applications of the satellite formula for metabelian ρ invariants, cutting up S^3 along the three satellite tori $T_\#$, T_0 and T_1 shown in Figure 12. (Note that each torus is isotopic to the torus depicted in Figure 11 with respect to the corresponding satellite operations.) Notice that we will need to work with the fundamental group of the resulting submanifolds, and so want to choose an appropriate basepoint. So, the tori have been chosen to coincide at a particular “spot”, and we choose the basepoint to be in this spot.

From the satellite formula for metabelian ρ invariants, we have

$$\rho(P_0(J) \# -P_1(J), \psi) = \rho(P_0(J), \psi_{P_0(J)}) + \rho(-P_1(J), \psi_{-P_1(J)}),$$

where $\psi_{P_0(J)}$ and $\psi_{-P_1(J)}$ are the homomorphisms obtained by restricting to the submanifolds which result from cutting S^3 along the “swallow-follow” torus $T_\#$ in Figure 12. We have the relation

$$\psi_{P_0(J)}(\mu_{P_0(J)}) = \psi_{-P_1(J)}(\mu_{-P_1(J)})$$

(since these are the same curve on $T_\#$).

After filling in the boundary components as we did in the discussion preceding [Lemma 6.3](#), we can cut along the other two satellite tori, T_0 and T_1 , and use the satellite formula for metabelian ρ invariants again to obtain

$$\rho(P_0(J) \# -P_1(J), \psi) = \rho(K_0, \psi_{K_0}) + \rho(J, \psi_J) + \rho(-K_1, \psi_{-K_1}) + \rho(-J, \psi_{-J})$$

with ψ_{K_0} , ψ_{-K_1} , ψ_J and ψ_{-J} coming from restrictions to submanifolds as before. From our gluings, we must have the identities

$$\psi_{K_0}(\eta_0) = \psi_J(\mu_J), \quad \psi_{K_0}(\mu_{K_0}) = \psi_{-K_1}(\mu_{K_1}), \quad \psi_{-K_1}(-\eta_1) = \psi_{-J}(\mu_{-J}).$$

Also, note that μ_{K_0} and μ_{-K_1} are the same curves as $\mu_{P_0(J)}$ and $\mu_{-P_1(J)}$ from before, and the middle relation follows from the relation coming from the first application of the satellite formula.

Now, since P_0 and P_1 have the same winding numbers, the concatenation $\eta_0 \cdot (-\eta_1)$ is nullhomologous in the complement of $K_0 \# -K_1$, and so lies in $\mathcal{A}(K_0 \# -K_1)$ (recall that $-\eta_1$ denotes the mirror image of η_1 inside the complement of the mirror image of K_1 in the connect sum $K_0 \# -K_1$). Moreover, by assumption, η_0 and η_1 cobound an immersed annulus in the complement of the slice disk for $K_0 \# -K_1$, and so $\eta_0 \cdot (-\eta_1)$ is nullhomotopic in the complement of the slice disk and so lies in the kernel of the inclusion-induced map on $H_1(-, \mathbb{Q}[t^{\pm 1}])$. Therefore, $\psi(\eta_0 \cdot (-\eta_1)) = 1$, and so

$$\psi_{K_0}(\eta_0) = \psi(\eta_0) = \psi(-\eta_1) = \psi_{-K_1}(-\eta_1).$$

Thus, since $\psi_{K_0}(\eta_0) = \psi_J(\mu_J)$ and $\psi_{-K_1}(-\eta_1) = \psi_{-J}(\mu_{-J})$, we must have

$$\psi_J(\mu_J) = \psi_{-J}(\mu_{-J}).$$

Therefore there are maps $\psi_{K_0 \# -K_1}$ and $\psi_{J \# -J}$ on $\pi_1(K_0 \# -K_1)^{\text{mab}}$ and $\pi_1(J \# -J)^{\text{mab}}$, respectively, whose restrictions to the appropriate submanifolds are ψ_{K_0} , ψ_{-K_1} , ψ_J and ψ_{-J} , and so

$$\begin{aligned} \rho(P_0(J) \# -P_1(J)) &= \rho(K_0, \psi_{K_0}) + \rho(J, \psi_J) + \rho(-K_1, \psi_{-K_1}) + \rho(-J, \psi_{-J}) \\ &= \rho(K_0 \# -K_1, \psi_{K_0 \# -K_1}) + \rho(J \# -J, \psi_{J \# -J}). \end{aligned}$$

Putting together the commutative diagrams corresponding to the various satellite operations as in [\(*\)](#), we obtain the commutative diagram shown in [Figure 13](#) (we have omitted the maps ψ_{K_0} , ψ_{-K_1} , ψ_J and ψ_{-J} to avoid cluttering up the diagram too much). The maps of (multiples of) Alexander modules are all inclusions of summands. This diagram commutes by the commutativity of the diagrams corresponding to the various satellite operations as in [\(*\)](#), plus the definitions of ψ , $\psi_{K_0 \# -K_1}$, $\psi_{J \# -J}$, ψ_{K_0} , ψ_{-K_1} , ψ_J and ψ_{-J} . In particular, the two pentagons with black arrows commute, and, since the maps of (multiples of) Alexander modules are inclusions of summands, the kernels of $\psi_{K_0 \# -K_1}$ and $\psi_{J \# -J}$ are exactly the images of our chosen metabolizers of $\mathcal{A}(K_0 \# -K_1)$ and $w\mathcal{A}(J \# -J)$, respectively.

Since these metabolizers came from a slice disk,

$$\begin{aligned} \rho(P_0(J) \# -P_1(J)) &= \rho(K_0, \psi_{K_0}) + \rho(J, \psi_J) + \rho(-K_1, \psi_{-K_1}) + \rho(-J, \psi_{-J}) \\ &= \rho(K_0 \# -K_1, \psi_{K_0 \# -K_1}) + \rho(J \# -J, \psi_{J \# -J}) \\ &= 0 + 0 = 0. \end{aligned}$$

□

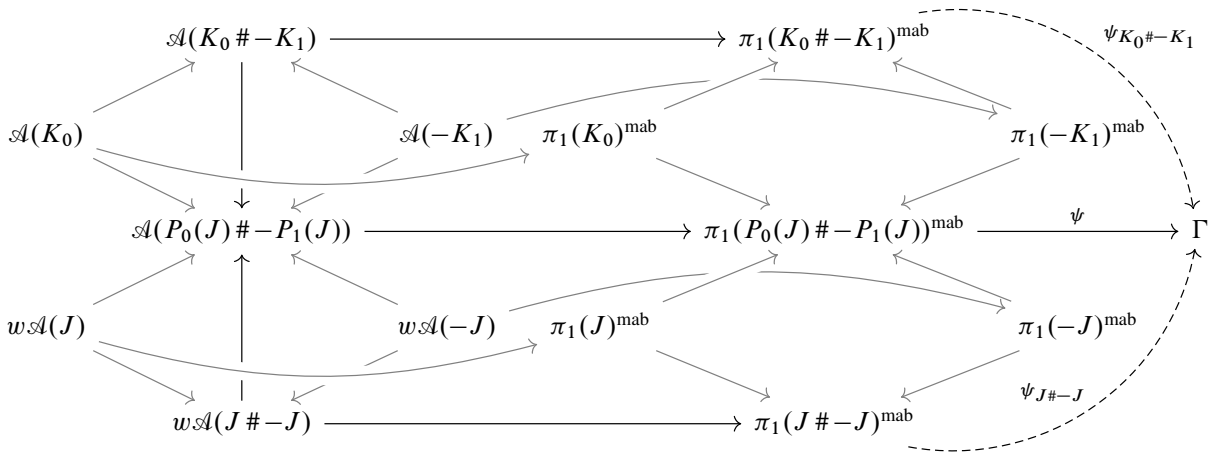


Figure 13

References

- [1] **A J Casson, C M Gordon**, *Cobordism of classical knots*, from “À la recherche de la topologie perdue” (L Guillou, A Marin, editors), Progr. Math. 62, Birkhäuser, Boston, MA (1986) 181–199 [MR](#) [Zbl](#)
- [2] **J Cheeger, M Gromov**, *Bounds on the von Neumann dimension of L^2 -cohomology and the Gauss–Bonnet theorem for open manifolds*, J. Differential Geom. 21 (1985) 1–34 [MR](#) [Zbl](#)
- [3] **T D Cochran, C W Davis, A Ray**, *Injectivity of satellite operators in knot concordance*, J. Topol. 7 (2014) 948–964 [MR](#) [Zbl](#)
- [4] **T D Cochran, B D Franklin, M Hedden, P D Horn**, *Knot concordance and homology cobordism*, Proc. Amer. Math. Soc. 141 (2013) 2193–2208 [MR](#) [Zbl](#)
- [5] **T Cochran, S Friedl, P Teichner**, *New constructions of slice links*, Comment. Math. Helv. 84 (2009) 617–638 [MR](#) [Zbl](#)
- [6] **T Cochran, S Harvey**, *The geometry of the knot concordance space*, Algebr. Geom. Topol. 18 (2018) 2509–2540 [MR](#) [Zbl](#)
- [7] **T D Cochran, S Harvey, P Horn**, *Filtering smooth concordance classes of topologically slice knots*, Geom. Topol. 17 (2013) 2103–2162 [MR](#) [Zbl](#)
- [8] **T D Cochran, S Harvey, C Leidy**, *Knot concordance and higher-order Blanchfield duality*, Geom. Topol. 13 (2009) 1419–1482 [MR](#) [Zbl](#)
- [9] **T D Cochran, S Harvey, C Leidy**, *Derivatives of knots and second-order signatures*, Algebr. Geom. Topol. 10 (2010) 739–787 [MR](#) [Zbl](#)
- [10] **T D Cochran, S Harvey, C Leidy**, *Primary decomposition and the fractal nature of knot concordance*, Math. Ann. 351 (2011) 443–508 [MR](#) [Zbl](#)
- [11] **T D Cochran, S Harvey, M Powell**, *Grope metrics on the knot concordance set*, J. Topol. 10 (2017) 669–699 [MR](#) [Zbl](#)
- [12] **T D Cochran, K E Orr, P Teichner**, *Knot concordance, Whitney towers and L^2 -signatures*, Ann. of Math. 157 (2003) 433–519 [MR](#) [Zbl](#)
- [13] **M M Cohen**, *A course in simple-homotopy theory*, Graduate Texts in Math. Vol. 10, Springer (1973) [MR](#) [Zbl](#)

- [14] **S Collins**, *Homology cobordism, smooth concordance, and the figure eight knot*, preprint (2022) [arXiv 2210.10308](#)
- [15] **A Conway**, *Homotopy ribbon discs with a fixed group*, *Algebr. Geom. Topol.* 24 (2024) 4575–4587 [MR](#) [Zbl](#)
- [16] **C W Davis**, **A Ray**, *Satellite operators as group actions on knot concordance*, *Algebr. Geom. Topol.* 16 (2016) 945–969 [MR](#) [Zbl](#)
- [17] **M H Freedman**, *A new technique for the link slice problem*, *Invent. Math.* 80 (1985) 453–465 [MR](#) [Zbl](#)
- [18] **M H Freedman**, **F Quinn**, *Topology of 4-manifolds*, Princeton Mathematical Series 39, Princeton Univ. Press (1990) [MR](#) [Zbl](#)
- [19] **M H Freedman**, **P Teichner**, *4-manifold topology, II: Dwyer’s filtration and surgery kernels*, *Invent. Math.* 122 (1995) 531–557 [MR](#) [Zbl](#)
- [20] **S Friedl**, **P Teichner**, *New topologically slice knots*, *Geom. Topol.* 9 (2005) 2129–2158 [MR](#) [Zbl](#)
- [21] **R E Gompf**, **K Miyazaki**, *Some well-disguised ribbon knots*, *Topology Appl.* 64 (1995) 117–131 [MR](#) [Zbl](#)
- [22] **R E Gompf**, **A I Stipsicz**, *4-manifolds and Kirby calculus*, Graduate Studies in Math. 20, Amer. Math. Soc., Providence, RI (1999) [MR](#) [Zbl](#)
- [23] **M Hedden**, **J Pinzón-Caicedo**, *Satellites of infinite rank in the smooth concordance group*, *Invent. Math.* 225 (2021) 131–157 [MR](#) [Zbl](#)
- [24] **A S Levine**, *Nonsurjective satellite operators and piecewise-linear concordance*, *Forum Math. Sigma* 4 (2016) art. id. e34 [MR](#) [Zbl](#)
- [25] **T Lidman**, **A N Miller**, **J Pinzón-Caicedo**, *Linking number obstructions to satellite homomorphisms*, *Quantum Topol.* (online publication July 2024)
- [26] **R A Litherland**, *Cobordism of satellite knots*, from “Four-manifold theory” (C Gordon, R Kirby, editors), *Contemp. Math.* 35, Amer. Math. Soc., Providence, RI (1984) 327–362 [MR](#) [Zbl](#)
- [27] **R A Litherland**, *Symmetries of twist-spun knots*, from “Knot theory and manifolds” (D Rolfsen, editor), *Lecture Notes in Math.* 1144, Springer (1985) 97–107 [MR](#) [Zbl](#)
- [28] **C Livingston**, **P Melvin**, *Abelian invariants of satellite knots*, from “Geometry and topology” (J Alexander, J Harer, editors), *Lecture Notes in Math.* 1167, Springer (1985) 217–227 [MR](#) [Zbl](#)
- [29] **C Livingston**, **S Naik**, *Introduction to knot concordance*, book project (2011)
- [30] **A N Miller**, *Homomorphism obstructions for satellite maps*, *Trans. Amer. Math. Soc. Ser. B* 10 (2023) 220–240 [MR](#) [Zbl](#)
- [31] **A N Miller**, **L Piccirillo**, *Knot traces and concordance*, *J. Topol.* 11 (2018) 201–220 [MR](#) [Zbl](#)
- [32] **J Milnor**, *Whitehead torsion*, *Bull. Amer. Math. Soc.* 72 (1966) 358–426 [MR](#) [Zbl](#)
- [33] **A Ray**, *Satellite operators with distinct iterates in smooth concordance*, *Proc. Amer. Math. Soc.* 143 (2015) 5005–5020 [MR](#) [Zbl](#)
- [34] **H Seifert**, *On the homology invariants of knots*, *Quart. J. Math. Oxford Ser.* 1 (1950) 23–32 [MR](#) [Zbl](#)
- [35] **K Yasui**, *Corks, exotic 4-manifolds and knot concordance*, preprint (2015) [arXiv 1505.02551](#)

Department of Mathematics, Rice University
Houston, TX, United States

alex.manchester@rice.edu

<https://sites.google.com/view/alexmanchester/home>

Received: 27 January 2023 Revised: 19 October 2023

Kauffman bracket intertwiners and the volume conjecture

ZHIHAO WANG

The volume conjecture relates the quantum invariant and the hyperbolic geometry. Bonahon, Wong and Yang introduced a new version of the volume conjecture by using the intertwiners between two isomorphic irreducible representations of the skein algebra. The intertwiners are built from surface diffeomorphisms; they formulated the volume conjecture when diffeomorphisms are pseudo-Anosov. We explicitly calculate all the intertwiners for the closed torus using an algebraic embedding from the skein algebra of the closed torus to a quantum torus, and show the limit superior related to the trace of these intertwiners is zero. Moreover, we consider the periodic diffeomorphisms for surfaces with negative Euler characteristic, and conjecture the corresponding limit is zero because the simplicial volume of the mapping tori for periodic diffeomorphisms is zero. For the once punctured torus, we make precise calculations for intertwiners and prove our conjecture.

14H10, 14H30, 14H45, 14H50, 14L10

1 Introduction

We first discuss irreducible representations for skein algebras of the closed torus and the once punctured torus, which is related to Bonahon and Wong's work [3; 4; 5; 6]. They explored the connection between irreducible representations of skein algebras and the character variety related to the fundamental group of a surface. In [Section 3](#), we give more detailed discussions about this connection for the closed torus and the once punctured torus.

A profound result of the skein algebra is the unicity theorem, which was conjectured by Bonahon and Wong and was proved by Frohman, Kania-Bartoszyńska and Lê [15]. Based on this result there is an increased focus on the Azumaya locus. Ganev, Jordan and Safronov proved that the smooth part of the character variety lives in the Azumaya locus when the surface is closed [16]. In [Section 3](#), we give an explicit description for the Azumaya locus for the skein algebra of the closed torus.

Let S be an oriented surface, let φ be a diffeomorphism for S , and let $q_n = e^{2\pi i/n}$ with $(q_n)^{1/2} = e^{\pi i/n}$ and n odd. Using these data, Bonahon, Wong and Yang built a sequence of intertwiners between irreducible representations of the skein algebra of S [7; 8]. When S has negative Euler characteristic and φ is pseudo-Anosov, they formulated the volume conjecture using these intertwiners:

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, \gamma}^{q_n}| = \frac{1}{4\pi} \text{vol}_{\text{hyp}}(M_{\varphi}),$$

where $\text{vol}_{\text{hyp}}(M_{\varphi})$ is the volume of the complete hyperbolic metric of the mapping torus M_{φ} .

We explicitly compute the intertwiners corresponding to all diffeomorphisms of the closed torus using an algebraic embedding from the skein algebra of the closed torus to a quantum torus [14]; see Section 3 for more details. The representation theory for this quantum torus is well studied. We prove almost all the irreducible representations of this quantum torus can be restricted to irreducible representations of the skein algebra of the closed torus. So intertwiners between two isomorphic irreducible representations of this quantum torus are also the intertwiners between irreducible representations for the skein algebra of the closed torus. These intertwiners are built when the quantum parameter q for the skein algebra is a primitive root of unity of odd order. We use Λ_n to denote the intertwiner obtained as above when the quantum parameter is $q_n = e^{2\pi i/n}$ with $(q_n)^{1/2} = e^{\pi i/n}$ and n odd. We also normalize Λ_n such that $|\det(\Lambda_n)| = 1$. Then we prove the following theorem; please refer to Theorem 4.15 for a more detailed version.

Theorem 1.1 *Let Λ_n be defined as above; then*

$$\limsup_{\text{odd } n \rightarrow \infty} \frac{\log(|\text{Trace } \Lambda_n|)}{n} = 0.$$

The volume conjecture was first introduced by Kashaev [19], and was rewritten and generalized to the nonhyperbolic case by Hitoshi Murakami and Jun Murakami [22] using the simplicial volume.

Bonahon, Wong and Yang only formulated the conjecture when the diffeomorphisms are pseudo-Anosov for surfaces with negative Euler characteristic. In this paper, we broaden the scope of the conjecture to include periodic diffeomorphisms. When φ is a periodic diffeomorphism for the surface S , the corresponding mapping torus M_φ is a Seifert manifold whose simplicial volume is zero. So we conjecture the limits are zero for periodic diffeomorphisms. We prove our conjecture for the once punctured torus, which serves as an example to confirm the limit is the simplicial volume of the corresponding mapping torus.

Let S be an oriented surface with negative Euler characteristic, and let φ be a periodic diffeomorphism for S . According to [13, page 371], φ fixes a point in the Teichmüller space of S . This fixed point in the Teichmüller space offers a smooth φ -invariant character γ (that is γ is a group homomorphism from $\pi_1(S)$ to $\text{SL}(2, \mathbb{C})$ such that $\gamma\varphi_*$ and γ have the same character, where φ_* is the isomorphism from $\pi_1(S)$ to $\pi_1(S)$ induced by φ). Suppose the quantum parameter for the skein algebra is $q_n = e^{2\pi i/n}$ with $(q_n)^{1/2} = e^{\pi i/n}$ and n odd. For each puncture v of S , we choose a complex number p_v such that $p_v = p_{\varphi(v)}$ and $T_n(p_v) = -\text{Trace}(\gamma(\alpha_v))$, where T_n is the n^{th} Chebyshev polynomial of the first type and α_v is the element in $\pi_1(S)$ going around puncture v . According to Theorem 2.1, we know γ and p_v uniquely determine an irreducible representation ρ of the skein algebra. Let $\varphi_\#$ be the isomorphism from the skein algebra of S to itself induced by φ . Since both γ and p_v are φ -invariant, we have ρ and $\rho\varphi_\#$ are isomorphic according to Theorem 2.2. Thus there exists the intertwiner $\Lambda_{\varphi, \gamma}^{q_n}$ between these two isomorphic irreducible representations. We normalize it such that $|\det(\Lambda_{\varphi, \gamma}^{q_n})| = 1$. Then we formulate the following conjecture, please refer to Conjecture 5.4 for a more detailed version.

Conjecture 1.2 *Let S be a surface with negative Euler characteristic, let φ be a periodic diffeomorphism for S , and let $\Lambda_{\varphi, \gamma}^{q_n}$ be defined as above; then*

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, \gamma}^{q_n}| = 0.$$

In [Theorem 5.5](#), we prove the limit in [Conjecture 1.2](#) is less than or equal to zero if it exists by using the periodic property. It seems like we are halfway there to prove our conjecture. But proving that the limit is greater than or equal to zero is harder, which is actually related to an interesting question raised by Gerald Myerson [\[24\]](#) and Terry Tao [\[27\]](#). By direct calculations and using some conclusions in [\[20; 24\]](#), we prove the above conjecture for some special cases:

Theorem 1.3 *For any surface with negative Euler characteristic, if φ is of order 2^m where m is any positive integer, then*

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, \gamma_\varphi}^{q_n}| = 0.$$

Theorem 1.4 *[Conjecture 1.2](#) holds if S is the once punctured torus.*

Plan of the paper: In [Section 2](#), we introduce the Kauffman bracket skein algebra, the classical shadow, the volume conjecture, and the Chekhov–Fork algebra. [Section 3](#) is about the discussion on the irreducible representations of skein algebras of the closed torus and the once punctured torus. In [Section 4](#), we calculate the intertwiners for the closed torus and prove [Theorem 1.1](#). In [Section 5](#), we formulate our conjecture for periodic diffeomorphisms and prove [Theorems 1.3](#) and [1.4](#).

Acknowledgements The idea of considering the periodic case was suggested by my supervisor Andrew James Kricker. We would like to thank Andrew James Kricker, Jeffrey Weenink, Roland van der Veen and Xiaoming Yu for constructive discussion and help. We wish to thank the referee most warmly for numerous suggestions that have improved the exposition of this paper. The research is supported by the NTU research scholarship.

2 Preliminaries

2.1 The $\text{SL}(2, \mathbb{C})$ character variety and the Kauffman bracket skein algebra

Let S be an oriented surface of finite type. The corresponding character variety

$$\mathcal{X}_{\text{SL}(2, \mathbb{C})}(S) = \text{Hom}(\pi_1(S), \text{SL}(2, \mathbb{C})) // \text{SL}(2, \mathbb{C})$$

is the set of the group homomorphisms from the fundamental group of S to $\text{SL}(2, \mathbb{C})$ with the equivalence relation that two homomorphisms are equivalent if and only if they have the same character [\[9; 12; 25\]](#).

The Kauffman bracket skein algebra $\text{SK}_{q^{1/2}}(S)$ of a surface S , as a vector space over the complex field $\mathbb{C}$, is generated by all isotopic framed links in $S \times [0, 1]$, subject to the skein relation

$$K_1 = q^{-1/2} K_\infty + q^{1/2} K_0,$$

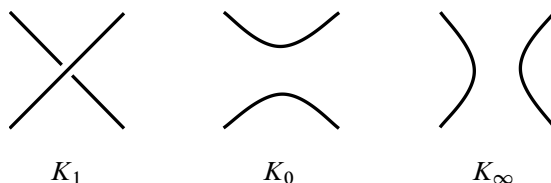


Figure 1: The Kauffman bracket skein relation.

where K_1 , K_∞ and K_0 are three links that differ in a small neighborhood as shown in Figure 1, and the trivial knot relation $K \sqcup \bigcirc = -(q + q^{-1})K$, where $\bigcirc$ is a simple knot bounding a disk that has no intersection with K . For any two links $[L_1]$ and $[L_2]$, the multiplication $[L_1][L_2]$ is defined by stacking L_2 above L_1 . Here $q^{1/2}$ is a nonzero complex number. The skein algebra $\text{SK}_{q^{1/2}}(S)$ is a quantization for the regular ring of the character variety $\mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$ [9].

2.2 Classical shadow and unicity theorem

We recall some notation and constructions for the classical shadow [3]. When q is a primitive n^{th} root of unity with n odd and $(q^{1/2})^n = -1$, Bonahon and Wong found a fascinating algebra homomorphism $T^{q^{1/2}}$ from $\text{SK}_{-1}(S)$ to $\text{SK}_{q^{1/2}}(S)$, called the Chebyshev homomorphism. Bonahon and Wong proved that $\text{Im}(T^{q^{1/2}})$ is contained in the center of $\text{SK}_{q^{1/2}}(S)$. If K is a simple knot with vertical framing, then $T^{q^{1/2}}([K]) = T_n([K])$ where T_n is the n^{th} Chebyshev polynomial of the first type.

Let $\rho: \text{SK}_{q^{1/2}}(S) \rightarrow \text{End}(V)$ be an irreducible representation of $\text{SK}_{q^{1/2}}(S)$. Then there exists an algebra homomorphism κ_ρ from $\text{SK}_{-1}(S)$ to $\mathbb{C}$ such that $\rho \circ T^{q^{1/2}}(X) = \kappa_\rho(X) \text{Id}_V$ for any X in $\text{SK}_{-1}(S)$. According to [9], there exists a unique character $[\gamma] \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$ such that $\text{Tr}^\gamma = \kappa_\rho$. Recall that Tr^γ is an algebra homomorphism from $\text{SK}_{-1}(S)$ to $\mathbb{C}$ defined by $\text{Tr}^\gamma([K]) = -\text{Trace } \gamma(K)$ where $[K]$ is a simple knot. For every puncture v , we use P_v to denote the loop going around this puncture. There is a complex number p_v such that $\rho([P_v]) = p_v \text{Id}_V$. Then an irreducible representation of $\text{SK}_{q^{1/2}}(S)$ gives a character $[\gamma]$, called the classical shadow of this irreducible representation, and puncture weights $\{p_v\}_v$, with the relation that $-\text{Trace } \gamma(\alpha_v) = T_n(p_v)$ where α_v denotes the element in the fundamental group of S going around the puncture v .

Theorem 2.1 [3; 5; 6; 7] *Let q be a primitive n^{th} root of unity with n odd and $(q^{1/2})^n = -1$. Then an irreducible representation $\rho: \text{SK}_{q^{1/2}} \rightarrow \text{End}(V)$ uniquely determines*

- (1) a character $[\gamma] \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$, represented by a group homomorphism $\gamma: \pi_1(S) \rightarrow \text{SL}(2, \mathbb{C})$;
- (2) a weight p_v associated to each puncture v of S such that $T_n(p_v) = -\text{Trace } \gamma(\alpha_v)$.

Conversely, every data of a character $\gamma \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$ and of puncture weights $p_v \in \mathbb{C}$ satisfying the above condition is realized by an irreducible representation $\rho: \text{SK}_{q^{1/2}}(S) \rightarrow \text{End}(V)$.

It turns out that every character in an open dense subset of $\mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$ corresponds to a unique irreducible representation of the skein algebra.

Theorem 2.2 [7; 15; 16] Suppose that $[\gamma]$ is in the smooth part of $\mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(S)$ or, equivalently, that it is realized by an irreducible homomorphism $\gamma: \pi_1(S) \rightarrow \mathrm{SL}(2, \mathbb{C})$. Then the irreducible representation $\rho: \mathrm{SK}_{q^{1/2}}(S) \rightarrow \mathrm{End}(V)$ in Theorem 2.1 is unique up to isomorphism of representations. This representation has dimension $\dim V = n^{3g+p-3}$ if S has genus g and p punctures.

2.3 Volume conjecture for surface diffeomorphisms

Bonahon, Wong and Yang constructed the so called Kauffman bracket intertwiners [7; 8]. They used these intertwiners to formulate the volume conjecture for surface diffeomorphisms. Here we recall their construction for Kauffman bracket intertwiners.

For a surface S , let φ be a diffeomorphism of S . Obviously φ induces an isomorphism φ_* from $\pi_1(S)$ to $\pi_1(S)$. Then φ_* induces an action on $\mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(S)$ defined by $\varphi^*([\gamma]) = [\gamma\varphi_*]$ where γ is a representative for $[\gamma]$. Although φ_* is only defined up to conjugation, φ^* is well defined. Actually the mapping class group $\mathrm{Mod}(S)$ acts on $\mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(S)$. We say an element $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(S)$ is invariant under a diffeomorphism φ , or the element it represents in $\mathrm{Mod}(S)$, if $\varphi^*([\gamma]) = [\gamma]$.

The algebra isomorphism induced by φ from $\mathrm{SK}_{q^{1/2}}(S)$ to itself is defined by $\varphi_{\#}([K]) = [\varphi \times \mathrm{Id}_{[0,1]}(K)]$ where K is a framed link in $S \times [0, 1]$. Actually the mapping class group $\mathrm{Mod}(S)$ acts on $\mathrm{SK}_{q^{1/2}}(S)$.

Let φ be any diffeomorphism for surface S , and let $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(S)$ be a φ -invariant smooth character. For each puncture v , select a complex number θ_v such that $\mathrm{Trace} \gamma(\alpha_v) = -e^{\theta_v} - e^{-\theta_v}$. Since $[\gamma]$ is φ -invariant, we can choose θ_v to be φ -invariant, that is, $\theta_v = \theta_{\varphi(v)}$. Then set $p_v = e^{\theta_v/n} + e^{-\theta_v/n}$; we have that $T_n(p_v) = -\mathrm{Trace}(\alpha_v)$ and $\{p_v\}_v$ are invariant under the action of φ . Suppose ρ is an irreducible representation associated to $[\gamma]$ and puncture weights p_v . Then $\rho \circ \varphi_{\#}$ is also an irreducible representation associated to $[\gamma]$ and puncture weights p_v . By the unicity theorem, we know there exists an intertwiner $\Lambda_{\varphi, \gamma}^q$ such that

$$\rho \circ \varphi_{\#}(X) = \Lambda_{\varphi, \gamma}^q \circ \rho(X) \circ (\Lambda_{\varphi, \gamma}^q)^{-1}$$

for every $X \in \mathrm{SK}_{q^{1/2}}(S)$. We normalize the intertwiner such that $|\det(\Lambda_{\varphi, \gamma}^q)| = 1$.

Conjecture 2.3 [7; 8] Let the pseudo-Anosov surface diffeomorphism $\varphi: S \rightarrow S$, the φ -invariant smooth character $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(S)$ and the φ -invariant puncture weights p_v as above be given. For every odd n , consider the primitive n^{th} root of unity $q_n = e^{2\pi i/n}$ and choose $(q_n)^{1/2} = e^{\pi i/n}$. Then

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\mathrm{Trace} \Lambda_{\varphi, \gamma}^{q_n}| = \frac{1}{4\pi} \mathrm{vol}_{\mathrm{hyp}}(M_{\varphi}),$$

where $\mathrm{vol}_{\mathrm{hyp}}(M_{\varphi})$ is the volume of the complete hyperbolic metric of the mapping torus M_{φ} .

2.4 Ideal triangulation and intertwiners obtained from Chekhov–Fock algebras

Let S be an oriented surface with punctures, and let $\tau = \{e_1, \dots, e_m\}$ be an ideal triangulation for S , where $e_1, \dots, e_m$ are nonisotopic disjoint embedded arcs in S connecting punctures such that all these arcs cut S into triangles. We call $e_1, \dots, e_m$ the edges of τ . An edge weight system for τ is an m -tuple,

$a = (a_1, \dots, a_m)$, where a_i is a nonzero complex number for each $1 \leq i \leq m$. The pair (τ, a) determines a character $[\bar{\gamma}]$ in $\mathcal{X}_{\text{PSL}(2, \mathbb{C})}(S)$; please refer to [1, Section 8] or [7, Section 3] for more details.

For each ideal triangulation τ , there is a Chekhov–Fock algebra $\mathcal{T}_\tau^q$ corresponding to τ , where q is a nonzero complex number. As an algebra over $\mathbb{C}$, the Chekhov–Fock algebra $\mathcal{T}_\tau^q$ is generated by $X_1^{\pm 1}, X_2^{\pm 1}, \dots, X_m^{\pm 1}$ subject to the relations

$$X_i X_i^{-1} = X_i^{-1} X_i = 1, \quad X_i X_j = q^{2\sigma_{ij}} X_j X_i.$$

Each X_i is associated to the i^{th} edge in the ideal triangulation τ , and σ_{ij} is an integer determined by τ ; see [1; 2; 21] for more details. If we replace q with $q^{1/4}$, we get the so called Chekhov–Fock square root algebra $\mathcal{T}_\tau^{q^{1/4}}$. It is well known that $\mathcal{T}_\tau^q$ is an Ore domain. We will use $\widehat{\mathcal{T}}_\tau^q$ to denote the ring of fractions of $\mathcal{T}_\tau^q$ (that is the localization over all nonzero elements).

Let τ and τ' be any two ideal triangulations for S . Then there is an algebra isomorphism $\Phi_{\tau\tau'}^q: \widehat{\mathcal{T}}_{\tau'}^q \rightarrow \widehat{\mathcal{T}}_\tau^q$, called be the Chekhov–Fock coordinate change isomorphism [21].

For an ideal triangulation τ , there are two operations.

- (1) **Reindexing** Obtain a new ideal triangulation τ' by reindexing all the edges in τ .
- (2) **Diagonal exchange** For any $1 \leq i \leq m$, define a new ideal triangulation $\tau' = \{e'_1, \dots, e'_m\}$, where $e'_j = e_j$ for every $j \neq i$ and e'_i is the other diagonal of the square formed by the two faces of τ that are adjacent to e_i .

Let τ be an ideal triangulation, let $a = (a_1, \dots, a_m)$ be an edge weight system for τ . Suppose $\tau' = \{e'_1, \dots, e'_m\}$ is obtained from τ by reindexing such that $e'_i = e_{\sigma(i)}$ for $1 \leq i \leq m$, where σ is a permutation for $\{1, \dots, m\}$. Then we define an edge weight system $a' = (a'_1, \dots, a'_m)$ for τ' by setting $a'_i = a_{\sigma(i)}$ for $1 \leq i \leq m$. If τ' is obtained from τ by the diagonal exchange, we define an edge weight system a' for τ' using formulas in [21, Proposition 3]. We will say a' is an edge weight system for τ' derived from the pair (τ, a) . Then (τ', a') determines the same character in $\mathcal{X}_{\text{PSL}(2, \mathbb{C})}(S)$ as (τ, a) [1; 7].

A sequence of ideal triangulations $\tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)}$ is called an ideal triangulation sweep if, for each $1 \leq i \leq k-1$, we have that $\tau^{(i+1)}$ is obtained from $\tau^{(i)}$ by reindexing or the diagonal exchange. A sequence of edge weight systems $a^{(0)}, a^{(1)}, \dots, a^{(k)}$ is called an edge weight system sweep for the ideal triangulation sweep $\tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)}$, if the edge weight system $a^{(i+1)}$ for $\tau^{(i+1)}$ is derived from $(\tau^{(i)}, a^{(i)})$ for each $0 \leq i \leq k-1$. Note that the sequence $a^{(0)}, a^{(1)}, \dots, a^{(k)}$ is completely determined by $a^{(0)}$. If in addition $a^{(0)} = a^{(k)}$, we call the sequence $a^{(0)}, a^{(1)}, \dots, a^{(k)}$ a periodic edge weight system for the ideal triangulation sweep $\tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)}$.

Suppose q is a primitive n^{th} root of unity with n odd. Let φ be an orientation preserving diffeomorphism for surface S , and let $\tau = \tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)} = \varphi(\tau)$ be an ideal triangulation sweep. Suppose $a = a^{(0)}, a^{(1)}, \dots, a^{(k)} = a$ is a periodic edge weight system for $\tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)}$ (the existence of the periodic edge weight system is guaranteed by [7, Lemma 11]), which defines a φ -invariant character

$[\bar{\gamma}] \in \mathcal{X}_{\text{PSL}(2, \mathbb{C})}(S)$. Then, for each puncture v , we can choose a nonzero complex number h_v such that $h_v = h_{\varphi(v)}$ for every puncture v and $(h_v)^n = a_{i_1} a_{i_2} \cdots a_{i_j}$ for every puncture v adjacent to the edges $e_{i_1}, e_{i_2}, \dots, e_{i_j}$. From [7, Proposition 13], we know a and puncture weights h_v uniquely determine an irreducible representation $\bar{\rho}: \mathcal{T}_\tau^q \rightarrow \text{End}(V)$ for the Chekhov–Fock algebra $\mathcal{T}_\tau^q$ such that $\bar{\rho}(X_i^n) = a_i$ for $1 \leq i \leq m$ and $\bar{\rho}(H_v) = h_v$ for each puncture v , where H_v is a central element in $\mathcal{T}_\tau^q$ associated to each puncture v . Let $\Phi_{\tau\varphi(\tau)}^q: \hat{\mathcal{T}}_{\varphi(\tau)}^q \rightarrow \hat{\mathcal{T}}_\tau^q$ be the Chekhov–Fock coordinate change isomorphism, and let $\Psi_{\varphi(\tau)\tau}^q: \hat{\mathcal{T}}_\tau^q \rightarrow \hat{\mathcal{T}}_{\varphi(\tau)}^q$ be the algebra isomorphism induced by φ . Then $\bar{\rho} \simeq \bar{\rho} \circ \Phi_{\tau\varphi(\tau)}^q \circ \Psi_{\varphi(\tau)\tau}^q$, so there exists an intertwiner $\bar{\Lambda}_{\varphi, \bar{\gamma}}^q$ such that

$$\bar{\rho} \circ \Phi_{\tau\varphi(\tau)}^q \circ \Psi_{\varphi(\tau)\tau}^q(X) = \bar{\Lambda}_{\varphi, \bar{\gamma}}^q \circ \bar{\rho}(X) \circ (\bar{\Lambda}_{\varphi, \bar{\gamma}}^q)^{-1}$$

for every $X \in \mathcal{T}_\tau^q$.

Under certain conditions, the trace of intertwiners obtained from Chekhov–Fock algebras equals the trace of intertwiners obtained from skein algebras; see [7, Theorem 16]. We will use this equality to calculate the trace of intertwiners obtained from skein algebras for the once punctured torus in Section 5.

From now on, we will always assume $q^{1/2}$ is a primitive n^{th} root of -1 with n odd.

3 Irreducible representation construction for $\text{SK}_{q^{1/2}}(T^2)$ and $\text{SK}_{q^{1/2}}(S_{1,1})$

In order to get Kauffman bracket intertwiners, we want to find the explicit irreducible representations associated to given characters and puncture weights. Here we construct irreducible representations for skein algebras of the closed torus T^2 and the once punctured torus $S_{1,1}$. In Section 4, we will use these irreducible representations to calculate intertwiners for the closed torus.

3.1 An algebraic embedding for $\text{SK}_{q^{1/2}}(T^2)$

Let $\mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$ be the algebra generated by X , X^{-1} , Y and Y^{-1} , subject to the relations $XY = qYX$, $XX^{-1} = X^{-1}X = 1$ and $YY^{-1} = Y^{-1}Y = 1$. Frohman and Gelca [14] built an algebraic embedding

$$G_{q^{1/2}}: \text{SK}_{q^{1/2}}(T^2) \rightarrow \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}, \quad (a, b)_T \mapsto \theta_{(a,b)} + \theta_{(-a,-b)},$$

where $(a, b)_T$ is the simple link associated to two integers a and b , and $\theta_{(a,b)} = q^{-ab/2} X^a Y^b$. If $\gcd(a, b) = 1$ (with the convention that $\gcd(\pm 1, 0) = \gcd(0, \pm 1) = 1$), then $(a, b)_T$ is represented by the simple knot (a, b) in $\mathbb{R}^2/\mathbb{Z}^2$ with vertical framing. If $\gcd(a, b) = k$, $a = a'k$ and $b = b'k$, then $(a, b)_T = T_k((a', b'))$ where T_k is the k^{th} Chebyshev polynomial of the first type. We have

$$\theta_{(a,b)} \theta_{(c,d)} = q^{1/2} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \theta_{(a+c, b+d)} \quad \text{and} \quad (\theta_{(a,b)})^{-1} = \theta_{(-a,-b)}.$$

Since $\theta_{(a,b)} + \theta_{(-a,-b)} = q^{-ab/2} (X^a Y^b + X^{-a} Y^{-b})$,

$$\text{Im } G_{q^{1/2}} = \text{span} \langle X^a Y^b + X^{-a} Y^{-b} \mid (a, b) \in \mathbb{Z} \times \mathbb{Z} \rangle.$$

Let $T_{q^{1/2}}$ be the Chebyshev homomorphism from the skein algebra $\mathrm{SK}_{-1}(T^2)$ to $\mathrm{SK}_{q^{1/2}}(T^2)$ defined in [3], and let $F_{q^{1/2}}: \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{-1} \rightarrow \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$ be defined by $X \mapsto X^n$ and $Y \mapsto Y^n$. It is easy to check that we have $F_{q^{1/2}}G_{-1} = G_{q^{1/2}}T_{q^{1/2}}$.

3.2 Irreducible representations for $\mathrm{SK}_{q^{1/2}}(T^2)$

Bonahon and Liu described the irreducible representations of $\mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$ [1]. Let V denote the n -dimensional vector space over the complex field with basis $e_0, e_1, \dots, e_{n-1}$, and let u and v be any two nonzero complex numbers. Set $\rho_{u,v}(X)e_i = uq^i e_i$ and $\rho_{u,v}(Y)e_i = ve_{i+1}$, where the indices are considered modulo n ; then $\rho_{u,v}$ is an irreducible representation. Any irreducible representation of $\mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$ is isomorphic to a representation $\rho_{u,v}$, and $\rho_{u,v} \simeq \rho_{u',v'}$ if and only if $u^n = (u')^n$ and $v^n = (v')^n$.

It is well known that $\pi_1(T^2) = \mathbb{Z} \oplus \mathbb{Z} = \mathbb{Z}\alpha \oplus \mathbb{Z}\beta$ where $\alpha = (1, 0)$ and $\beta = (0, 1)$. For any $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(T^2)$, $[\gamma]$ has a representative γ such that

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}$$

because $\pi_1(T^2)$ is commutative.

For any given character $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(T^2)$, the following theorem offers a representation of $\mathrm{SK}_{q^{1/2}}(T^2)$ whose classical shadow is $[\gamma]$. For this theorem, we use the fact that $ab + a + b \equiv \gcd(a, b) \pmod{2}$ for any two integers a and b (recall that $\gcd(\pm 1, 0) = \gcd(0, \pm 1) = 1$).

Theorem 3.1 Choose $u, v \in \mathbb{C}$ such that $u^n = -\lambda_1$, $v^n = -\lambda_2$ or $u^n = -\lambda_1^{-1}$, $v^n = -\lambda_2^{-1}$; then the classical shadow of $\rho_{u,v}G_{q^{1/2}}$ is $[\gamma]$.

Proof To show the classical shadow of $\rho_{u,v}G_{q^{1/2}}$ is $[\gamma]$, it suffices to show that

$$\rho_{u,v}G_{q^{1/2}}(T_{q^{1/2}}((a, b)_T)) = \mathrm{Tr}^\gamma((a, b)_T)\mathrm{Id}_V$$

for all $(a, b)_T \in \mathrm{SK}_{-1}(T^2)$. First we have

$$\begin{aligned} \rho_{u,v}G_{q^{1/2}}(T_{q^{1/2}}((a, b)_T)) &= \rho_{u,v}(F_{q^{1/2}}G_{-1}((a, b)_T)) \\ &= \rho_{u,v}(F_{q^{1/2}}(\theta_{(a,b)} + \theta_{(-a,-b)})) \\ &= \rho_{u,v}(\theta_{(na,nb)} + \theta_{(-na,-nb)}) \\ &= \rho_{u,v}((-1)^{ab}X^{na}Y^{nb} + (-1)^{ab}X^{-na}Y^{-nb}) \\ &= (-1)^{ab}[(\rho_{u,v}(X))^{na}(\rho_{u,v}(Y))^{nb} + (\rho_{u,v}(X))^{-na}(\rho_{u,v}(Y))^{-nb}] \\ &= (-1)^{ab}[(u^n)^a(v^n)^b + (u^n)^{-a}(v^n)^{-b}]\mathrm{Id}_V \\ &= (-1)^{ab+a+b}[\lambda_1^a\lambda_2^b + \lambda_1^{-a}\lambda_2^{-b}]\mathrm{Id}_V. \end{aligned}$$

Suppose $\gcd(a, b) = d$ and $a = a'd, b = b'd$; then we have

$$\begin{aligned}
 \mathrm{Tr}^\gamma((a, b)_T) &= \mathrm{Tr}^\gamma(T_d((a', b')) = T_d(\mathrm{Tr}^\gamma((a', b'))) \\
 &= T_d(-\mathrm{Trace}(\gamma((a', b')))) \\
 &= T_d(-\mathrm{Trace}(\gamma(a'\alpha + b'\beta))) \\
 &= T_d(-\mathrm{Trace}((\gamma(\alpha))^{a'}(\gamma(\beta))^{b'})) \\
 &= T_d((- \lambda_1^{a'} \lambda_2^{b'} + (-\lambda_1^{-a'} \lambda_2^{-b'})) \\
 &= (-\lambda_1^{a'} \lambda_2^{b'})^d + (-\lambda_1^{-a'} \lambda_2^{-b'})^d \\
 &= (-1)^d [\lambda_1^{da'} \lambda_2^{db'} + \lambda_1^{-da'} \lambda_2^{-db'}] \\
 &= (-1)^{ab+a+b} [\lambda_1^a \lambda_2^b + \lambda_1^{-a} \lambda_2^{-b}].
 \end{aligned}$$

□

We can easily get the following theorem by using the representation theory.

Theorem 3.2 Under the same assumption as in [Theorem 3.1](#), we have the following conclusions:

- (a) If $\lambda_1 \neq \pm 1$ or $\lambda_2 \neq \pm 1$, the representation $\rho_{u,v} G_{q^{1/2}}$ is irreducible.
- (b) If $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$, then V has only two irreducible subrepresentations, V_1 and V_2 , $V = V_1 \oplus V_2$, $\dim(V_1) = (n+1)/2$, and $\dim(V_2) = (n-1)/2$; especially

$$V_1 = \mathrm{span}\langle e_0, e_1 + e_{n-1}, e_2 + e_{n-2}, \dots, e_{(n-1)/2} + e_{(n+1)/2} \rangle,$$

$$V_2 = \mathrm{span}\langle e_1 - e_{n-1}, e_2 - e_{n-2}, \dots, e_{(n-1)/2} - e_{(n+1)/2} \rangle$$

if $u = \pm 1$ and $v = \pm 1$.

Remark 3.3 The Azumaya locus of $\mathrm{SK}_{q^{1/2}}(T^2)$ is a subset of $\mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(T^2)$; an element in $\mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(T^2)$ lives in the Azumaya locus if it corresponds to a unique irreducible representation of $\mathrm{SK}_{q^{1/2}}(T^2)$ (the correspondence is the one in [Theorem 2.1](#)). We know the PI-dimension of $\mathrm{SK}_{q^{1/2}}(T^2)$ is n . Then a character $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(T^2)$ lives in the Azumaya locus if and only if there exists an irreducible representation of $\mathrm{SK}_{q^{1/2}}(T^2)$ of dimension n whose classical shadow is $[\gamma]$. So by [Theorem 3.2](#), $[\gamma]$ lives in the Azumaya locus if and only if $\lambda_1 \neq \pm 1$ or $\lambda_2 \neq \pm 1$, where

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}.$$

A contemporaneous paper by Karuo and Korinman [\[18\]](#) considered instead the case when $q^{1/2}$ is an odd order root of unity; both cases were studied through similar methods. They proved the character lives in the Azumaya locus of the skein algebra of a closed surface if and only if the character is noncentral.

In [\[4\]](#), Bonahon and Wong proved the Witten–Reshetikhin–Turaev representation of the Kauffman bracket skein algebra is irreducible and whose classical shadow is the trivial character. For the closed torus T^2 , we use V_{T^2} to denote the Witten–Reshetikhin–Turaev representation of $\mathrm{SK}_{q^{1/2}}(T^2)$. We know $\dim V_{T^2} = \frac{1}{2}(n-1)$ with basis $v_1, v_2, \dots, v_{(n-1)/2}$ where v_k is the skein in the solid torus represented by

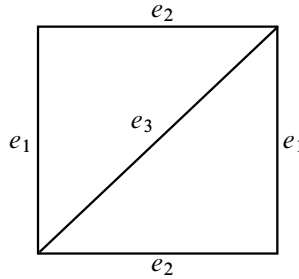


Figure 2

$2k - 2$ nontrivial parallel closed curves, which are parallel to the core of the solid torus, with the $(2k - 2)^{\text{nd}}$ Jones–Wenzel idempotent inserted. In [Theorem 3.2](#) with $\lambda_1 = \lambda_2 = 1$, V_2 and V_{T^2} are isomorphic as representations for $\text{SK}_{q^{1/2}}(T^2)$, and the isomorphism is given by

$$e_{2k-1} - e_{n-2k+1} \mapsto v_k \quad \text{for all } 1 \leq k \leq \frac{1}{2}(n-1).$$

3.3 Irreducible representations for $\text{SK}_{q^{1/2}}(S_{1,1})$

We want to find the explicit irreducible representations of $\text{SK}_{q^{1/2}}(S_{1,1})$ corresponding to given characters and puncture weights. Let $\mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}]$ be the algebra generated by X_1, X_2 and X_3 subject to the relations

$$X_1 X_2 = q X_2 X_1, \quad X_2 X_3 = q X_3 X_2, \quad X_3 X_1 = q X_1 X_2, \quad X_i X_i^{-1} = X_i^{-1} X_i = 1.$$

We have $\mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}] = \mathcal{T}_{\tau}^{q^{1/4}}(S_{1,1})$ where τ is the ideal triangulation in [Figure 2](#).

We define the skeins K_1, K_2 and K_3 in the skein algebra $\text{SK}_{q^{1/2}}(S_{1,1})$ using [Figure 3](#).

According to [\[10\]](#), the algebra $\text{SK}_{q^{1/2}}(S_{1,1})$ is generated by K_1, K_2 and K_3 subject to the relations

$$q^{-1/2} K_1 K_2 - q^{1/2} K_2 K_1 = (q^{-1} - q) K_3,$$

$$q^{-1/2} K_2 K_3 - q^{1/2} K_3 K_2 = (q^{-1} - q) K_1,$$

$$q^{-1/2} K_3 K_1 - q^{1/2} K_1 K_3 = (q^{-1} - q) K_2.$$

Let P be the loop around the puncture in $S_{1,1}$. Then

$$P = q^{-1/2} K_1 K_2 K_3 - q^{-1} K_1^2 - q K_2^2 - q^{-1} K_3^2 + q + q^{-1}.$$

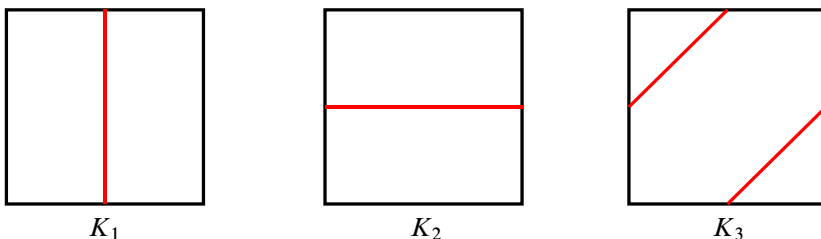


Figure 3

Lemma 3.4 *There is an algebraic embedding $F: S_{q^{1/2}}(S_{1,1}) \rightarrow \mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}]$ such that*

$$\begin{aligned} F(K_1) &= [X_2 X_3] + [X_2^{-1} X_3^{-1}] + [X_2 X_3^{-1}], \\ F(K_2) &= [X_3 X_1] + [X_3^{-1} X_1^{-1}] + [X_3 X_1^{-1}], \\ F(K_3) &= [X_1 X_2] + [X_1^{-1} X_2^{-1}] + [X_1 X_2^{-1}], \\ F(P) &= [X_1^2 X_2^2 X_3^2] + [X_1^{-2} X_2^{-2} X_3^{-2}]. \end{aligned}$$

Proof Actually, F is just the quantum trace map constructed in [2, Theorem 11] if we regard

$$\mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}]$$

as the Chekhov–Fock square root algebra associated to the ideal triangulation in Figure 2 where X_i corresponds to e_i for $i = 1, 2, 3$. $\square$

Let V be an n dimensional vector space over the complex field with basis $w_0, w_1, \dots, w_{n-1}$. We can define a representation $\rho_{r_1, r_2, r_3}: \mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}] \rightarrow \text{End}(V)$ such that

$$\begin{aligned} \rho_{r_1, r_2, r_3}(X_1)w_i &= r_1 q^i w_i, \\ \rho_{r_1, r_2, r_3}(X_2)w_i &= r_2 q^{-i} w_{i+1}, \\ \rho_{r_1, r_2, r_3}(X_3)w_i &= r_3 w_{i-1}, \end{aligned}$$

where r_1, r_2 and r_3 are nonzero complex numbers. We can get $\rho_{r_1, r_2, r_3}([X_1 X_2 X_3]) = r_1 r_2 r_3 q^{1/2} \text{Id}_V$.

Lemma 3.5 *For any three nonzero complex numbers r_1, r_2 and r_3 , ρ_{r_1, r_2, r_3} is an irreducible representation of $\mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}]$. Furthermore, every irreducible representation of $\mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}]$ is isomorphic to a representation ρ_{r_1, r_2, r_3} , and $\rho_{r_1, r_2, r_3} \simeq \rho_{s_1, s_2, s_3}$ if and only if*

$$r_1^n = s_1^n, \quad r_2^n = s_2^n, \quad r_3^n = s_3^n, \quad r_1 r_2 r_3 = s_1 s_2 s_3.$$

For any $\gamma \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(S_{1,1})$ and a nonzero complex number p such that $T_n(p) = -\text{Trace } \gamma(P)$ where P is the loop going around the only puncture in $S_{1,1}$, let $t_i = -\text{Trace } \gamma(K_i)$ for $i = 1, 2, 3$. According to [26], we have

$$T_n(p) = -t_1 t_2 t_3 - t_1^2 - t_2^2 - t_3^2 + 2.$$

Lemma 3.6 *Let x and y be two indeterminates such that $xy = q^{-2}yx$. Then*

$$T_n(x + x^{-1} + y) = x^{-n} + x^n + y^n$$

for $n \geq 1$.

For any given character $[\gamma] \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(S_{1,1})$, the following theorem offers a representation of $\text{SK}_{q^{1/2}}(S_{1,1})$ whose classical shadow is $[\gamma]$.

Theorem 3.7 With the above notation, $\rho_{r_1, r_2, r_3} F$ is a representation of $\text{SK}_{q^{1/2}}(S_{1,1})$. The classical shadow of $\rho_{r_1, r_2, r_3} F$ is γ and $\rho_{r_1, r_2, r_3} F(P) = p\text{Id}_V$ if and only if we have the following equations:

$$(1) \quad \begin{aligned} r_2^n r_3^n + r_2^{-n} r_3^{-n} + r_2^n r_3^{-n} &= -t_1, \\ r_3^n r_1^n + r_3^{-n} r_1^{-n} + r_3^n r_1^{-n} &= -t_2, \\ r_1^n r_2^n + r_1^{-n} r_2^{-n} + r_1^n r_2^{-n} &= -t_3, \\ r_1^2 r_2^2 r_3^2 q + r_1^{-2} r_2^{-2} r_3^{-2} q^{-1} &= p. \end{aligned}$$

Proof It is easy to see

$$\begin{aligned} [X_2 X_3][X_2 X_3^{-1}] &= q^{-2}[X_2 X_3^{-1}][X_2 X_3], \\ [X_3 X_1][X_3 X_1^{-1}] &= q^{-2}[X_3 X_1^{-1}][X_3 X_1], \\ [X_1 X_2][X_1 X_2^{-1}] &= q^{-2}[X_1 X_2^{-1}][X_1 X_2]. \end{aligned}$$

From Lemma 3.6, we get

$$(2) \quad \begin{aligned} T_n(\rho_{r_1, r_2, r_3} F(K_1)) &= T_n(\rho_{r_1, r_2, r_3} ([X_2 X_3] + [X_2^{-1} X_3^{-1}] + [X_2 X_3^{-1}])) \\ &= \rho_{r_1, r_2, r_3} (T_n([X_2 X_3] + [X_2^{-1} X_3^{-1}] + [X_2 X_3^{-1}])) \\ &= \rho_{r_1, r_2, r_3} ([X_2^n X_3^n] + [X_2^{-n} X_3^{-n}] + [X_2^n X_3^{-n}]) \\ &= \rho_{r_1, r_2, r_3} (-(X_2^n X_3^n + X_2^{-n} X_3^{-n} + X_2^n X_3^{-n})) \\ &= -(r_2^n r_3^n + r_2^{-n} r_3^{-n} + r_2^n r_3^{-n})\text{Id}_V. \end{aligned}$$

Similarly we can get

$$(3) \quad \begin{aligned} T_n(\rho_{r_1, r_2, r_3} F(K_2)) &= -(r_3^n r_1^n + r_3^{-n} r_1^{-n} + r_3^n r_1^{-n})\text{Id}_V, \\ T_n(\rho_{r_1, r_2, r_3} F(K_3)) &= -(r_1^n r_2^n + r_1^{-n} r_2^{-n} + r_1^n r_2^{-n})\text{Id}_V. \end{aligned}$$

And

$$(4) \quad \begin{aligned} \rho_{r_1, r_2, r_3} F(P) &= \rho_{r_1, r_2, r_3} ([X_1^2 X_2^2 X_1^2] + [X_1^{-2} X_2^{-2} X_1^{-2}]) \\ &= \rho_{r_1, r_2, r_3} ([X_1^2 X_2^2 X_1^2]) + \rho_{r_1, r_2, r_3} ([X_1^{-2} X_2^{-2} X_1^{-2}]) \\ &= (r_1^2 r_2^2 r_3^2 q + r_1^{-2} r_2^{-2} r_3^{-2} q^{-1})\text{Id}_V. \end{aligned}$$

From (2), (3), (4) and the fact that K_1 , K_2 and K_3 generate the algebra $\text{SK}_{-1}(S_{1,1})$, we can get the conclusions in Theorem 3.7. $\square$

Remark 3.8 At first glance, it seems like, in (1), we may not be able to get solutions, but actually the forth one is a consequence of first three equations because we have the relation $T_n(p) = -t_1 t_2 t_3 - t_1^2 - t_2^2 - t_3^2 + 2$. In fact, to get solutions, we only need to solve the equations

$$\begin{aligned} yz + y^{-1} z^{-1} + yz^{-1} &= -t_1, \\ zx + z^{-1} x^{-1} + zx^{-1} &= -t_2, \\ xy + x^{-1} y^{-1} + xy^{-1} &= -t_3. \end{aligned}$$

Let $Y_i = X_i^2$ for $i = 1, 2, 3$; then

$$Y_1 Y_2 = q^4 Y_2 Y_1, Y_2 Y_3 = q^4 Y_3 Y_2, Y_3 Y_1 = q^4 Y_1 Y_3, Y_i Y_i^{-1} = Y_i^{-1} Y_i = 1.$$

The subalgebra of $\mathbb{C}_q[X_1^{\pm 1}, X_2^{\pm 1}, X_3^{\pm 1}]$ generated by $Y_1^{\pm 1}, Y_2^{\pm 1}$ and $Y_3^{\pm 1}$ is $\mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}]$.

Here we recall a lemma from [7] for irreducible representations for $\mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}]$. Let V be an n dimensional vector space over $\mathbb{C}$ with basis $w_0, w_1, \dots, w_{n-1}$. For any three nonzero complex numbers y_1, y_2 and y_3 , define $\rho_{y_1, y_2, y_3}: \mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}] \rightarrow \text{End}(V)$ such that

$$\begin{aligned} \rho_{y_1, y_2, y_3}(Y_1)(w_i) &= y_1 q^{4i} w_i, \\ \rho_{y_1, y_2, y_3}(Y_2)(w_i) &= y_2 q^{-2i} w_{i+1}, \\ \rho_{y_1, y_2, y_3}(Y_3)(w_i) &= y_3 q^{-2i} w_{i-1}. \end{aligned} \quad (5)$$

Lemma 3.9 [7] (1) For any three nonzero complex numbers y_1, y_2 and y_3 , the representation ρ_{y_1, y_2, y_3} is irreducible.

(2) Every irreducible representation of $\mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}]$ is isomorphic to a representation ρ_{y_1, y_2, y_3} .

(3) The representations ρ_{y_1, y_2, y_3} and $\rho_{y'_1, y'_2, y'_3}$ are isomorphic if and only if $y_1^n = (y'_1)^n, y_2^n = (y'_2)^n, y_3^n = (y'_3)^n$ and $y_1 y_2 y_3 = y'_1 y'_2 y'_3$.

4 Calculation of intertwiners for the closed torus

4.1 Construction of intertwiners for the closed torus

The mapping class group of the closed torus is $\text{SL}(2, \mathbb{Z})$ [13]. For any $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$, we hope to find invariant characters under A . For a $[\gamma] \in \mathcal{R}_{\text{SL}(2, \mathbb{C})}(T^2)$, we choose a representative with

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}$$

where α and β denote loops $(1, 0)$ and $(0, 1)$ in $\mathbb{R}^2/\mathbb{Z}^2$ respectively. We have $[\gamma]$ is invariant under A if and only if $\text{Trace}(\gamma(A_*(z))) = \text{Trace}(\gamma(z))$ for all $z \in \pi_1(T^2)$. For any $(k_1, k_2) \in \mathbb{Z} \oplus \mathbb{Z}$, we have

$$A_*(k_1 \alpha + k_2 \beta) = (k_1, k_2) \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (k_1 a + k_2 c, k_1 b + k_2 d) = (k_1 a + k_2 c) \alpha + (k_1 b + k_2 d) \beta,$$

$$\begin{aligned} \gamma(A_*(k_1 \alpha + k_2 \beta)) &= \gamma[(k_1 a + k_2 c) \alpha + (k_1 b + k_2 d) \beta] \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix}^{k_1 a + k_2 c} \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}^{k_1 b + k_2 d} \\ &= \begin{pmatrix} \lambda_1^{k_1 a + k_2 c} \lambda_2^{k_1 b + k_2 d} & 0 \\ 0 & \lambda_1^{-(k_1 a + k_2 c)} \lambda_2^{-(k_1 b + k_2 d)} \end{pmatrix}, \\ \gamma(k_1 \alpha + k_2 \beta) &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix}^{k_1} \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}^{k_2} = \begin{pmatrix} \lambda_1^{k_1} \lambda_2^{k_2} & 0 \\ 0 & \lambda_1^{-k_1} \lambda_2^{-k_2} \end{pmatrix}. \end{aligned}$$

Then it is easy to show that $[\gamma]$ is A -invariant if and only if $\lambda_1 = \lambda_1^a \lambda_2^b$, $\lambda_2 = \lambda_1^c \lambda_2^d$ or $\lambda_1 = \lambda_1^{-a} \lambda_2^{-b}$, $\lambda_2 = \lambda_1^{-c} \lambda_2^{-d}$.

A also induces two algebra isomorphisms, $F_{A,+}$ and $F_{A,-}$, from $\mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$ to itself defined by

$$F_{A,+}(\theta_{(i,j)}) = \theta_{(i,j)A}, \quad F_{A,-}(\theta_{(i,j)}) = \theta_{(-i,-j)A}.$$

$F_{A,+}$ and $F_{A,-}$ are well defined because

$$\theta_{(i,j)A} \theta_{(k,l)A} = q^{1/2} \begin{bmatrix} (i,j)A \\ (k,l)A \end{bmatrix} \theta_{(i+k,j+l)A} = q^{1/2} \begin{bmatrix} i & j \\ k & l \end{bmatrix} [A] \theta_{(i+k,j+l)A} = q^{1/2} \begin{bmatrix} i & j \\ k & l \end{bmatrix} \theta_{(i+k,j+l)A},$$

and similarly

$$\theta_{(-i,-j)A} \theta_{(-k,-l)A} = q^{1/2} \begin{bmatrix} i & j \\ k & l \end{bmatrix} \theta_{(-i-k,-j-l)A}.$$

From Section 3, we know there is an embedding

$$G_{q^{1/2}}: \text{SK}_{q^{1/2}}(T^2) \rightarrow \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}.$$

For the following discussion we will omit the subscript for $G_{q^{1/2}}$ when there is no confusion.

Lemma 4.1 *The following diagram is commutative:*

$$\begin{array}{ccc} \text{SK}_{q^{1/2}}(T^2) & \xrightarrow{G} & \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}} \\ \downarrow A_{\#} & & \downarrow F_A \\ \text{SK}_{q^{1/2}}(T^2) & \xrightarrow{G} & \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}} \end{array}$$

where F_A is $F_{A,+}$ or $F_{A,-}$.

Proof We only prove the case when $F_A = F_{A,+}$.

First we show $A_{\#}((k, l)_T) = ((k, l)A)_T$. Assume $\gcd(k, l) = j$, $k = k'j$ and $l = l'j$; then

$$A_{\#}((k, l)_T) = A_{\#}(T_j((k', l')))) = T_j(A_{\#}((k', l')))) = T_j((k', l')A) = T_j(ak' + cl', bk' + dl').$$

There exist integers u and v such that $uk + vl = j$, so $\begin{bmatrix} k & l \\ -v & u \end{bmatrix} = j$ and $\det\left(\begin{bmatrix} k & l \\ -v & u \end{bmatrix} A\right) = j$. Then

$$\det\left(\begin{pmatrix} k & l \\ -v & u \end{pmatrix} A\right) = \begin{bmatrix} ak + cl & bk + dl \\ -v' & u' \end{bmatrix} = (ak + cl)u' + (bk + dl)v' = j.$$

We also have $j \mid (ak + cl)$ and $j \mid (bk + dl)$. Thus $\gcd(ak + cl, bk + dl) = j$, $ak + cl = j(ak' + cl')$ and $bk + dl = j(bk' + dl')$. Then

$$A_{\#}((k, l)_T) = T_j(ak' + cl', bk' + dl') = (ak + cl, bk + dl)_T = ((k, l)A)_T.$$

So

$$GA_{\#}((k, l)_T) = G(((k, l)A)_T) = \theta_{(k,l)A} + \theta_{(k,l)A}^{-1},$$

$$F_A G((k, l)_T) = F_A(\theta_{(k,l)} + \theta_{(k,l)}^{-1}) = \theta_{(k,l)A} + \theta_{(k,l)A}^{-1}.$$

We have $GA_{\#} = F_A G$ because all $(k, l)_T$ span the skein algebra. □

The following two theorems give the intertwiners for the closed surface for all the diffeomorphisms. We will give explicit formulas for these intertwiners and their trace in the following subsections.

Theorem 4.2 *In the diagram*

$$(6) \quad \begin{array}{ccccc} \mathrm{SK}_{q^{1/2}}(T^2) & \xrightarrow{G} & \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}} & \xrightarrow{\rho_{u,v}} & \mathrm{End}(V) \\ \downarrow A_{\#} & & \downarrow F_A & & \downarrow G_{\Lambda} \\ \mathrm{SK}_{q^{1/2}}(T^2) & \xrightarrow{G} & \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}} & \xrightarrow{\rho_{u,v}} & \mathrm{End}(V) \end{array}$$

suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$ and $F_A = F_{A,+}$. Let $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(T^2)$ with

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}$$

where $\lambda_1 = \lambda_1^a \lambda_2^b$ and $\lambda_2 = \lambda_1^c \lambda_2^d$, and let u and v be two complex numbers such that $u^n = -\lambda_1$ and $v^n = -\lambda_2$. We have the following conclusions:

- (a) $[\gamma]$ is invariant under A .
- (b) The classical shadow of $\rho_{u,v}G$ is $[\gamma]$.
- (c) $\rho_{u,v}F_A \simeq \rho_{u,v}$.
- (d) From (c), we know there exists an intertwiner $\Lambda_{n,+}$ such that $\rho_{u,v}F_A(Z) = \Lambda_{n,+}\rho_{u,v}(Z)\Lambda_{n,+}^{-1}$ for all $Z \in \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$. Then this intertwiner induces an intertwiner between two irreducible representations of the skein algebra.

Proof Items (a) and (b) are already shown in the previous discussion.

To prove (c), we have

$$\begin{aligned} \rho_{u,v}F_A(X^n) &= \rho_{u,v}F_A(\theta_{(n,0)}) = \rho_{u,v}(\theta_{(n,0)A}) \\ &= \rho_{u,v}(\theta_{(na,nb)}) \\ &= \rho_{u,v}((-1)^{ab}X^{na}Y^{nb}) \\ &= (-1)^{ab}u^{na}v^{nb}\mathrm{Id}_V \\ &= (-1)^{ab}(-\lambda_1)^a(-\lambda_2)^b\mathrm{Id}_V \\ &= (-1)^{ab+a+b}\lambda_1^a\lambda_2^b\mathrm{Id}_V = -\lambda_1\mathrm{Id}_V = u^n\mathrm{Id}_V. \end{aligned}$$

Similarly we can show $\rho_{u,v}F_A(Y^n) = v^n\mathrm{Id}_V$, thus $\rho_{u,v}F_A \simeq \rho_{u,v}$.

For (d), if $\lambda_1 \neq \pm 1$ or $\lambda_2 \neq \pm 1$, [Theorem 3.2](#) implies that $\Lambda_{n,+}$ itself is the intertwiner between two irreducible representations of the skein algebra. If $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$, [Theorem 3.2](#) implies that V has only two irreducible subrepresentations, V_1 and V_2 , with $\dim(V_1) = (n+1)/2$ and $\dim(V_2) = (n-1)/2$. We have that $\Lambda_{n,+}(V_1)$ is an irreducible subrepresentation of V and $\dim(\Lambda_{n,+}(V_1)) = (n+1)/2$. Then $\Lambda_{n,+}(V_1) = V_1$. Thus $\Lambda_{n,+}|_{V_1}$ is an intertwiner for V_1 . Similarly $\Lambda_{n,+}|_{V_2}$ is an intertwiner for V_2 . $\square$

Theorem 4.3 In the diagram (6), suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z})$ and $F_A = F_{A,-}$. Let $[\gamma] \in \mathcal{X}_{\mathrm{SL}(2, \mathbb{C})}(T^2)$ with

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}$$

where $\lambda_1 = \lambda_1^{-a} \lambda_2^{-b}$ and $\lambda_2 = \lambda_1^{-c} \lambda_2^{-d}$, and let u and v be two complex numbers such that $u^n = -\lambda_1^{-1}$ and $v^n = -\lambda_2^{-1}$. We have the following conclusions:

- (a) $[\gamma]$ is invariant under A .
- (b) The classical shadow of $\rho_{u,v}G$ is $[\gamma]$.
- (c) $\rho_{u,v}F_A \simeq \rho_{u,v}$.
- (d) From (c), we know there exists an intertwiner $\Lambda_{n,-}$ such that $\rho_{u,v}F_A(Z) = \Lambda_{n,-}\rho_{u,v}(Z)\Lambda_{n,-}^{-1}$ for all $Z \in \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$. Then this intertwiner induces an intertwiner between two irreducible representations of the skein algebra.

Proof The proof is the same as in Theorem 4.2. □

Note that a rescaling of the intertwiner $\Lambda_{n,+}$ in Theorem 4.2 such that $|\det(\Lambda_{n,+})| = 1$ makes $|\mathrm{Trace} \Lambda_{n,+}|$ independent of the choice of u and v . The same thing holds for the intertwiner in Theorem 4.3.

For the following discussion, we always require F_A to be $F_{A,+}$ unless specified otherwise (parallel results hold for $F_{A,-}$). From the above discussion, we know there exists an intertwiner $\Lambda_n \in \mathrm{End}(V)$ such that the diagram (6) commutes, where $G_\Lambda(B) = \Lambda_n B \Lambda_n^{-1}$ for all $B \in \mathrm{End}(V)$. Next we are going to find an intertwiner $\bar{\Lambda}_n$ under the assumption in Theorem 4.2.

4.2 Calculation for intertwiners

Under the assumption of Theorem 4.2, we have $\rho_{u,v}F_A \simeq \rho_{u,v}$. For any $a \in V$ and $Z \in \mathbb{C}[X^{\pm 1}, Y^{\pm 1}]_{q^{1/2}}$, we use $Z \cdot a$ and $Z \star a$ to denote $\rho_{u,v}(Z)(a)$ and $\rho_{u,v}F_A(Z)(a)$ respectively. Then we are trying to find $\bar{\Lambda}_n \in \mathrm{End}(V)$ such that $\bar{\Lambda}_n(X \cdot a) = X \star (\bar{\Lambda}_n(a))$ and $\bar{\Lambda}_n(Y \cdot a) = Y \star (\bar{\Lambda}_n(a))$ for all $a \in V$.

Remark 4.4 Assume $\gcd(b, n) = m$ and $n = n'm$. There exist two integers r and s such that $br + sn = m$. Then we have

$$\begin{aligned} (7) \quad (v^{n'b} u^{n'(a-1)} q^{ab(n')^2/2})^m &= v^{mn'b} u^{mn'(a-1)} q^{abm(n')^2/2} \\ &= v^{nb} u^{n(a-1)} q^{abnn'/2} \\ &= (-\lambda_2)^b (-\lambda_1)^{a-1} (-1)^{abn'} = (-1)^{ab+b+a-1} \lambda_2^b \lambda_1^{a-1} = 1, \end{aligned}$$

and $q^{an'}$ is a primitive m^{th} root of unity. Then there exists a unique integer $0 \leq k_0 \leq m-1$ such that $(v^{n'b} u^{n'(a-1)} q^{ab(n')^2/2}) q^{an'k_0} = 1$ and $(v^{n'b} u^{n'(a-1)} q^{ab(n')^2/2}) q^{an'k} \neq 1$ for $k \neq k_0, 0 \leq k \leq m-1$. We set $r_{k_0} = 1$ and $r_k = 0$ for $k \neq k_0, 0 \leq k \leq m-1$, and define $r_{k+tb} = r_k v^{tb} u^{t(a-1)} q^{a(tk+bt^2/2)}$ for

all $0 \leq k \leq m-1$ and $t \in \mathbb{Z}$, where we consider all indices modulo n . Since $\gcd(b, n) = m$ and $n = mn'$, we can reach all the indices. It is an easy check that $r_{k_1+t_1b} = r_{k_2+t_2b}$ if $k_1 + t_1b \equiv k_2 + t_2b \pmod{n}$. Then r_k is well defined for each $0 \leq k \leq n-1$.

It is easy to check that we have $r_{k+tb} = r_k v^{tb} u^{t(a-1)} q^{atk+abt^2/2}$ for all $k, t \in \mathbb{Z}$. Actually we have

$$r_{k_0+tb} = v^{tb} u^{t(a-1)} q^{a(tk_0+bt^2/2)} \quad \text{for all } 0 \leq t \leq n' - 1,$$

and all other r_k are 0. We have

$$(v^b u^{a-1})^n = (v^n)^b (u^n)^{a-1} = (-\lambda_2)^b (-\lambda_1)^{a-1} = (-1)^{a+b-1} \lambda_1^{a-1} \lambda_2^b = (-1)^{a+b-1}.$$

Then we get $|r_k| = 0$ or 1 for all $k \in \mathbb{Z}$. From $br + sn = m$, we get $tm = tbr + tsn$ for all $t \in \mathbb{Z}$. Then we have

$$r_{k_0+tm} = r_{k_0+trb} = v^{trb} u^{tr(a-1)} q^{a(trk_0+bt^2r^2/2)} \quad \text{for all } 0 \leq t \leq n' - 1,$$

and all other r_k are 0.

The following lemma offers an explicit formula for the intertwiner constructed in [Theorem 4.2\(d\)](#).

Lemma 4.5 Under the assumption of [Theorem 4.2](#), suppose $\bar{\Lambda}_n \in \text{End}(V)$ and

$$\bar{\Lambda}_n(e_t) = \sum_{0 \leq k \leq n-1} (\bar{\Lambda}_n)_{k,t} e_k$$

for all $0 \leq t \leq n-1$, where

$$(\bar{\Lambda}_n)_{k,t} = r_{k-td} (v^{(d-1)} u^c)^t q^{c(tk-dt^2/2)}.$$

Then $\bar{\Lambda}_n$ satisfies the conditions in [Theorem 4.2\(d\)](#).

Proof From direct calculations, we can get $\bar{\Lambda}_n(X \cdot e_t) = X \star (\bar{\Lambda}_n(e_t))$ and $\bar{\Lambda}_n(Y \cdot e_t) = Y \star (\bar{\Lambda}_n(e_t))$ for all $0 \leq t \leq n-1$. $\square$

We have $(v^{d-1} u^c)^n = (v^n)^{d-1} (u^n)^c = (-\lambda_2)^{d-1} (-\lambda_1)^c = (-1)^{c+d-1} \lambda_1^c \lambda_2^{d-1} = (-1)^{c+d-1}$. Then we can get $|(\bar{\Lambda}_n)_{k,t}| = 0$ or 1. We have $(\bar{\Lambda}_n)_{k,t} = 0$ if and only if $r_{k-td} = 0$. Then it is easy to show that $(\bar{\Lambda}_n)_{ld+km+k_0, l+tm}$, for $0 \leq l \leq m-1$ and $0 \leq k, t \leq n'-1$, are the only nonzero entries.

For each $0 \leq l \leq m-1$, we define an $n' \times n'$ matrix B^l such that $(B^l)_{k,t} = (\bar{\Lambda}_n)_{ld+km+k_0, l+tm}$ for all $0 \leq k, t \leq n'-1$. Then by Laplace expansion, we know $|\det(\bar{\Lambda}_n)| = \prod_{0 \leq l \leq m} |\det(B^l)|$.

From pure calculations, we can get $|\det(B^l)| = (n')^{n'/2}$. Then we have

$$|\det(\bar{\Lambda}_n)| = \prod_{0 \leq l \leq m} |\det(B^l)| = ((n')^{n'/2})^m = (n')^{mn'/2} = (n')^{n/2}.$$

Furthermore, $|\det((n')^{-1/2} \bar{\Lambda}_n)| = 1$.

Remark 4.6 If $\tilde{\Lambda}_n$ is the intertwiner in [Theorem 4.3](#), and we still suppose $\gcd(b, n) = m$, $br + sn = m$ and $n = n'm$, then

$$r_{k_0-tm} = (v^{-b}u^{-a-1})^{tr}q^{-trak_0+abt^2r^2/2} \quad \text{for all } t \in \mathbb{Z}$$

with all other $r_k = 0$, and

$$(\tilde{\Lambda}_n)_{k,t} = r_{k+td}(v^{-d-1}u^{-c})^tq^{-tck-cdt^2/2} \quad \text{for all } 0 \leq k, t \leq n-1,$$

where $0 \leq k_0 \leq m-1$ such that

$$(v^{-b}u^{-a-1})^{n'}q^{ab(n')^2/2}q^{-n'ak_0} = 1.$$

Also we can get $|\det(\tilde{\Lambda}_n)| = (n')^{n/2}$.

4.3 On the trace of intertwiners

Bonahon, Wong and Yang only formulated the conjecture when the mapping tori are hyperbolic. So they considered surfaces with negative Euler characteristic because the mapping tori for the closed torus can never be hyperbolic. Since the simplicial volume of mapping tori for the closed torus is zero — see [\[13, page 380\]](#) — we expect the corresponding limit to be zero. In [Theorem 4.15](#), we can show the limit superior is zero for any diffeomorphism. But the limits are not zero for some cases; see [Example 4.12](#). Some diffeomorphisms even do not have invariant characters that live in the Azumaya locus, but the intertwiners in [Theorems 4.2 and 4.3](#) are very close to intertwiners constructed in [\[7\]](#).

When we consider the intertwiners in [Theorems 4.2 and 4.3](#), we fix the mapping class A and the A -invariant character $[\gamma]$. In this subsection we will use (l, s) to denote $\gcd(l, s)$ for any two integers l and s .

Theorem 4.7 *If we require $|\det(\Lambda_n)| = 1$ for the intertwiner in [Theorem 4.2](#), then $|\text{Trace } \Lambda_n| \leq n^{3/2}$.*

Proof Since any two intertwiners in [Theorem 4.2](#) are different by a scalar multiplication or by conjugation and we require $|\det(\Lambda_n)| = 1$, the absolute value $|\text{Trace } \Lambda_n|$ is independent of the choice of intertwiners. Let $\Lambda_n = (n')^{-1/2}\bar{\Lambda}_n$; then $|\det(\Lambda_n)| = 1$. Since $|(\bar{\Lambda}_n)_{k,t}| = 0$ or 1 for all $0 \leq k, t \leq n-1$ and each row has exactly n' nonzero entries, we have the absolute value of every eigenvalue of $\bar{\Lambda}_n$ is not more than n' . Then

$$|\text{Trace}(\Lambda_n)| = |\text{Trace}((n')^{-1/2}\bar{\Lambda}_n)| \leq (n')^{-1/2}(nn') = (n')^{1/2}n \leq n^{3/2}. \quad \square$$

Theorem 4.8 *If we require $|\det(\Lambda_n)| = 1$ for the intertwiner in [Theorem 4.3](#), then $|\text{Trace } \Lambda_n| \leq n^{3/2}$.*

Proof It is similar to the proof for [Theorem 4.7](#). $\square$

Lemma 4.9 *Let k be any integer; then we have*

$$\left| \sum_{0 \leq t \leq n-1} (-q^{1/2})^{kt^2} \right| = \sqrt{(k, n)n}.$$

Recall that $q^{1/2}$ is a primitive n^{th} root of -1 .

Proof In [23] this result is proved for $k = 2$. Using the same trick, we can prove this generalized lemma. $\square$

In the following of this section, we always assume $q^{1/2} = e^{\pi i/n}$ unless especially specified. Next we are going to calculate $\text{Trace } \Lambda_n$, where Λ_n is the intertwiner in Theorem 4.2 or 4.3 with $|\det \Lambda_n| = 1$. First we give detailed discussion on the invariant character. Recall that for any $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z})$ and a character $[\gamma] \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(T^2)$ with

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix},$$

we have that $[\gamma]$ is A -invariant if and only if $1 = \lambda_1^{a-1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d-1}$ or $1 = \lambda_1^{a+1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d+1}$.

Remark 4.10 We will provide a detailed discussion only for the case when $1 = \lambda_1^{a-1} \lambda_2^b$ and $1 = \lambda_1^c \lambda_2^{d-1}$. Suppose $\lambda_1 = \alpha_1 e^{i\theta_1}$ and $\lambda_2 = \alpha_2 e^{i\theta_2}$; then we get

$$(8) \quad \begin{aligned} 1 &= \alpha_1^{a-1} \alpha_2^b, \quad 1 = \alpha_1^c \alpha_2^{d-1}, \\ (a-1)\theta_1 + b\theta_2 &= 2k_1\pi, \quad c\theta_1 + (d-1)\theta_2 = 2k_2\pi. \end{aligned}$$

Since $u^n = -\lambda_1$ and $v^n = -\lambda_2$, we can suppose $u = -\alpha_1^{1/n} e^{i\theta_1/n} q^{r_1}$ and $v = -\alpha_2^{1/n} e^{i\theta_2/n} q^{r_2}$ where both r_1 and r_2 are integers. Then we have

$$(9) \quad \begin{aligned} u^{a-1} v^b &= (-1)^{a+b-1} q^{r_1(a-1)+r_2b} e^{(i/n)((a-1)\theta_1+b\theta_2)} = (-1)^{a+b-1} q^{r_1(a-1)+r_2b+k_1}, \\ u^c v^{d-1} &= (-1)^{c+d-1} q^{r_1c+r_2(d-1)} e^{(i/n)(c\theta_1+(d-1)\theta_2)} = (-1)^{c+d-1} q^{r_1c+r_2(d-1)+k_2}. \end{aligned}$$

Define $s_1 = r_1(a-1) + r_2b + k_1$ and $s_2 = r_1c + r_2(d-1) + k_2$. Then $u^{a-1} v^b = (-1)^{a+b-1} q^{s_1}$ and $u^c v^{d-1} = (-1)^{c+d-1} q^{s_2}$.

From $1 = \lambda_1^{a-1} \lambda_2^b$ and $1 = \lambda_1^c \lambda_2^{d-1}$, we can get $1 = \lambda_1^{(a-1)c} \lambda_2^{bc}$ and $1 = \lambda_1^{(a-1)c} \lambda_2^{(a-1)(d-1)}$. Thus we have

$$\lambda_2^{bc} = \lambda_2^{(a-1)(d-1)} = \lambda_2^{ad-(a+d)+1} \implies 1 = \lambda_2^{ad-bc-(a+d)+1} = \lambda_2^{2-(a+d)}.$$

If $a+d \neq 2$, then λ_2 is a root of unity with $\lambda_2^{2-(a+d)} = 1$. Similarly we can show λ_1 is also a root of unity, with $\lambda_1^{2-(a+d)} = 1$, under the assumption $a+d \neq 2$.

We look at the case when $(b, n) = 1$, and suppose $br + sn = 1$. Then we have

$$q^{br} = q, \quad q^{br/2} = (-1)^s q^{1/2}, \quad (-1)^s = (-1)^{br+1}.$$

When $(b, n) = 1$, we can choose $k_0 = 1$ and set $r_{k_0} = r_0 = 1$. Then we have

$$r_{tb} = v^{tb} u^{t(a-1)} q^{abt^2/2} \quad \text{for all } t \in \mathbb{Z}.$$

For any $k \in \mathbb{Z}$, we have $k = krb + ksn$. Then

$$r_k = r_{krb} = v^{krb} u^{kr(a-1)} q^{abk^2r^2/2}.$$

From the above discussion, we know we can choose Λ_n to be $n^{-1/2}\bar{\Lambda}_n$. We have

$$\begin{aligned}
 (\bar{\Lambda}_n)_{t,t} &= r_{t-td}(v^{d-1}u^c)^t q^{ct^2-cdt^2/2} \\
 &= (-1)^{cdt} r_{t-td} q^{s_2 t} q^{ct^2-cdt^2/2} \\
 &= (-1)^{cdt} (v^b u^{a-1})^{r(t-td)} q^{abr^2(t-td)^2/2} q^{s_2 t} q^{ct^2-cdt^2/2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} q^{s_1 r(t-td)} q^{abr^2(t-td)^2/2} q^{s_2 t} q^{ct^2-cdt^2/2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} (q^{s_1 r(1-d)} q^{s_2})^t ((q^{br/2})^{ar(1-d)^2} q^{c-cd/2})^{t^2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} (q^{s_1 r(1-d)} q^{s_2})^t (((-1)^s q^{1/2})^{ar(1-d)^2} q^{c-cd/2})^{t^2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} (-1)^{tsar(1-d)^2} (q^{s_1 r(1-d)} q^{s_2})^t (q^{ar/2} q^{-ard} q^{ard^2/2} q^{c-cd/2})^{t^2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} (-1)^{tsar(1-d)^2} (q^{s_1 r(1-d)} q^{s_2})^t (q^{ar/2} q^{-r} q^{-rbc} q^{rd/2} q^{rbcd/2} q^{c-cd/2})^{t^2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} (-1)^{tsar(1-d)^2} (q^{s_1 r(1-d)} q^{s_2})^t (q^{ar/2} q^{-r} q^{-c} q^{rd/2} ((-1)^s q^{1/2})^{cd} q^{c-cd/2})^{t^2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} (-1)^{tsar(1-d)^2} (-1)^{tscd} (q^{s_1 r(1-d)} q^{s_2})^t (q^{(a+d-2)r/2})^{t^2} \\
 &= (-1)^{cdt} (-1)^{abr(t-td)} (-1)^{t(br+1)ar(1-d)} (-1)^{t(br+1)cd} (q^{s_1 r(1-d)} q^{s_2})^t (q^{(a+d-2)r/2})^{t^2} \\
 &= (-1)^{tar} (-1)^{tard} (-1)^{trd(ad+1)} (q^{s_1 r(1-d)} q^{s_2})^t (q^{(a+d-2)r/2})^{t^2} \\
 &= (-1)^{tar} (-1)^{trd} (q^{s_1 r(1-d)} q^{s_2})^t (q^{(a+d-2)r/2})^{t^2} \\
 &= ((-1)^{(a+d-2)r})^t (q^{s_1 r(1-d)} q^{s_2})^t (q^{(a+d-2)r/2})^{t^2}.
 \end{aligned}$$

Since $q^{s_2} = q^{rs_2b}$, we have

$$(10) \quad (\bar{\Lambda}_n)_{t,t} = ((-1)^{(a+d-2)r})^t q^{(r/2)((a+d-2)t^2+2(s_1(1+d)-s_2b)t)},$$

and

$$(11) \quad \text{Trace } \Lambda_n = n^{-1/2} \sum_{0 \leq t \leq n-1} ((-1)^{(a+d-2)r})^t q^{(r/2)((a+d-2)t^2+2(s_1(1+d)-s_2b)t)}.$$

Remark 4.11 Here we state the parallel results for $1 = \lambda_1^{a+1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d+1}$, $u^n = -\lambda_1^{-1}$, and $v^n = -\lambda_2^{-1}$.

Suppose $\lambda_1 = \alpha_1 e^{i\theta_1}$ and $\lambda_2 = \alpha_2 e^{i\theta_2}$; then we get

$$\begin{aligned}
 1 &= \alpha_1^{a+1} \alpha_2^b, \quad 1 = \alpha_1^c \alpha_2^{d+1}, \\
 (a+1)\theta_1 + b\theta_2 &= 2k_1\pi, \quad c\theta_1 + (d+1)\theta_2 = 2k_2\pi.
 \end{aligned}$$

Since $u^n = -\lambda_1^{-1}$ and $v^n = -\lambda_2^{-1}$, we can suppose $u = -\alpha_1^{-1/n} e^{-i\theta_1/n} q^{r_1}$ and $v = -\alpha_2^{-1/n} e^{-i\theta_2/n} q^{r_2}$ where both r_1 and r_2 are integers. Then we have

$$\begin{aligned}
 u^{a+1} v^b &= (-1)^{a+b+1} q^{r_1(a+1)+r_2b-k_1}, \\
 u^c v^{d+1} &= (-1)^{c+d+1} q^{r_1c+r_2(d+1)-k_2}.
 \end{aligned}$$

Similarly we set $s_1 = r_1(a+1) + r_2b - k_1$ and $s_2 = r_1c + r_2(d+1) - k_2$. Then $u^{a+1}v^b = (-1)^{a+b+1}q^{s_1}$ and $u^c v^{d+1} = (-1)^{c+d+1}q^{s_2}$. If $a+d \neq -2$, then $\alpha_1 = \alpha_2 = 1$.

For the case when $(b, n) = 1$ and $br + sn = 1$, we have

$$(\tilde{\Lambda}_n)_{t,t} = ((-1)^{(a+d+2)r})^t q^{(r/2)((a+d+2)t^2 + 2(s_1(1+d) - s_2b)t)}$$

and

$$(12) \quad \text{Trace } \Lambda_n = n^{-1/2} \sum_{0 \leq t \leq n-1} ((-1)^{(a+d+2)r})^t q^{(r/2)((a+d+2)t^2 + 2(s_1(1+d) - s_2b)t)}.$$

Example 4.12 Let $A = \begin{pmatrix} 2 & 1 \\ -7 & -3 \end{pmatrix}$. If we try to solve $1 = \lambda_1^{a-1} \lambda_2^b$ and $1 = \lambda_1^c \lambda_2^{d-1}$, we get

$$(13) \quad \theta_1 + \theta_2 = 2\pi, \quad -7\theta_1 - 4\theta_2 = 2\pi.$$

We have $\theta_1 = -\frac{10\pi}{3}$ and $\theta_2 = \frac{16\pi}{3}$; thus $\lambda_1 = e^{2\pi i/3}$ and $\lambda_2 = e^{4\pi i/3}$. So if we set $u = -e^{-10\pi i/3n}$ and $v = -e^{16\pi i/3n}$, then $u^{a-1}v^b = q$ and $u^c v^{d-1} = -q$. We have $s_1 = s_2 = 1$. Since $b = 1$, we get $r = 1$. Then from (11), we have

$$\text{Trace } \Lambda_n = n^{-1/2} \sum_{0 \leq t \leq n-1} (-1)^t q^{(-3t^2 + 10t)/2}.$$

Note that when n is a multiple of 3, we have $\text{Trace } \Lambda_n = 0$.

Example 4.13 Let A be the same matrix as above. But this time we try to solve $1 = \lambda_1^{a+1} \lambda_2^b$ and $1 = \lambda_1^c \lambda_2^{d+1}$; then we get $\lambda_1^{2+a+d} = \lambda_2^{2+a+d} = 1$. Since $2+a+d = 1$, we have $\lambda_1 = \lambda_2 = 1$. If we set $u = v = -1$, then $s_1 = s_2 = 0$. From (11), we have

$$|\text{Trace } \Lambda_n| = n^{-1/2} \left| \sum_{0 \leq t \leq n-1} (-1)^t q^{t^2/2} \right| = 1.$$

Lemma 4.14 Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where $(b, n) = 1$ and $|a+d| = 2$. Then we have the following statements:

- (1) If $a+d = 2$ and Λ_n is the intertwiner obtained in Theorem 4.2 such that $|\det(\Lambda_n)| = 1$, then $|\text{Trace } \Lambda_n| = \sqrt{n}$.
- (2) If $a+d = -2$ and Λ_n is the intertwiner obtained in Theorem 4.3 such that $|\det(\Lambda_n)| = 1$, then $|\text{Trace } \Lambda_n| = \sqrt{n}$.

Proof We only prove the statement (1) (the proof for statement (2) is similar). Let $[\gamma] \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(T^2)$, with

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix},$$

be any A -invariant character.

We use the same notation as in Remark 4.10. Then we have

$$\begin{aligned} s_1(1-d) + s_2b &= r_1(a-1)(1-d) + r_2b(1-d) + k_1(1-d) + r_1cb + r_2(d-1)b + k_2b \\ &= k_1(1-d) + k_2b. \end{aligned}$$

From (8), we can get

$$\begin{aligned} 2\pi((1-d)k_1 + k_2b) &= (1-d)2\pi k_1 + b2\pi k_2 \\ &= (1-d)(a-1)\theta_1 + b(1-d)\theta_2 + bc\theta_1 + b(d-1)\theta_2 = 0. \end{aligned}$$

Thus we have

$$(1-d)k_1 + k_2b = s_1(1-d) + s_2b = 0.$$

From (11), we know

$$\text{Trace } \Lambda_n = n^{-1/2} \sum_{0 \leq t \leq n-1} ((-1)^{(a+d-2)r})^t q^{(r/2)((a+d-2)t^2 + 2(s_1(1-d) + s_2b)t)} = n^{-1/2} n = \sqrt{n}. \quad \square$$

The following theorem shows the limit superior related to the trace of intertwiners for any diffeomorphism of the closed torus is zero, which equals the simplicial volume of the corresponding mapping torus.

Theorem 4.15 Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be any fixed element in the mapping class group for the closed torus, and let $[\gamma]$ be any fixed A -invariant character with

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}.$$

If $1 = \lambda_1^{a-1} \lambda_2^b$ and $1 = \lambda_1^c \lambda_2^{d-1}$, let $\{\Lambda_n\}_{n \in 2\mathbb{Z}_{\geq 0} + 1}$ be intertwiners obtained in Theorem 4.2 such that $|\det(\Lambda_n)| = 1$ for all $n \in 2\mathbb{Z}_{\geq 0} + 1$. If $1 = \lambda_1^{a+1} \lambda_2^b$ and $1 = \lambda_1^c \lambda_2^{d+1}$, let $\{\Lambda_n\}_{n \in 2\mathbb{Z}_{\geq 0} + 1}$ be intertwiners obtained in Theorem 4.3 such that $|\det(\Lambda_n)| = 1$ for all $n \in 2\mathbb{Z}_{\geq 0} + 1$. Then we have

$$\limsup_{\text{odd } n \rightarrow \infty} \frac{\log(|\text{Trace } \Lambda_n|)}{n} = 0.$$

Proof Since $[\gamma]$ is A -invariant, we have $1 = \lambda_1^{a-1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d-1}$ or $1 = \lambda_1^{a+1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d+1}$. We look at the case when $1 = \lambda_1^{a-1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d-1}$. Then we can set Λ_n to be $a_n \bar{\Lambda}_n$, where $a_n = |\det(\bar{\Lambda}_n)|^{-1/n}$.

Case I Suppose $b = 0$. In this case we know $|\det(\bar{\Lambda}_n)| = 1$ since $n' = 1$.

We have $A = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$ or $\begin{pmatrix} -1 & 0 \\ c & -1 \end{pmatrix}$. We first consider the case when $A = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$. If $c = 0$, it is trivial. So suppose $c \neq 0$. Since we have $\lambda_1^{a-1} \lambda_2^b = 1$, $\lambda_1^c \lambda_2^{d-1} = 1$, we get $\lambda_1^c = 1$. Suppose $\lambda_1 = e^{i\theta}$; then we get $\theta c = 2k\pi$ where k is an integer. Since $u^n = -\lambda_1$, we can choose $u = -e^{i\theta/n} q^r$ where r is an integer. Then we have

$$u^c = (-1)^c e^{i\theta c/n} q^{cr} = (-1)^c e^{2k\pi i/n} q^{cr} = (-1)^c q^{k+cr}.$$

Note that $|\text{Trace } \Lambda_n|$ is independent of the choice of r .

From Remark 4.4 and Lemma 4.5, Λ_n is a diagonal matrix, and

$$(\bar{\Lambda}_n)_{t,t} = (v^{(d-1)} u^c)^t q^{c(t^2 - dt^2/2)}.$$

Then we have

$$\text{Trace } \Lambda_n = \sum_{0 \leq t \leq n-1} v^{t(d-1)} u^{tc} q^{c dt^2/2} = \sum_{0 \leq t \leq n-1} (-1)^{ct} q^{ct^2/2} q^{(k+cr)t}.$$

Let $\{n_i\}_{i \in \mathbb{N}}$ be a subsequence of $2\mathbb{Z}_{\geq 0} + 1$ such that $(n_i, c) = 1$ for all i . Then for every i there exists r such that $k + cr \equiv 0 \pmod{n_i}$; thus

$$|\text{Trace } \Lambda_{n_i}| = \left| \sum_{0 \leq t \leq n_i - 1} (-1)^{ct} q^{ct^2/2} \right| = \sqrt{(n_i, c)n_i} = \sqrt{n_i} \geq 1.$$

Thus we have

$$0 \leq \limsup_{\text{odd } n \rightarrow \infty} \frac{\log(|\text{Trace } \Lambda_n|)}{n}.$$

According to [Theorem 4.7](#), we also have

$$\limsup_{\text{odd } n \rightarrow \infty} \frac{\log(|\text{Trace } \Lambda_n|)}{n} \leq \limsup_{\text{odd } n \rightarrow \infty} \frac{\log(n^{3/2})}{n} = 0.$$

We look at the case when $A = \begin{pmatrix} -1 & 0 \\ c & -1 \end{pmatrix}$. Then we get $(\lambda_1)^2 = 1$ and $\lambda_1 = \pm 1$. We can choose $u = \pm 1$. From [Remark 4.4](#) and [Lemma 4.5](#) we get

$$(\bar{\Lambda}_n)_{t,k} = r_{k-td} (v^{(d-1)} u^c)^t q^{c(tk-dt^2/2)},$$

where $r_k = 1$ if k is a multiple of n and it is zero otherwise. Then $(\bar{\Lambda}_n)_{t,t} \neq 0$ if and only if $r_{2t} \neq 0$ if and only if $n \mid (2t)$, which means there is only one nonzero diagonal element. Then we get $|\text{Trace } \Lambda_n| = 1$ for any n , which proves this special case.

Case II Suppose $b \neq 0$.

We first consider the subcase when $a + d \neq 2$. From the above discussion we know λ_1 and λ_2 are both roots of unity; thus we can suppose $\lambda_1 = e^{i\theta_1}$, $\lambda_2 = e^{i\theta_2}$ and we can get [\(8\)](#) where θ_1 , θ_2 , k_1 and k_2 are determined by γ . Since $u^n = -\lambda_1$ and $v^n = -\lambda_2$, we can write $u = -e^{i\theta_1/n} q^{r_1}$ and $v = -e^{i\theta_2/n} q^{r_1}$. Then we have [\(9\)](#). Note that $|\text{Trace } \Lambda_n|$ is independent of the choice of r_1 and r_2 .

Since $b \neq 0$ and $2 - (a + d) \neq 0$, let $\{n_j\}_{j \in \mathbb{N}}$ be a subsequence of $2\mathbb{Z}_{\geq 0} + 1$ such that

$$(n_j, b) = (n_j, 2 - (a + d)) = 1.$$

Since $\begin{bmatrix} a-1 & b \\ c & d-1 \end{bmatrix} = 2 - (a + d)$ and $(n_j, 2 - (a + d)) = 1$, the following equations always have solutions in $\mathbb{Z}_{n_j}$:

$$r_1(a-1) + r_2b + k_1 = 0, \quad r_1c + r_2(d-1) + k_2 = 0.$$

Thus for every j , there always exist integers r_1 and r_2 such that $s_1 = s_2 = 0$ in $\mathbb{Z}_{n_j}$. Then, from [\(11\)](#),

$$\begin{aligned} |\text{Trace } \Lambda_{n_j}| &= n_j^{-1/2} \left| \sum_{0 \leq t \leq n_j - 1} ((-1)^{(a+d-2)r})^t q^{(r/2)((a+d-2)t^2 + 2(s_1(1-d) + s_2b)t)} \right| \\ &= n_j^{-1/2} \left| \sum_{0 \leq t \leq n_j - 1} ((-1)^{(a+d-2)r})^t q^{(r/2)((a+d-2)t^2)} \right| \\ &= n_j^{-1/2} \left| \sum_{0 \leq t \leq n_j - 1} (-q^{1/2})^{r(a+d-2)t^2} \right| = n_j^{-1/2} \sqrt{(r(a+d-2), n_j)n_j} = 1. \end{aligned}$$

Then we have

$$0 \leq \limsup_{\text{odd } n \rightarrow \infty} \frac{\log(|\text{Trace } \Lambda_n|)}{n}.$$

By [Theorem 4.7](#),

$$\limsup_{\text{odd } n \rightarrow \infty} \frac{\log(|\text{Trace } \Lambda_n|)}{n} = 0.$$

The other subcase is when $a + d = 2$. Since $b \neq 0$, let $\{n_k\}_{k \in \mathbb{N}}$ be a subsequence of $2\mathbb{Z}_{\geq 0} + 1$ such that $(n_k, b) = 1$ for all k . From [Lemma 4.14](#), we have $|\text{Trace } \Lambda_{n_k}| = \sqrt{n_k} \geq 1$. Similarly, we get

$$\limsup_{\text{odd } n \rightarrow \infty} \frac{\log(|\text{Trace } \Lambda_n|)}{n} = 0. \quad \square$$

From now on we discuss the periodic mapping class. Recall that $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is periodic if and only if $|a + d| \in \{0, 1\}$. Suppose $[\gamma]$ is an A -invariant character with

$$\gamma(\alpha) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1^{-1} \end{pmatrix} \quad \text{and} \quad \gamma(\beta) = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}.$$

Then we have $1 = \lambda_1^{a-1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d-1}$ or $1 = \lambda_1^{a+1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d+1}$. For the case when $1 = \lambda_1^{a-1} \lambda_2^b$, $1 = \lambda_1^c \lambda_2^{d-1}$, the above discussion implies $1 = \lambda_1^{2-(a+d)} = \lambda_2^{2-(a+d)}$. So if $a + d = 0$, we have $\lambda_1^2 = \lambda_2^2 = 1$; thus $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$. Hence there is no A -invariant character living in the Azumaya locus if $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$. But we can still get intertwiners in [Theorems 4.2](#) and [4.3](#), although $\lambda_1 = \pm 1$ and $\lambda_2 = \pm 1$. Now we consider intertwiners if we choose $\lambda_1 = \lambda_2 = 1$.

Theorem 4.16 *Let A be a periodic mapping class, and let Λ_n be the intertwiner obtained in [Theorem 4.2](#) or [4.3](#) by using the trivial A -invariant character, that is $\lambda_1 = \lambda_2 = 1$, and we require $|\det(\Lambda_n)| = 1$. We have the following conclusions:*

- (1) *If $a + d = 1$ and Λ_n is obtained in [Theorem 4.2](#), then $|\text{Trace}(\Lambda_n)| = 1$ for any odd n .*
- (2) *If $a + d = -1$ and Λ_n is obtained in [Theorem 4.3](#), then $|\text{Trace}(\Lambda_n)| = 1$ for any odd n .*
- (3) *If $a + d = 0$ and Λ_n is obtained in [Theorem 4.2](#) or [4.3](#), then $|\text{Trace}(\Lambda_n)| = 1$ for any odd n .*

Proof Suppose $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $(b, n) = m$, $br + sn = m$ and $n = mn'$. Since $\lambda_1 = \lambda_2 = 1$, we can set $u = v = -1$. From the previous discussion we know

$$(\bar{\Lambda}_n)_{k,t} = (-1)^{cdt} r_{k-t} d q^{c(tk-dt^2/2)}$$

for all $0 \leq k, t \leq n-1$, where

$$r_{tm} = (-1)^{abrt} q^{abt^2r^2/2} \quad \text{for all } 0 \leq t \leq n' - 1,$$

and all other r_k are 0. For the l^{th} column, we have $\{(\bar{\Lambda}_n)_{ld+km,l}\}_{0 \leq k \leq n'-1}$ are the only nonzero entries. Then the l^{th} column contains a nonzero diagonal entry if and only if $ld + km \equiv l \pmod{n}$ for some $0 \leq k \leq n' - 1$. It is easy to show $ld + km \equiv l \pmod{n}$ for some $0 \leq k \leq n' - 1$ if and only if $m \mid (ld - l)$.

Now we suppose $(m, d-1) = 1$. Then the l^{th} column of $\bar{\Lambda}_n$ contains a nonzero diagonal entry if and only if $m \mid l$. Thus $(\bar{\Lambda}_n)_{tm,tm}$, $0 \leq t \leq n'-1$, are the only nonzero diagonal entries, and

$$\begin{aligned} (\bar{\Lambda}_n)_{tm,tm} &= (-1)^{cdtm} r_{tm-tmd} q^{c(t^2m^2-dt^2m^2/2)} \\ &= (-1)^{cdtm} (-1)^{abr(t-d)} q^{abr^2(1-d)^2t^2/2} q^{c(t^2m^2-dt^2m^2/2)}. \end{aligned}$$

After a similar calculation as in [Remark 4.10](#), we can get

$$(\bar{\Lambda}_n)_{tm,tm} = (-1)^{(ar+dr-2r)t} (q^{m/2})^{(ar+dr-2r)t^2}.$$

Then we have

$$\text{Trace } \bar{\Lambda}_n = \sum_{0 \leq t \leq n'-1} (-1)^{(ar+dr-2r)t} (q^{m/2})^{(ar+dr-2r)t^2}.$$

Similarly if $(m, d+1) = 1$, then

$$\text{Trace } \tilde{\Lambda}_n = \sum_{0 \leq t \leq n'-1} (-1)^{(ar+dr+2r)t} (q^{m/2})^{(ar+dr+2r)t^2}.$$

- (1) Since the intertwiner is obtained in [Theorem 4.2](#), we can set $\Lambda_n = (n')^{-1/2} \bar{\Lambda}_n$. We have $d-1 = -a$ because $a+d = 1$. Then we get $(d-1, m) = 1$ because $(a, b) = 1$ and m is a divisor of b . Then from the above discussion, we get

$$\begin{aligned} |\text{Trace } \Lambda_n| &= (n')^{-1/2} \left| \sum_{0 \leq t \leq n'-1} (-1)^{(ar+dr-2r)t} (q^{m/2})^{(ar+dr-2r)t^2} \right| \\ &= (n')^{-1/2} \left| \sum_{0 \leq t \leq n'-1} (-1)^{-rt} (q^{m/2})^{-rt^2} \right| \\ &= (n')^{-1/2} \sqrt{(-r, n') n'} = 1. \end{aligned}$$

- (2) The proof is similar to that of (1).

- (3) First we show $(m, d-1) = (m, d+1) = 1$ if $a+d = 0$. From $ad-bc = 1$, we get $-bc = d^2 + 1$. Suppose $p \mid m$ and $p \mid d-1$. Then $p \mid (d^2 + 1)$ and $p \mid (d^2 - 1)$. Thus we get $p \mid 2$, which means $p = 1$ because $(m, 2) = 1$. Similarly we can show $(m, d+1) = 1$. If $\Lambda_n = (n')^{-1/2} \bar{\Lambda}_n$, then

$$\begin{aligned} |\text{Trace } \Lambda_n| &= (n')^{-1/2} \left| \sum_{0 \leq t \leq n'-1} (-1)^{(ar+dr-2r)t} (q^{m/2})^{(ar+dr-2r)t^2} \right| \\ &= (n')^{-1/2} \left| \sum_{0 \leq t \leq n'-1} (q^m)^{-rt^2} \right| \\ &= (n')^{-1/2} \sqrt{(-r, n') n'} = 1. \end{aligned}$$

If $\Lambda_n = (n')^{-1/2} \tilde{\Lambda}_n$, we can similarly show that $|\text{Trace } \Lambda_n| = 1$. □

5 The volume conjecture for surface diffeomorphisms: periodic case

5.1 Preliminaries for the volume conjecture for periodic surface diffeomorphisms

If we want to formulate the parallel conjecture for periodic diffeomorphisms as in [7; 8], we have to find a good invariant character that lives in the smooth part of $\mathcal{R}_{\mathrm{SL}(2, \mathbb{C})}(S)$.

Lemma 5.1 [11] *Let $A, B \in \mathrm{SL}(2, \mathbb{C})$. If $\mathrm{Trace}([A, B]) = 2$ where $[A, B] = ABA^{-1}B^{-1}$, then $G = \langle A, B \rangle \leq \mathrm{SL}(2, \mathbb{C})$ is not free of rank two where $\langle A, B \rangle$ is the group generated by A and B .*

Lemma 5.2 *Let G be a subgroup of $\mathrm{SL}(2, \mathbb{C})$ freely generated by two elements, and let R be the subalgebra of $\mathrm{Mat}(2, \mathbb{C})$ generated by G , where $\mathrm{Mat}(2, \mathbb{C})$ is the algebra of all 2 by 2 complex matrices. Then $R = \mathrm{Mat}(2, \mathbb{C})$.*

Proof Suppose G is freely generated by A and B . We know there exists $X \in \mathrm{GL}(2, \mathbb{C})$ such that $XAX^{-1} = \begin{pmatrix} u & v \\ 0 & u^{-1} \end{pmatrix}$ and $XBX^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then XGX^{-1} is a free subgroup generated by XAX^{-1} and XBX^{-1} , and XRX^{-1} is the subalgebra generated by XGX^{-1} . Since $XRX^{-1} = \mathrm{Mat}(2, \mathbb{C})$ if and only if $R = \mathrm{Mat}(2, \mathbb{C})$, we can assume $A = \begin{pmatrix} u & v \\ 0 & u^{-1} \end{pmatrix}$ and $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(I) Suppose $v = 0$. Then $A = \begin{pmatrix} u & 0 \\ 0 & u^{-1} \end{pmatrix}$ and $u^2 \neq 1$, otherwise $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, which contradicts the fact that G is freely generated by A and B . We also get $b \neq 0$ and $c \neq 0$, since otherwise $\mathrm{Trace}([A, B]) = 2$, which contradicts the fact that A and B freely generate G by Lemma 5.1. Since $\begin{pmatrix} u & 0 \\ 0 & u^{-1} \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in R$ and $u \neq \pm 1$, we have $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \in R$. Then $\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix} \in R$. By multiplication, $\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} \in R$, which implies $R = \mathrm{Mat}(2, \mathbb{C})$ since $b \neq 0$ and $c \neq 0$.

(II) Suppose $v \neq 0$. In this case we should have $c \neq 0$, otherwise $\mathrm{Trace}([A, B]) = 2$, which is a contradiction.

If $u = \pm 1$, then $A = \begin{pmatrix} u & v \\ 0 & u \end{pmatrix} \in R$. Remember we also have $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in R$, which implies $\begin{pmatrix} 0 & v \\ 0 & 0 \end{pmatrix} \in R$. Furthermore, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in R$ because $v \neq 0$. Since $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in R$, we have $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \in R$, and also $\begin{pmatrix} a-d & 0 \\ c & 0 \end{pmatrix} \in R$. By multiplication, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a-d & 0 \\ c & 0 \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & 0 \end{pmatrix} \in R$, which implies $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in R$ because $c \neq 0$. We have $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in R$, so $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in R$. Remember we also have $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in R$ and $c \neq 0$, which implies $R = \mathrm{Mat}(2, \mathbb{C})$.

If $u \neq \pm 1$, then $\begin{pmatrix} u-u^{-1} & v \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & v \\ 0 & u^{-1}-u \end{pmatrix} \in R$, so $\begin{pmatrix} 1 & k \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -k \\ 0 & 1 \end{pmatrix} \in R$ where $k = v/(u-u^{-1})$. Then from multiplication, we get

$$\begin{pmatrix} 1 & k \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+kc & b+kd \\ 0 & 0 \end{pmatrix} \in R.$$

Next we want to show $b+kd \neq k(a+kc)$. Suppose the contrary. Then $b+kd = k(a+kc) = ka+k^2c$. With $k = v/(u-u^{-1})$, we can get

$$b + \frac{dv}{u-u^{-1}} = \frac{av}{u-u^{-1}} + \frac{cv^2}{(u-u^{-1})^2} \implies 2b = bu^2 + bu^{-2} + dvu - dvu^{-1} - avu + avu^{-1} - cv^2.$$

Then we get

$$\begin{aligned} \text{Trace}([A, B]) &= ad + acuv + cdvu^{-1} + c^2v^2 - cbu^2 - cdv - cbu^{-2} - cau^{-1}v + ad \\ &= 2ad - c(-auv - dvu^{-1} - cv^2 + bu^2 + duv + bu^{-2} + au^{-1}v) = 2ad - 2cb = 2. \end{aligned}$$

Since $\text{Trace}([A, B]) = 2$ is a contradiction, we have $b + kd \neq k(a + kc)$. We then get $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in R$ because $\begin{pmatrix} a+kc & b+kd \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & k \\ 0 & 0 \end{pmatrix} \in R$. We also have $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in R$, so $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in R$. Thus, $R = \text{Mat}(2, \mathbb{C})$ because $c \neq 0$. $\square$

Proposition 5.3 *Let $\gamma: \pi_1(S) \rightarrow \text{SL}(2, \mathbb{C})$ be a representative of an element in the character variety $\mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$. Then γ is irreducible if $\text{Im } \gamma$ contains a subgroup of $\text{SL}(2, \mathbb{C})$ freely generated by two elements. In particular, γ is irreducible if S has negative Euler characteristic and γ is injective.*

Proof This is a direct consequence of Lemma 5.2. $\square$

5.2 Statement of the conjecture

To get the intertwiner, we first have to get a φ -invariant smooth character $\gamma \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$. In [13, page 371] it is proved that every periodic diffeomorphism fixes a point in the Teichmüller space. This means there is a discrete and faithful group homomorphism $\bar{\gamma}: \pi_1(S) \rightarrow \text{PSL}(2, \mathbb{R})$ such that $\bar{\gamma}\varphi_*$ is conjugate to $\bar{\gamma}$ by an element in $\text{PSL}(2, \mathbb{R})$, where φ_* is the isomorphism from $\pi_1(S)$ to $\pi_1(S)$ induced by φ .

Since $\text{PSL}(2, \mathbb{R}) \subset \text{PSL}(2, \mathbb{C})$, we can regard $\bar{\gamma}\varphi_*$ and $\bar{\gamma}$ as two elements in $\mathcal{X}_{\text{PSL}(2, \mathbb{C})}(S)$. Then $\bar{\gamma}\varphi_*$ is conjugate to $\bar{\gamma}$ by an element in $\text{PSL}(2, \mathbb{C})$. Thus $\bar{\gamma}$ can be extended to a group homomorphism from $\pi_1(M_\varphi)$ to $\text{PSL}(2, \mathbb{C})$; we use $\hat{\gamma}$ to denote this homomorphism. Then we can lift $\hat{\gamma}$ to a group homomorphism $\tilde{\gamma}$ from $\pi_1(M_\varphi)$ to $\text{SL}(2, \mathbb{C})$. The restriction of $\tilde{\gamma}$ to $\pi_1(S)$ is φ -invariant, and we use γ to denote this group homomorphism. Note that γ is a group homomorphism from $\pi_1(S)$ to $\text{SL}(2, \mathbb{C})$. Let ε be the projection from $\text{SL}(2, \mathbb{C})$ to $\text{PSL}(2, \mathbb{C})$; then we have $\varepsilon\tilde{\gamma} = \hat{\gamma}$. Furthermore,

$$\varepsilon\gamma = \varepsilon\tilde{\gamma}|_{\pi_1(S)} = \hat{\gamma}|_{\pi_1(S)} = \bar{\gamma}.$$

Since $\bar{\gamma}$ is injective, γ is injective. From Proposition 5.3, we know γ is irreducible. Thus we get a φ -invariant smooth character $\gamma \in \mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$. From now on, we use γ_φ to denote γ and $\bar{\gamma}_\varphi$ to denote $\bar{\gamma}$. For every puncture v in S , we know $\text{Trace } \gamma_\varphi(\alpha_v) = \pm 2$ where α_v is the loop going around puncture v . If $\text{Trace } \gamma_\varphi(\alpha_v) = 2$, we choose $p_v = -(q + q^{-1})$. Then

$$T_n(p_v) = (-q)^n + (-q^{-1})^n = -1 - 1 = -\text{Trace } \gamma_\varphi(\alpha_v).$$

If $\text{Trace } \gamma_\varphi(\alpha_v) = -2$, we choose $p_v = 1 + 1$. Then

$$T_n(p_v) = 1^n + 1^n = 1 + 1 = -\text{Trace } \gamma_\varphi(\alpha_v).$$

Since $\text{Trace } \gamma_\varphi(\alpha_v) = \text{Trace } \gamma_\varphi(\varphi(\alpha_v)) = \text{Trace } \gamma_\varphi(\alpha_{\varphi(v)})$, we have $p_v = p_{\varphi(v)}$. So now we have everything we want. Then we obtain the Kauffman bracket intertwiner $\Lambda_{\varphi, \gamma_\varphi}^q$ associated to these data. We require $|\det(\Lambda_{\varphi, \gamma_\varphi}^q)| = 1$. With the fixed S , φ , γ_φ and $\{p_v\}_v$, we have $|\text{Trace } \Lambda_{\varphi, r}^q|$ is only related to q .

Conjecture 5.4 Suppose S is an oriented surface with negative Euler characteristic, and φ is a periodic diffeomorphism for S . Let γ_φ be the φ -invariant smooth character defined as in the second paragraph of this subsection. For each puncture v , let p_v be the complex number defined as in the third paragraph of this subsection. Let $q_n = e^{2\pi i/n}$ with $(q_n)^{1/2} = e^{\pi i/n}$. Then

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| = 0.$$

5.3 Proofs for the conjecture for some special cases

In the remaining part of this paper, we will present some results related to our conjecture. Especially, we will give a proof for our conjecture when the surface S is the once punctured torus.

In the following theorem, we use the periodic property of the diffeomorphisms to prove that the limit in [Conjecture 5.4](#) is less than or equal to zero if it exists.

Theorem 5.5 If $\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}|$ exists, the limit is less than or equal to zero.

Proof Let $\rho: \text{SK}_{q_n^{1/2}}(S) \rightarrow \text{End}(V)$ be an irreducible representation of the skein algebra associated to γ_φ and weight system $\{p_v\}_v$. From the definition of intertwiners $\Lambda_{\varphi, r_\varphi}^{q_n}$, we know

$$\rho \varphi_\#(X) = \Lambda_{\varphi, r_\varphi}^{q_n} \circ \rho(X) \circ (\Lambda_{\varphi, r_\varphi}^{q_n})^{-1}$$

for all $X \in \text{SK}_{q_n^{1/2}}(S)$. We have

$$\rho(\varphi^2)_\#(X) = \rho \varphi_\#(\varphi_\#(X)) = \Lambda_{\varphi, r_\varphi}^{q_n} \circ \rho \varphi_\#(X) \circ (\Lambda_{\varphi, r_\varphi}^{q_n})^{-1} = (\Lambda_{\varphi, r_\varphi}^{q_n})^2 \circ \rho \varphi_\#(X) \circ (\Lambda_{\varphi, r_\varphi}^{q_n})^{-2}.$$

Then it is easy to show that, with any integer j , we have

$$\rho(\varphi^j)_\#(X) = (\Lambda_{\varphi, r_\varphi}^{q_n})^j \circ \rho \varphi_\#(X) \circ (\Lambda_{\varphi, r_\varphi}^{q_n})^{-j}.$$

Since φ is periodic, there exists a positive integer k such that $\varphi^k = \text{Id}_S$. Then we have

$$\rho(X) = \rho(\varphi^k)_\#(X) = (\Lambda_{\varphi, r_\varphi}^{q_n})^k \circ \rho(X) \circ (\Lambda_{\varphi, r_\varphi}^{q_n})^{-k}$$

for all $X \in \text{SK}_{q_n^{1/2}}(S)$. We must have $(\Lambda_{\varphi, r_\varphi}^{q_n})^k = \lambda I$ because ρ is irreducible, where I is the identity matrix and λ is a nonzero complex number. But we require $|\det(\Lambda_{\varphi, r_\varphi}^{q_n})| = 1$; thus $|\lambda| = 1$. Actually we can always choose a good $\Lambda_{\varphi, r_\varphi}^{q_n}$ such that $(\Lambda_{\varphi, r_\varphi}^{q_n})^k = I$. Since $x^k - 1$ has no multiple roots, $\Lambda_{\varphi, r_\varphi}^{q_n}$ is always diagonalizable. All its eigenvalues are k^{th} roots of unity. Then

$$\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n} = \sum_{0 \leq i \leq n-1} \lambda_i,$$

where $\lambda_i^k = 1$ for all $0 \leq i \leq n-1$.

We have $|\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| \leq n$. So if the limit exists, the limit is less than or equal to zero. □

From the proof of [Theorem 5.5](#), we know $|\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}|$ is simply the absolute value of the sum of roots of unity. We are only concerned with how small $|\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}|$ can be because of [Theorem 5.5](#). Actually this problem was already asked by Myerson [\[24\]](#) and Tao [\[27\]](#). For any two positive integers k and n , let $f(n, k)$ be the least absolute value of a nonzero sum of n (not necessarily distinct) k^{th} roots of unity. Myerson gave the lower bound for all positive integers k and n :

$$(14) \quad f(n, k) \geq n^{-k}.$$

According to [\[20\]](#), we know $\text{Trace } \Lambda_{\varphi, r}^{q_n} \neq 0$ if the order of φ is 2^m for some positive integer m .

Theorem 5.6 *If φ is of order 2^m where m is any positive integer, then for any surface with negative Euler characteristic,*

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| = 0.$$

Proof Since for any odd n , we have $\text{Trace } \Lambda_{\varphi, r}^{q_n} \neq 0$. Then

$$n^{-k} \leq f(n, k) \leq |\text{Trace } \Lambda_{\varphi, r}^{q_n}|,$$

where $k = 2^m$. So we get

$$\frac{1}{n} \log n^{-k} \leq \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| \leq \frac{1}{n} \log n.$$

Then $\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| = 0$. □

Proposition 5.7 *If φ is of order p^m where p is any positive prime number and m is any positive integer, then for any surface with negative Euler characteristic,*

$$\limsup_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| = 0.$$

Proof The proof is similar to that of [Theorem 5.6](#). □

For the following discussion, we will use some notation and terminology in [\[7\]](#). Suppose the surface S has at least one puncture, that is, it has ideal triangulations. Let τ be an ideal triangulation of S , and let φ be any periodic map of S . Suppose $\tau = \tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)} = \varphi(\tau)$ is an ideal triangulation sweep. Since φ fixes a point in the Teichmüller space, there exists a periodic edge weight system $a = a^{(0)}, a^{(1)}, \dots, a^{(k)} = a \in (\mathbb{R}_{>0})^e$ where a is the shear parameter corresponding to this fixed point in the Teichmüller space. Then $[\gamma_\varphi] \in \mathcal{X}_{\text{PSL}(2, \mathbb{C})}(S)$ is the character associated to the weight system a . From the above discussion, we know $[\gamma_\varphi]$ can be lift to a smooth φ -invariant character $[\gamma_\varphi]$ in $\mathcal{X}_{\text{SL}(2, \mathbb{C})}(S)$.

We also have $a_{i_1} a_{i_2} \cdots a_{i_j} = 1$, where $e_{i_1}, e_{i_2}, \dots, e_{i_j}$ are all the edges connecting to a common vertex, because a corresponds to a complete hyperbolic structure. If $\text{Trace } \gamma_\varphi(\alpha_v) = 2$, set $h_v = q^2$. Then $h_v^n = 1$ and $p_v^2 = h_v + h_v^{-1} + 2$. If $\text{Trace } \gamma_\varphi(\alpha_v) = -2$, set $h_v = 1$. Then $h_v^n = 1$ and $p_v^2 = h_v + h_v^{-1} + 2$. Obviously $h_v = h_{\varphi(v)}$ for any puncture v . Proposition 15 of [\[7\]](#) implies that we can obtain an intertwiner $\bar{\Lambda}_{\varphi, \bar{r}_\varphi}^q$ with $|\det(\bar{\Lambda}_{\varphi, \bar{r}_\varphi}^q)| = 1$. According to [\[7, Theorem 16\]](#), we have $|\text{Trace } \bar{\Lambda}_{\varphi, \bar{r}_\varphi}^q| = |\text{Trace } \Lambda_{\varphi, r_\varphi}^q|$.

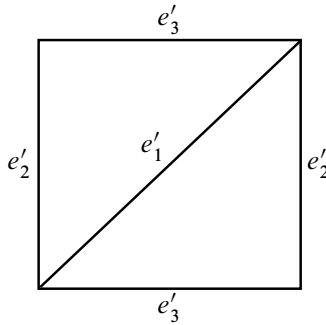


Figure 4

For the once punctured torus $S_{1,1}$, we only have one puncture v . Let $\alpha = K_2$ and $\beta = K_1$ denote two elements in $\pi_1(S_{1,1})$; see Figure 3. It is well known that α and β freely generate $\pi_1(S_{1,1})$. Let α_v be the loop around v . Then $\alpha_v = \beta\alpha\beta^{-1}\alpha^{-1}$. From Lemma 5.1, we have

$$\text{Trace } \gamma_\varphi(\alpha_v) = \text{Trace } \gamma_\varphi(\beta\alpha\beta^{-1}\alpha^{-1}) = \text{Trace } \gamma_\varphi(\beta)\gamma_\varphi(\alpha)\gamma_\varphi(\beta)^{-1}\gamma_\varphi(\alpha)^{-1} \neq 2$$

because γ_φ is injective. Thus we must have $\text{Trace } \gamma_\varphi(\alpha_v) = -2$, which means $h_v = 1$.

Lemma 5.8 *Let the surface be $S_{1,1}$. Then Conjecture 5.4 holds if φ is $\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$.*

Proof We only prove the case when $\varphi = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$ (the proof for the other one is similar). Let τ be the ideal triangulation in Figure 2. Then $\varphi(\tau)$ is the ideal triangulation in Figure 4.

Thus from τ to $\varphi(\tau)$ is relabeling. Suppose the shear parameter for τ is $a^\tau = (a_1, a_2, a_3)$; then the shear parameter for $\varphi(\tau)$ is $a^{\varphi(\tau)} = (a_3, a_1, a_2)$. From $a^\tau = a^{\varphi(\tau)}$, we get $a_1 = a_2 = a_3$. Since we also have $a_1^2 a_2^2 a_3^2 = 1$ and $a_i \in \mathbb{R}_{>0}$, we have $a_1 = a_2 = a_3 = 1$.

Recall that the Chekhov–Fock algebra associated to the ideal triangulation τ is $\mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}]$, where Y_i corresponds to edge e_i for $i = 1, 2, 3$. The algebra $\mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}]$ is generated by Y_1, Y_2 and Y_3 , and subject to the relations

$$Y_1 Y_2 = q^4 Y_2 Y_1, \quad Y_2 Y_3 = q^4 Y_3 Y_2, \quad Y_3 Y_1 = q^4 Y_1 Y_3, \quad Y_i Y_i^{-1} = Y_i^{-1} Y_i = 1.$$

Define the irreducible representation ρ of $\mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}]$ as $\rho_{1,1,1}$ in Lemma 3.9, that is, set $y_1 = y_2 = y_3 = 1$ in (5). Then

$$\rho(Y_1^n) = \text{Id}_V = a_1 \text{Id}_V, \quad \rho(Y_2^n) = \text{Id}_V = a_2 \text{Id}_V, \quad \rho(Y_3^n) = \text{Id}_V = a_3 \text{Id}_V$$

and

$$\rho(H_v) = \rho([Y_1^2 Y_2^2 Y_3^2]) = \text{Id}_V = h_v \text{Id}_V.$$

It is easy to calculate that $\Phi_{\tau\varphi(\tau)}^{q_n} \Psi_{\varphi(\tau)\tau}^{q_n}$ is actually an isomorphism from $\mathbb{C}_{q^4}[Y_1^{\pm 1}, Y_2^{\pm 1}, Y_3^{\pm 1}]$ to itself and

$$\Phi_{\tau\varphi(\tau)}^{q_n} \Psi_{\varphi(\tau)\tau}^{q_n}(Y_1) = Y_3, \quad \Phi_{\tau\varphi(\tau)}^{q_n} \Psi_{\varphi(\tau)\tau}^{q_n}(Y_2) = Y_1, \quad \Phi_{\tau\varphi(\tau)}^{q_n} \Psi_{\varphi(\tau)\tau}^{q_n}(Y_3) = Y_2.$$

We use ρ' to denote the irreducible representation $\rho \Phi_{\tau\varphi(\tau)}^{q_n} \Psi_{\varphi(\tau)\tau}^{q_n}$. Then ρ is isomorphic to ρ' .

For each $0 \leq k \leq n-1$, set

$$v_k = \sum_{0 \leq i \leq n-1} q_n^{k^2+i^2+4ik+i-k} w_i.$$

Then we have

$$\rho'(Y_1)(v_k) = q_n^{4k} v_k, \quad \rho'(Y_2)(v_k) = q_n^{-2k} v_{k+1}, \quad \rho'(Y_3)(v_k) = q_n^{-2k} v_{k-1}.$$

Define invertible operator Λ for V such that $\Lambda(w_k) = v_k$ for all $0 \leq k \leq n-1$. Then, for all $0 \leq k \leq n-1$, we have

$$\rho'(Y_1)(\Lambda(w_k)) = \rho'(Y_1)(v_k) = q_n^{4k} v_k = \Lambda(q_n^{4k} w_k) = \Lambda(\rho(Y_1)w_k).$$

Thus we get $\rho'(Y_1) = \Lambda \circ \rho(Y_1) \circ \Lambda^{-1}$. Similarly, $\rho'(Y_2) = \Lambda \circ \rho(Y_2) \circ \Lambda^{-1}$ and $\rho'(Y_3) = \Lambda \circ \rho(Y_3) \circ \Lambda^{-1}$. Thus Λ is the intertwiner. As a matrix, we have $\Lambda_{i,k} = q_n^{k^2+i^2+4ik+i-k}$.

From direct calculation, we get $|\det(\Lambda)| = n^{n/2}$. Thus we can set $\bar{\Lambda}_{\varphi, \bar{r}_\varphi}^{q_n} = n^{-1/2} \Lambda$. Then

$$|\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| = |\text{Trace } \bar{\Lambda}_{\varphi, \bar{r}_\varphi}^{q_n}| = n^{-1/2} \sum_{0 \leq i \leq n-1} q_n^{6i^2} = n^{-1/2} \sqrt{(6, n)n} = \sqrt{(6, n)}.$$

Obviously we get

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| = 0. \quad \square$$

Remark 5.9 In the proof of [Lemma 5.8](#), when we try to find the periodic edge weight system for the triangulation sweep τ , $\varphi(\tau)$, we require $a_i \in \mathbb{R}_{>0}$ because we want to get the fixed character corresponding to a point in the Teichmüller space. Actually we still get the same intertwiner Λ as in [Lemma 5.8](#) without requiring $a_i \in \mathbb{R}_{>0}$, that is, for any periodic edge weight system, the intertwiner we get is Λ . This means [Lemma 5.8](#) still holds when we choose any other φ -invariant smooth character (without the restriction for only choosing the one corresponding to a fixed point in the Teichmüller space). Readers can check the same arguments hold for [Theorems 5.6](#) and [5.13](#).

Let ϕ be a pseudo-Anosov map for S , and let f be any diffeomorphism for surface S . Then $f\phi f^{-1}$ is also a pseudo-Anosov map. Then we have the following conclusion:

Lemma 5.10 *Let ϕ be any pseudo-Anosov map for S , and let f be any diffeomorphism for S . If [Conjecture 2.3](#) holds for ϕ , then it also holds for $f\phi f^{-1}$.*

Proof We will use the same notation as in [Conjecture 2.3](#). Let f_*^{-1} be the isomorphism from $\pi_1(S)$ to itself induced by f^{-1} . Then $[\gamma f_*^{-1}]$ is a smooth $f\phi f^{-1}$ -invariant character. Set $\theta'_v = \theta_{f^{-1}(v)}$; then θ'_v are invariant under the action of $f\phi f^{-1}$ and

$$\text{Trace } \gamma f_*^{-1}(\alpha_v) = \text{Trace } \gamma(\alpha_{f^{-1}(v)}) = -e^{\theta_{f^{-1}(v)}} - e^{-\theta_{f^{-1}(v)}} = -e^{\theta'_v} - e^{-\theta'_v}.$$

Set $p'_v = e^{\theta'_v/n} + e^{-\theta'_v/n} = e^{\theta_{f^{-1}(v)/n}} + e^{-\theta_{f^{-1}(v)/n}} = p_{f^{-1}(v)}$; then

$$T_n(p'_v) = -\text{Trace } \gamma f_*^{-1}(\alpha_v).$$

Recall that we use $f_{\#}^{-1}$ to denote the isomorphism from $\mathrm{SK}_{q^{1/2}}(S)$ to itself induced by f^{-1} . Let ρ be the irreducible representation associated to $[\gamma]$ and puncture weights p_v . Then $\rho f_{\#}^{-1}$ is an irreducible representation associated to the character $[\gamma f_{*}^{-1}]$ and puncture weights p'_v .

With the assumption for [Conjecture 2.3](#), we have

$$\rho \phi_{\#}(X) = \Lambda_{\phi, r}^{q_n} \circ \rho(X) \circ (\Lambda_{\phi, r}^{q_n})^{-1}$$

for any element $X \in \mathrm{SK}_{q^{1/2}}(S)$ and $|\det(\Lambda_{\phi, r}^{q_n})| = 1$. Then we get

$$\rho f_{\#}^{-1}(f \phi f^{-1})_{\#}(X) = \rho f_{\#}^{-1} f_{\#} \phi_{\#} f_{\#}^{-1}(X) = \rho \phi_{\#} f_{\#}^{-1}(X) = \Lambda_{\phi, r}^{q_n} \circ \rho(f_{\#}^{-1}(X)) \circ (\Lambda_{\phi, r}^{q_n})^{-1}.$$

Thus we get $\Lambda_{f \phi f^{-1}, r f_{*}^{-1}}^{q_n} = \Lambda_{\phi, r}^{q_n}$, and

$$\begin{aligned} \lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\mathrm{Trace} \Lambda_{f \phi f^{-1}, r f_{*}^{-1}}^{q_n}| &= \lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\mathrm{Trace} \Lambda_{\phi, r}^{q_n}| \\ &= \frac{1}{4\pi} \mathrm{vol}_{\mathrm{hyp}}(M_{\phi}) = \frac{1}{4\pi} \mathrm{vol}_{\mathrm{hyp}}(M_{f \phi f^{-1}}). \end{aligned} \quad \square$$

From [\[8\]](#), we know [Conjecture 2.3](#) holds for $\phi = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

Corollary 5.11 [Conjecture 2.3](#) holds for all $f \phi f^{-1}$ where f is any element in $\mathrm{GL}(2, \mathbb{Z})$.

Let φ be a periodic map for S , and let g be any diffeomorphism for S . Then $g \varphi g^{-1}$ is also a periodic map. The same discussion as in [Lemma 5.10](#) implies the following conclusion.

Lemma 5.12 Let φ be any periodic map for S , and let g be any diffeomorphism for S . If [Conjecture 5.4](#) holds for φ , then it also holds for $g \varphi g^{-1}$.

The following theorem shows [Conjecture 5.4](#) holds for the once punctured torus. This confirms the relation between the intertwiner and the simplicial volume of the corresponding mapping torus.

Theorem 5.13 [Conjecture 5.4](#) holds for the once punctured torus.

Proof Let φ be any periodic map for $S_{1,1}$. Then the order of φ could be 1, 2, 3, 4 or 6. According to [Theorem 5.6](#), [Conjecture 5.4](#) holds if the order of φ is 2 or 4.

If the order of φ is 1, then φ is just the identity map. In this case, we can just choose the intertwiner to be the identity operator. Then [Conjecture 5.4](#) holds trivially.

We look at the case when the order of φ is 3 or 6. For these two cases, we have $|\mathrm{Trace} \varphi| = 1$. According to [\[17\]](#), we know there exists an element $g \in \mathrm{GL}(2, \mathbb{Z})$ such that $\varphi = g \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} g^{-1}$ or $\varphi = g \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} g^{-1}$. By [Lemmas 5.8](#) and [5.12](#), [Conjecture 5.4](#) holds for these two cases. $\square$

Remark 5.14 From the proof of [Theorem 5.6](#), we know if we can show $\text{Trace } \Lambda_{\varphi, r}^{q_n} \neq 0$ after n is big enough, then we can prove

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \Lambda_{\varphi, r_\varphi}^{q_n}| = 0.$$

Remark 5.15 From [Section 2.4](#), we know the periodic edge weight system $a = a^{(0)}, a^{(1)}, \dots, a^{(k)} = a$ for the ideal triangulation sweep $\tau = \tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)} = \varphi(\tau)$ and φ -invariant puncture weights h_v can give us the intertwiner $\bar{\Lambda}_{\varphi, \bar{\gamma}}^q$ such that

$$\bar{\rho} \circ \Phi_{\tau\varphi(\tau)}^q \circ \Psi_{\varphi(\tau)\tau}^q(X) = \bar{\Lambda}_{\varphi, \bar{\gamma}}^q \circ \bar{\rho}(X) \circ (\bar{\Lambda}_{\varphi, \bar{\gamma}}^q)^{-1}$$

for every $X \in \mathcal{T}_\tau^q$.

It is easy to verify that $(\Phi_{\tau\varphi(\tau)}^q \circ \Psi_{\varphi(\tau)\tau}^q)^m = (\Phi_{\tau\varphi^m(\tau)}^q \circ \Psi_{\varphi^m(\tau)\tau}^q)$, and

$$a = a^{(0)}, a^{(1)}, \dots, a^{(k)}, \dots, a^{(0)}, a^{(1)}, \dots, a^{(k)} = a$$

is the periodic edge weight system for the ideal triangulation sweep

$$\tau = \tau^{(0)}, \tau^{(1)}, \dots, \tau^{(k)} = \varphi(\tau), \varphi(\tau^{(1)}), \dots, \varphi(\tau^{(k)}) = \varphi^2(\tau), \dots$$

$$\dots, \varphi^{m-1}(\tau) = \varphi^{m-1}(\tau^{(0)}), \varphi^{m-1}(\tau^{(1)}), \dots, \varphi^{m-1}(\tau^{(k)}) = \varphi^m(\tau)$$

and $h_{\varphi^m(v)} = h_v$.

Suppose φ is periodic with order m , then

$$(\bar{\Lambda}_{\varphi, \bar{\gamma}}^q)^m \circ \bar{\rho}(X) \circ (\bar{\Lambda}_{\varphi, \bar{\gamma}}^q)^{-m} = \bar{\rho} \circ \Phi_{\tau\varphi^m(\tau)}^q \circ \Psi_{\varphi^m(\tau)\tau}^q(X) = \bar{\rho}(X)$$

for every $X \in \mathcal{T}_\tau^q$. Then $(\bar{\Lambda}_{\varphi, \bar{\gamma}}^q)^m$ is a scalar matrix since $\bar{\rho}$ is irreducible. Actually we can choose good $\bar{\Lambda}_{\varphi, \bar{\gamma}}^q$ such that $(\bar{\Lambda}_{\varphi, \bar{\gamma}}^q)^m$ is the identity matrix. We have all the eigenvalues of $\bar{\Lambda}_{\varphi, \bar{\gamma}}^q$ are m^{th} roots of unity, and $|\text{Trace } \bar{\Lambda}_{\varphi, \bar{\gamma}}^q| = 0$ or $|\text{Trace } \bar{\Lambda}_{\varphi, \bar{\gamma}}^q| \geq n^{-m}$.

From [\[7, Lemma 11\]](#), we know the complex dimension of the space of all periodic edge weight systems for the fixed ideal triangulation is more than or equal to 1. Thus this space is connected. In a local open subset of this space, we can choose φ -invariant puncture weights such that these puncture weights smoothly vary according to periodic edge weight systems. Then $|\text{Trace } \bar{\Lambda}_{\varphi, \bar{\gamma}}^q|$ smoothly varies according to periodic edge weight systems in a local open subset by using the similar argument in [\[6, Complement 10\]](#). Since this space is connected and 0 is an isolated point in the image, we have $|\text{Trace } \bar{\Lambda}_{\varphi, \bar{\gamma}}^q| = 0$ for all periodic edge weight systems with the chosen puncture weights, or $|\text{Trace } \bar{\Lambda}_{\varphi, \bar{\gamma}}^q| \geq n^{-m}$ for all periodic edge weight systems with the chosen puncture weights.

If we can find one periodic edge weight system with the chosen puncture weights such that

$$\lim_{\text{odd } n \rightarrow \infty} \frac{1}{n} \log |\text{Trace } \bar{\Lambda}_{\varphi, \bar{\gamma}}^{q_n}| = 0,$$

we can conclude that the above equation is true for every periodic edge weight system with the chosen puncture weights.

References

- [1] **F Bonahon, X Liu**, *Representations of the quantum Teichmüller space and invariants of surface diffeomorphisms*, *Geom. Topol.* 11 (2007) 889–937 [MR](#) [Zbl](#)
- [2] **F Bonahon, H Wong**, *Quantum traces for representations of surface groups in $SL_2(\mathbb{C})$* , *Geom. Topol.* 15 (2011) 1569–1615 [MR](#) [Zbl](#)
- [3] **F Bonahon, H Wong**, *Representations of the Kauffman bracket skein algebra, I: Invariants and miraculous cancellations*, *Invent. Math.* 204 (2016) 195–243 [MR](#) [Zbl](#)
- [4] **F Bonahon, H Wong**, *The Witten–Reshetikhin–Turaev representation of the Kauffman bracket skein algebra*, *Proc. Amer. Math. Soc.* 144 (2016) 2711–2724 [MR](#) [Zbl](#)
- [5] **F Bonahon, H Wong**, *Representations of the Kauffman bracket skein algebra, II: Punctured surfaces*, *Algebr. Geom. Topol.* 17 (2017) 3399–3434 [MR](#) [Zbl](#)
- [6] **F Bonahon, H Wong**, *Representations of the Kauffman bracket skein algebra, III: Closed surfaces and naturality*, *Quantum Topol.* 10 (2019) 325–398 [MR](#) [Zbl](#)
- [7] **F Bonahon, H Wong, T Yang**, *Asymptotics of quantum invariants of surface diffeomorphisms, I: Conjecture and algebraic computations*, preprint (2021) [arXiv 2112.12852](#)
- [8] **F Bonahon, H Wong, T Yang**, *Asymptotics of quantum invariants of surface diffeomorphisms, II: The figure-eight knot complement*, preprint (2022) [arXiv 2203.05730](#)
- [9] **D Bullock**, *Rings of $SL_2(\mathbb{C})$ -characters and the Kauffman bracket skein module*, *Comment. Math. Helv.* 72 (1997) 521–542 [MR](#) [Zbl](#)
- [10] **D Bullock, J H Przytycki**, *Multiplicative structure of Kauffman bracket skein module quantizations*, *Proc. Amer. Math. Soc.* 128 (2000) 923–931 [MR](#) [Zbl](#)
- [11] **M Conder**, *Discrete and free subgroups of SL_2* , PhD thesis, University of Cambridge (2020) Available at <https://doi.org/10.17863/CAM.60353>
- [12] **M Culler, P B Shalen**, *Varieties of group representations and splittings of 3-manifolds*, *Ann. of Math.* 117 (1983) 109–146 [MR](#) [Zbl](#)
- [13] **B Farb, D Margalit**, *A primer on mapping class groups*, Princeton Math. Ser. 49, Princeton Univ. Press (2012) [MR](#) [Zbl](#)
- [14] **C Frohman, R Gelca**, *Skein modules and the noncommutative torus*, *Trans. Amer. Math. Soc.* 352 (2000) 4877–4888 [MR](#) [Zbl](#)
- [15] **C Frohman, J Kania-Bartoszyńska, T Lê**, *Unicity for representations of the Kauffman bracket skein algebra*, *Invent. Math.* 215 (2019) 609–650 [MR](#) [Zbl](#)
- [16] **I Ganey, D Jordan, P Safronov**, *The quantum Frobenius for character varieties and multiplicative quiver varieties*, *J. Eur. Math. Soc.* (online publication February 2024)
- [17] **O N Karpenkov**, *Continued fractions and $SL(2, \mathbb{Z})$ conjugacy classes: elements of Gauss reduction theory*, from “Geometry of continued fractions”, *Algor. Computat. Math.* 26, Springer (2013) 67–85 [Zbl](#)
- [18] **H Karuo, J Korinman**, *Azumaya loci of skein algebras*, preprint (2022) [arXiv 2211.13700](#)
- [19] **R M Kashaev**, *The hyperbolic volume of knots from the quantum dilogarithm*, *Lett. Math. Phys.* 39 (1997) 269–275 [MR](#) [Zbl](#)
- [20] **T Y Lam, K H Leung**, *On vanishing sums of roots of unity*, *J. Algebra* 224 (2000) 91–109 [MR](#) [Zbl](#)

- [21] **X Liu**, *The quantum Teichmüller space as a noncommutative algebraic object*, J. Knot Theory Ramifications 18 (2009) 705–726 [MR](#) [Zbl](#)
- [22] **H Murakami, J Murakami**, *The colored Jones polynomials and the simplicial volume of a knot*, Acta Math. 186 (2001) 85–104 [MR](#) [Zbl](#)
- [23] **H Murakami, T Ohtsuki, M Okada**, *Invariants of three-manifolds derived from linking matrices of framed links*, Osaka J. Math. 29 (1992) 545–572 [MR](#) [Zbl](#)
- [24] **G Myerson**, *How small can a sum of roots of unity be?*, Amer. Math. Monthly 93 (1986) 457–459 [MR](#) [Zbl](#)
- [25] **J H Przytycki, A S Sikora**, *On skein algebras and $Sl_2(\mathbb{C})$ -character varieties*, Topology 39 (2000) 115–148 [MR](#) [Zbl](#)
- [26] **N Takenov**, *Representations of the Kauffman skein algebra of small surfaces*, preprint (2015) [arXiv 1504.04573](#)
- [27] **T Tao**, *How small can a sum of a few roots of unity be?*, MathOverflow question (2010) Available at <https://mathoverflow.net/q/46068>

Bernoulli Institute, University of Groningen
Groningen, Netherlands

zhihao003@e.ntu.edu.sg

Received: 9 February 2023 Revised: 15 February 2024

Horizontal decompositions, II

PAOLO LISCA
ANDREA PARMA

We complete the classification of the smooth, closed, oriented 4-manifolds having Euler characteristic less than four and a horizontal handlebody decomposition of genus one. We use the classification result to find a large family of rational homology ball smoothings of cyclic quotient singularities which can be smoothly embedded into the complex projective plane. Our family contains all such rational balls previously known to embed into $\mathbb{CP}^2$ and infinitely many more. We also show that a rational ball of our family admits an almost-complex embedding in $\mathbb{CP}^2$ if and only if it admits a symplectic embedding.

[57K40](#); [57K43](#), [57R17](#)

1 Introduction

Let $B_{p,q}$, where $p > q \geq 0$ are coprime integers, be the rational homology ball smoothing of the cyclic quotient singularity $\frac{1}{p^2}(pq - 1, 1)$. The rational balls $B_{p,q}$ were used in the smooth rational blow-down construction of Fintushel and Stern [4] and in its symplectic counterpart by Symington [17]. They played an important role in the construction of exotic 4-manifolds, starting with the papers of Fintushel and Stern [5] and Park [15]. The problem of embedding the $B_{p,q}$'s, smoothly or symplectically, in a 4-manifold was considered by Khodorovskiy [8; 9; 10]. As pointed out by Evans and Smith [3, Section 2], the existence of holomorphic embeddings $B_{p,q} \hookrightarrow \mathbb{CP}^2$ for certain (p, q) 's follows from results in algebraic geometry — see Hacking and Prokhorov [7] — and explicit symplectic embeddings can be constructed. Moreover, Evans and Smith [3] showed that no further symplectic embeddings exist beyond those which are already known. Owens [14] showed that infinitely many $B_{p,q}$'s admit smooth embeddings in $\mathbb{CP}^2$ but by [3] they cannot be symplectically embedded. The authors extended Owens' family [11] and proved the nonexistence of almost complex embeddings [13] without relying on [3]. The present paper is a natural continuation of [12], where we introduced certain handlebody decompositions that we call horizontal, classified the closed 4-manifolds with the simplest horizontal decompositions and in doing so we recovered infinitely many of the known smooth embeddings $B_{p,q} \hookrightarrow \mathbb{CP}^2$.

This paper consists of two parts. In the first part we complete the classification of the smooth, closed, oriented 4-manifolds X with Euler characteristic $\chi(X) \leq 3$ and a horizontal decomposition of genus 1. In the second part we exploit the classification result to determine the rational balls $B_{p,q}$ which admit smooth embeddings in $\mathbb{CP}^2$ induced by a genus-1 horizontal decomposition. It turns out that in this way

we are able to find the largest family so far of $B_{p,q}$'s smoothly embedded into the complex projective plane. The family contains all such rational balls previously known to embed into $\mathbb{CP}^2$ and many more. We show that each rational ball of the family admits a symplectic embedding in $\mathbb{CP}^2$ if and only if it admits an almost-complex embedding.

In order to state our results we need to recall the definition of a horizontal decomposition [12]. Let $\mathcal{H}$ be a handlebody decomposition of a smooth, oriented 4-dimensional cobordism $X: \partial_- X \rightarrow \partial_+ X$. Suppose that $\mathcal{H}$ has h_i handles of index i and let $\partial_{\pm}^{h_i} X = \partial_{\pm} X \#^{h_i} S^1 \times S^2$ for $i \in \{1, 3\}$. Let $X_L: \partial_-^{h_1} X \rightarrow \partial_+^{h_3} X$ be the cobordism obtained by attaching the 2-handles along a framed link $L = \bigcup_{i=1}^{h_2} \gamma_i \subset \partial_-^{h_1} X$. According to [12], the link L and the decomposition $\mathcal{H}$ are called *horizontal* if, for some Heegaard decomposition $H_g \cup \Sigma_g \times [0, 1] \cup H'_g$ of $\partial_-^{h_1} X$, we have

- $\gamma_i \subset \mathfrak{S}_i := \Sigma_g \times \{t_i\}$ for some $0 < t_1 < \dots < t_{h_2} < 1$;
- the i^{th} 2-handle of $\mathcal{H}$ is attached to γ_i with framing $\text{fr}(\gamma_i) = \text{fr}_{\mathfrak{S}_i}(\gamma_i) \pm 1$.

Note that the 3-manifold $\partial_+^{h_3} X$ is obtained from $\partial_-^{h_1} X$ by 3-dimensional surgery on each γ_i with the same framing $\text{fr}(\gamma_i)$. We say that $\mathcal{H}$ is a horizontal decomposition of type (g, h_1, h_2, h_3) , or a (g, h_1, h_2, h_3) -decomposition. As observed in [12, Section 1], the inequality $g \geq \max\{h_1, h_3\}$ always holds. We say that a smooth, closed 4-manifold $\hat{X}$ has a horizontal decomposition if $X := \hat{X} \setminus \{B^4 \cup B^4\}$ has a such a decomposition when viewed as a cobordism $S^3 \rightarrow S^3$.

Together with [12, Theorem 1.7], the following result yields a complete classification of the smooth, closed, oriented 4-manifolds $\hat{X}$ having Euler characteristic $\chi(\hat{X}) \leq 3$ and a horizontal decomposition of genus 1.

Theorem 1.1 *Let $\hat{X}$ be a smooth, closed 4-manifold admitting a $(1, h_1, h_2, 1)$ -decomposition. Then $\hat{X}$ is diffeomorphic, possibly after reversing orientation, to*

$$S^4, \quad S^1 \times S^3, \quad \pm \mathbb{CP}^2 \# S^1 \times S^3, \quad \pm \mathbb{CP}^2 \# \pm \mathbb{CP}^2 \# S^1 \times S^3 \quad \text{or} \quad B_{p,q} \cup -B_{p,q} \quad \text{if } \chi(\hat{X}) \leq 2,$$

$$\mathbb{CP}^2, \quad \mathbb{CP}^2 \# B_{p,q} \cup -B_{p,q}, \quad \mathbb{CP}^2 \# 2\overline{\mathbb{CP}}^2 \# S^1 \times S^3 \quad \text{or} \quad 3\mathbb{CP}^2 \# S^1 \times S^3 \quad \text{if } \chi(\hat{X}) = 3,$$

for some $p > q \geq 0$.

From now on we concentrate on $(1, h_1, h_2, 1)$ -decompositions of a 4-manifold $\hat{X}$ assuming $\hat{X} \cong \pm \mathbb{CP}^2$. Then we have $3 = \chi(\hat{X}) = 1 - h_1 + h_2$. Since $h_1 \leq g = 1$, the pair (h_1, h_2) is equal to either $(0, 2)$ or $(1, 3)$. The proof of [12, Lemma 5.1] shows that a $(1, 1, 3, 1)$ -decomposition of a 4-manifold X induces a smooth embedding in X of the disjoint union of three 4-manifolds, each consisting of a 1-handle and a 2-handle attached to $\partial_+(S^1 \times D^3 \setminus B^4) = S^1 \times S^2$ along a simple closed curve sitting on the standard genus-1 Heegaard torus. Using this fact, in the following Theorem 1.2 we describe essentially all the smooth embeddings of $B_{p,q}$'s induced by $(1, 1, 3, 1)$ -decompositions of $\mathbb{CP}^2$.

We introduce some notation in order to give a more concise statement of Theorem 1.2. When we write $B_{p,q}$ we shall always assume that p and q are coprime integers, but not necessarily satisfying $p > q \geq 0$.

When $p \neq 0$, by $B_{p,q}$ we shall denote the rational homology ball $B_{|p|,\bar{q}}$, where $\bar{q}$ is the unique integer congruent to q modulo $|p|$ and satisfying $|p| > \bar{q} \geq 0$. Note that if $p = \pm 1$ then $q = 0$ and $B_{\pm 1,0} = B^4$. If $p = 0$ then $q = \pm 1$ and by $B_{p,q}$ we shall denote $S^1 \times D^3 \# q\mathbb{CP}^2$.

Let $\{x_n(a, b)\}_{n \geq 0}$ be the integer sequence defined by

$$x_0(a, b) = a, \quad x_1(a, b) = b, \quad x_{n+2}(a, b) = x_{n+1}(a, b) + x_n(a, b), \quad n \geq 2,$$

and denote by $F_n := x_n(0, 1)$ and $L_n := x_n(2, 1)$ the n^{th} Fibonacci and Lucas number, respectively.

Theorem 1.2 A disjoint union $\pm B_{p_1,q_1} \cup \pm B_{p_2,q_2} \cup \pm B_{p_3,q_3}$ with $\max\{|p_i|\} > 1$ is induced by a $(1, 1, 3, 1)$ -decomposition of $\mathbb{CP}^2$ if and only if it coincides with one of the following:

- (1) $B_{p_1,q_1} \cup B_{p_2,q_2} \cup B_{p_3,q_3}$, where (p_1, p_2, p_3) is a Markov triple and $p_k q_i \equiv \pm 3p_j \pmod{p_i}$, where (i, j, k) is a permutation of $(1, 2, 3)$;
- (2) $B_{x_1,q_1} \cup -B_{x_2,q_2} \cup -B_{x_3,q_3}$, where $x_1^2 + x_1 x_2 x_3 = x_2^2 + x_3^2$,
 $x_2 q_1 \equiv \pm x_3 \pmod{x_1}$, $x_3 q_2 \equiv \pm x_1 \pmod{(x_1 x_3 + x_2)}$, and $x_2 q_3 \equiv \pm x_1 \pmod{x_3}$;
- (3) $B_{F_a,q_1} \cup -B_{F_{b-\varepsilon a},q_2} \cup B_{F_b,q_3}$ where a and b are odd, coprime integers, $\varepsilon \in \{\pm 1\}$, $b \neq \varepsilon a$,
 $q_1 \equiv \pm L_b / F_{b-\varepsilon a} \pmod{F_a}$, $q_2 \equiv \pm L_b / F_a \pmod{F_{b-\varepsilon a}}$ and $q_3 \equiv \pm L_a / F_{b-\varepsilon a} \pmod{F_b}$.

Remark 1.3 The fact that the disjoint unions of Theorem 1.2(1) embed in $\mathbb{CP}^2$ is already known [7].

Denote by $\mathcal{F}_i$, for $i = 1, 2, 3$, the family of rational homology balls $B_{p,q}$ which admit an orientation-preserving embedding in $\mathbb{CP}^2$ in view of Theorem 1.2(1).

Theorem 1.4 Suppose that $|p| > 1$ and the rational ball $B_{p,q}$ admits an orientation-preserving embedding in $\mathbb{CP}^2$ induced by a $(1, h_1, h_2, 1)$ -decomposition. Then:

- (1) $B_{p,q} \in \mathcal{F}_1$ if and only if $q^2 \equiv -9 \pmod{p}$.
- (2) $B_{p,q} \in \mathcal{F}_2$ if and only if $q^2 \equiv -1 \pmod{p}$.
- (3) Let

$$S_p = \{x \in \mathbb{Z}/p\mathbb{Z} \mid x \equiv q^2 \pmod{p} \text{ for some } B_{p,q} \in \mathcal{F}_3\}.$$

Then, the sequence $\{|S_p|\}_p$ is unbounded.

Remark 1.5 In [11] we constructed infinitely many embeddings $B_{p,q} \hookrightarrow \mathbb{CP}^2$, extending a family of embeddings constructed by Owens [14]. It is not hard to check that our constructions implicitly used horizontal decompositions of $\mathbb{CP}^2$ — in fact, those constructions led us to the discovery of horizontal decompositions. At the end of [11, Section 3] we showed that all the balls shown to embed in $\mathbb{CP}^2$ satisfy $q^2 \equiv -1 \pmod{p}$. In view of Remark 1.3, Theorem 1.4 implies that all $B_{p,q}$'s known so far to embed in $\mathbb{CP}^2$ belong to $\mathcal{F}_1 \cup \mathcal{F}_2$ and that $(1, h_1, h_2, 1)$ -decompositions of $\mathbb{CP}^2$ induce infinitely many more embeddings of $B_{p,q}$'s.

Corollary 1.6 *Let $B_{p,q}$ be a rational ball which admits an orientation-preserving embedding in $\mathbb{CP}^2$ induced by a $(1, h_1, h_2, 1)$ -decomposition. Then $B_{p,q}$ admits an almost-complex embedding in $\mathbb{CP}^2$ if and only if it embeds symplectically in $\mathbb{CP}^2$ and belongs to $\mathcal{F}_1$.*

Proof By [13, Corollary 1.3], if $B_{p,q}$ admits an almost-complex embedding in $\mathbb{CP}^2$ then $q \equiv -9 \pmod{p}$; therefore by Theorem 1.4 $B_{p,q} \in \mathcal{F}_1$, and by Remark 1.3 it embeds symplectically in $\mathbb{CP}^2$. The converse follows from the well-known fact that symplectic embeddings are almost-complex. $\square$

The paper is organized as follows. In Section 2 we prove Theorem 1.1 when $\chi(\hat{X}) \leq 2$ and in Section 3 we prove Theorem 1.1 when $\chi(\hat{X}) = 3$. In Section 4 we prove Theorem 1.2 and in Section 5 we prove Theorem 1.4.

Acknowledgements The authors wish to thank the referee for their accurate report, which helped us to improve the exposition. The present work is part of the MIUR-PRIN projects 2017JZ2SW5 and 2022NMPLT8. Lisca is a member of the GNSAGA research group (INdAM).

2 Proof of Theorem 1.1 when $\chi(\hat{X}) \leq 2$

Let $\hat{X}$ be a smooth, oriented cobordism admitting a $(1, h_1, h_2, 1)$ -decomposition. In this section we prove the first part of Theorem 1.1, that is we show that if $\chi(\hat{X}) \leq 2$ then $\hat{X}$ is diffeomorphic, possibly after reversing orientation, to

$$S^4, \quad S^1 \times S^3, \quad \pm \mathbb{CP}^2 \# S^1 \times S^3, \quad \pm \mathbb{CP}^2 \# \pm \mathbb{CP}^2 \# S^1 \times S^3 \quad \text{or} \quad B_{p,q} \cup -B_{p,q}$$

for some $p > q \geq 0$.

We recall that to an n -component link $L \subset S^1 \times S^2$, horizontal with respect to the standard genus-1 Heegaard decomposition of $S^1 \times S^2$, one can associate an element of mapping class group $\text{Map}(\Sigma_1)$ of the genus-1 surface, factorized as a product of Dehn twists [12, Section 1]. This is done as follows. Let $T \subset S^1 \times S^2$ be the standard Heegaard torus and $\gamma_1, \dots, \gamma_n$ the components of L , viewed as sitting on T . Then, to γ_i we associate the Dehn twist $\tau_i := \tau_{\gamma_i} \in \text{Map}(T)$ for $i = 1, \dots, n$. By definition, the factorization associated to L is

$$F_L = (\tau_n^{\delta_n}, \dots, \tau_1^{\delta_1}),$$

where $\delta_i = \text{fr}_{\mathcal{S}_i}(\gamma_i) - \text{fr}(\gamma_i) \in \{\pm 1\}$. The product $m_L = \tau_n^{\delta_n} \cdots \tau_1^{\delta_1}$ is called the *monodromy* of L .

Now we prove an auxiliary result, which will be also used later on. Recall (see eg [16, Section 1.2]) that each handlebody decomposition $\mathcal{H}$ on a smooth, oriented cobordism $X: \partial_- X \rightarrow \partial_+ X$ can be viewed as coming from a Morse function $f: X \rightarrow [0, 1]$. The handlebody decomposition $\overline{\mathcal{H}}$ associated with $1 - f: X \rightarrow [0, 1]$ is sometimes called the *dual* of $\mathcal{H}$. When X has dimension 4 each i -handle of $\mathcal{H}$ corresponds to a $4 - i$ -handle of $\overline{\mathcal{H}}$. In particular, the cocores of the 2-handles of $\mathcal{H}$ are the cores of the 2-handles of $\overline{\mathcal{H}}$, and vice versa.

Lemma 2.1 *Let $X: \partial_- X \rightarrow \partial_+ X$ be a smooth, oriented 4-dimensional cobordism with a horizontal handlebody decomposition $\mathcal{H}$. Let $\bar{\mathcal{H}}$ be the dual handlebody decomposition on X , viewed as an oriented cobordism $\bar{X}: -\partial_+ X \rightarrow -\partial_- X$. Then $\bar{\mathcal{H}}$ can be made horizontal via a framed isotopy of the attaching circles of its 2-handles. Moreover, if $\mathcal{H}$ is of type (g, h_1, h_2, h_3) then $\bar{\mathcal{H}}$ is of type (g, h_3, h_2, h_1) .*

Proof We assume that we are in the setup of [Section 1](#). Recall that the i^{th} 2-handle of $\mathcal{H}$ is attached to γ_i with framing $\text{fr}(\gamma_i) = \text{fr}_{\Sigma_i}(\gamma_i) \pm 1$ and as a result the 3-manifold $\partial_+^{h_3} X$ is obtained from $\partial_+^{h_1} X$ by 3-dimensional surgery on each γ_i with framing $\text{fr}(\gamma_i)$. Consider, for each $i = 1, \dots, h_2$, a parallel copy γ'_i of γ_i inside Σ_i . We may assume that the surgeries are performed using tubular neighborhoods of the γ_i 's disjoint from the γ'_i 's, and therefore that $\gamma'_i \subset \partial_+^{h_3} X$. It is easy to check that, once endowed with the framing $\text{fr}(\gamma'_i) := \text{fr}_{\Sigma_i}(\gamma'_i) \mp 1$, the framed knot γ'_i is framed isotopic to a 0-framed meridian of γ_i — also viewed inside $\partial_+^{h_3} X$. In other words, the pair $(\gamma'_i, \text{fr}(\gamma'_i))$ prescribes, up to framed isotopy, how to attach the i^{th} 2-handle of $\bar{\mathcal{H}}$. Clearly the link $L' = \bigcup_{i=1}^{h_2} \gamma'_i \subset \partial_+^{h_3} X$ is horizontal with respect to a Heegaard decomposition of $-\partial_+^{h_3} X$ and the statement follows easily. $\square$

We now continue with the first part of [Theorem 1.1](#). Let X be the complement of two disjoint balls in $\hat{X}$, viewed as a cobordism $S^3 \rightarrow S^3$ and endowed with a horizontal decomposition. Clearly $\chi(X) = h_2 - h_1 - 1$.

Assume first $\chi(X) < 0$. Since $h_1 \leq g = 1$, $\chi(X) < 0$ implies $(h_1, h_2) \in \{(0, 0), (1, 0), (1, 1)\}$. The case $(h_1, h_2) = (0, 0)$ is not possible because a single 3-handle cannot give a cobordism from S^3 to S^3 . If $(h_1, h_2) = (1, 0)$ then $\hat{X}$ is the double of $S^1 \times D^3$, ie $\hat{X} \cong S^1 \times S^3$. If $(h_1, h_2) = (1, 1)$ the 2-handle defines a cobordism $X_\gamma: S^1 \times S^2 \rightarrow S^1 \times S^2$, where $\gamma \subset S^1 \times S^2$ is a simple closed curve sitting on the standard genus-1 Heegaard torus T of $S^1 \times S^2$ and $\text{fr}_T(\gamma) \pm 1$. Then, γ is the boundary of a compressing disc in H_1 and it is easy to check that $\hat{X} \cong \pm \mathbb{C}P^2 \# S^1 \times S^3$.

If $\chi(X) = 0$ the pair (h_1, h_2) is equal to either $(0, 1)$ or $(1, 2)$. In the first case, by [Lemma 2.1](#) the upside-down cobordism $\bar{X}: S^3 \rightarrow S^3$ admits a horizontal decomposition of type $(1, 1, 1, 0)$; therefore by [\[12, Theorem 1\]](#) $\bar{X} \cong S^3 \times [0, 1]$ and $\hat{X} \cong S^4$.

Now suppose $(h_1, h_2) = (1, 2)$. Then, the 2-handles define a cobordism $X_L: S^1 \times S^2 \rightarrow S^1 \times S^2$, where $L \subset S^1 \times S^2$ is a 2-component link, horizontal with respect to the standard genus-1 Heegaard decomposition of $S^1 \times S^2$. The two components γ_1 and γ_2 of L sit on parallel copies of the standard Heegaard torus $T \subset S^1 \times S^2$ and the factorization of L is given by

$$(2-1) \quad F_L = (\tau_2^{\delta_2}, \tau_1^{\delta_1}),$$

where $\delta_i = \text{fr}_T(\gamma_i) - \text{fr}(\gamma_i) \in \{\pm 1\}$ is equal to minus the relative framing of γ_i and the monodromy of L is $m_L = \tau_2^{\delta_2} \tau_1^{\delta_1}$ and $\tau_i := \tau_{\gamma_i}$ for $i = 1, 2$.

Let $\lambda \subset T$ be a simple closed curve which is the boundary of a compressing disc in H_1 . Then (λ, λ) is a Heegaard diagram for $S^1 \times S^2$ and by [\[12, Lemma 3.3\]](#) $(m_L(\lambda), \lambda)$ is a Heegaard diagram for $\partial_+ X_L = S^1 \times S^2$. In order to have $H_1(\partial_+ X_L; \mathbb{Z}) = \mathbb{Z}$ we need $m_L(\lambda) \cdot \lambda = 0$. From now on we

fix arbitrary orientations of γ_1 and γ_2 and we abuse notation by using γ_1 and γ_2 to also denote the corresponding homology classes in $H_1(T; \mathbb{Z})$. Define

$$x := \gamma_1 \cdot \lambda, \quad y := \gamma_2 \cdot \lambda \quad \text{and} \quad n := \gamma_2 \cdot \gamma_1.$$

If $x = y = 0$ we have $\gamma_1 = \gamma_2 = \pm\lambda$ and X has a handlebody decomposition with a 1-handle, two 2-handles attached along ± 1 -framed unlinked trivial knots in $S^1 \times S^2$ and a 3-handle. It is easy to check that $\hat{X}$ is orientation-preserving diffeomorphic to $-\delta_1 \mathbb{C}P^2 \# -\delta_2 \mathbb{C}P^2 \# S^1 \times S^3$. Hence, from now on we assume $(x, y) \neq (0, 0)$. Moreover, we choose the orientations of γ_1 and γ_2 so that $n \geq 0$. Since

$$m_L(\lambda) = \tau_2^{\delta_2}(\tau_1^{\delta_1}(\lambda)) = \tau_2^{\delta_2}(\lambda + \delta_1 x \gamma_1) = \lambda + \delta_1 x \gamma_1 + \delta_2 y \gamma_2 + \delta_1 \delta_2 n x \gamma_2,$$

taking the intersection product with $\delta_1 \delta_2 \lambda$ yields

$$(2-2) \quad \delta_2 x^2 + \delta_1 y^2 + nxy = 0.$$

Equation (2-2) implies that x and y divide each other; thus $|x| = |y| \neq 0$, and dividing by $|xy| = x^2 = y^2$ it follows that

$$\delta_2 + \delta_1 \pm n = 0.$$

We conclude $n \in \{0, 2\}$. If $n = 0$ we have $\gamma_1 = \pm\gamma_2$ and $\delta_1 \delta_2 = -1$; therefore F_L is either of the form $(\tau_\gamma, \tau_\gamma^{-1})$ or $(\tau_\gamma^{-1}, \tau_\gamma)$, and in both cases m_L is the identity. In fact, the components γ_1 and γ_2 of L , which are the attaching circles of the two 2-handles, cobound an annulus A of the form $\gamma \times [0, 1] \subset T \times [0, 1] \subset Y_1$, and they have framings with respect to A which are opposite and equal to 1 in absolute value. Sliding for instance γ_1 over γ_2 changes L into the framed link consisting of γ_2 and a 0-framed meridian of γ_2 . By [6, Example 4.6.3], this shows that $\hat{X}$ is the double of $\hat{X}_{\gamma_2} := S^1 \times D^3 \cup X_{L_2}$. We can view $\hat{X}_{\gamma_2} \setminus B^4$ as a cobordism $S^3 \rightarrow \partial_+ X_{\gamma_2}$ with a horizontal decomposition of type $(1, 1, 1, 0)$. By [12, Proposition 3.1] it follows that $\hat{X}_{\gamma_2} \cong \pm B_{p,q}$ for some $p > q \geq 0$. Therefore $\hat{X} \cong B_{p,q} \cup -B_{p,q}$.

If $n = 2$ we must have $\delta_1 = \delta_2 = -\text{sgn}(xy)$, so that $\gamma_1 = x\mu + a\lambda$ and $\gamma_2 = -\delta_1(x\mu + b\lambda)$ for some $a, b \in \mathbb{Z}$ (both coprime with x), where $\mu, \lambda \subset T$ is a symplectic basis and x can be assumed to be positive up to reversing the orientation of both γ_1 and γ_2 . Now, $n = \gamma_2 \cdot \gamma_1 = 2$ implies $-\delta_1(a - b)x = 2$. We observe that $x = 2$ gives a contradiction, because this would force both a and b to be odd, and the equality $-\delta_1(a - b) = 1$ would be false modulo 2. Therefore $x = 1$ and $b - a = 2\delta_1$. Twisting the 1-handle as in [12, Figure 3] replaces μ with μ plus a multiple of λ . This allows us to assume $\gamma_1 = \mu$ and, consequently, $\gamma_2 = -\delta_1(\mu + (b - a)\lambda) = -\delta_1\mu - 2\lambda$. A Kirby calculus picture of $\hat{X}$ consists of a dotted unknot U , a $-\delta_1$ -framed meridian of U corresponding to $\tau_1^{\delta_1}$ and another meridian of U , corresponding to $\tau_2^{\delta_2}$, whose Seifert framing is $2\delta_1 - \delta_1 = \delta_1$. Sliding one meridian over the other we get a picture for the double of B^4 ; therefore in this case $\hat{X} = S^4$.

3 Proof of Theorem 1.1 when $\chi(\hat{X}) = 3$

The purpose of this section is to prove the following.

Theorem 3.1 Let $\hat{X}$ be a smooth, oriented cobordism admitting a $(1, h_1, h_2, 1)$ -decomposition and such that $\chi(\hat{X}) = 3$. Then $\hat{X}$ is diffeomorphic, possibly after reversing orientation, to

$$\mathbb{CP}^2, \quad \mathbb{CP}^2 \# B_{p,q} \cup -B_{p,q}, \quad \mathbb{CP}^2 \# 2\overline{\mathbb{CP}}^2 \# S^1 \times S^3, \quad \text{or} \quad 3\mathbb{CP}^2 \# S^1 \times S^3$$

for some $p > q \geq 0$.

Let X be the complement of two disjoint balls in $\hat{X}$, viewed as a cobordism $S^3 \rightarrow S^3$ and endowed with the horizontal decomposition. Then $1 = \chi(X) = -h_1 + h_2 - 1$. Since $h_1 \leq g = 1$, the pair (h_1, h_2) is equal to either $(0, 2)$ or $(1, 3)$. In the first case, the 2-handles define a cobordism $X_L: S^3 \rightarrow S^1 \times S^2$, where $L \subset S^3$ is a 2-component link, horizontal with respect to the standard genus-1 Heegaard decomposition of S^3 . By [Lemma 2.1](#) the upside-down cobordism $\bar{X}: S^3 \rightarrow S^3$ admits a horizontal decomposition of type $(1, 1, 2, 0)$; therefore by [\[12, Theorem 1\]](#) $\hat{X} \cong \pm \mathbb{CP}^2$. This proves [Theorem 3.1](#) when $(h_1, h_2) = (0, 2)$.

When $(h_1, h_2) = (1, 3)$ the proof of [Theorem 3.1](#) is much more involved and is organized as follows. In [Section 3.1](#) we prove an auxiliary result and in [Section 3.2](#) we use it to show that to the monodromy factorization of a horizontal decomposition are associated solutions of certain Diophantine equations — [\(3-2\)](#) and [\(3-3\)](#). Recall that, given an n -tuple $F = (g_n, \dots, g_1)$ of elements of a group G , a *Hurwitz move* on F is a transformation of the form

$$F = (\dots, g_{i+1}, g_i, \dots) \mapsto F' = (\dots, g_{i+1}g_i g_{i+1}^{-1}, g_{i+1}, \dots)$$

or

$$F = (\dots, g_{i+1}, g_i, \dots) \mapsto F' = (\dots, g_i, g_i^{-1}g_{i+1}g_i, \dots).$$

In [Section 3.3](#) we show how the solutions associated to [\(3-2\)](#) change under the application of a Hurwitz move to the factorization. Hurwitz moves do not change the cobordism X [\[12, Proposition 4.6\]](#), so the basic idea is to use the equations to find a sequence of Hurwitz moves which simplify the factorization. We also show how sets of solutions corresponding to various parameters are related, we define the notion of weakly minimal solution and we observe that applying Hurwitz moves each solution can be made weakly minimal. As for [\[12, Theorem 1.7\]](#), the inspiration for this proof came from a paper by Denis Auroux [\[1\]](#). In [Section 3.4](#) we determine, for the relevant values of the parameters, the weakly minimal solutions of [\(3-2\)](#). Moreover, for each such solution we recover the corresponding “weakly minimal” factorization. In [Sections 3.5](#) and [3.6](#) we determine the cobordisms represented by the weakly minimal factorizations and identify each closed 4-manifold $\hat{X}$ from the corresponding Kirby diagram. We proceed in a similar way in [Sections 3.7, 3.8](#) and [3.9](#). Equation [\(3-3\)](#) is not used in this section but will be used in [Section 4](#) to prove [Theorem 1.2](#).

3.1 A trace formula

Let $(V, \cdot)$ be a 2-dimensional, real symplectic vector space. Given $v \in V$, denote by $\tau_v: V \rightarrow V$ the transvection defined by

$$\tau_v(w) = w + (v \cdot w)v, \quad w \in V.$$

Note that $\tau_v^n(w) = w + n(v \cdot w)v$ for each $n \in \mathbb{Z}$.

Lemma 3.2 Given $u, v, w \in V$ and $l, n, m \in \mathbb{Z}$, let $\Phi = \tau_u^l \tau_v^n \tau_w^m$. Then

$$nm(v \cdot w)^2 + lm(u \cdot w)^2 + ln(u \cdot v)^2 = lmn(w \cdot u)(w \cdot v)(v \cdot u) + 2 - \text{tr}(\Phi).$$

Proof Once we fix $l, n, m \in \mathbb{Z}$, the two sides of the equality are continuous functions of u, v and w ; therefore it is enough to prove that they agree when v and w are linearly independent. Using $\{v, w\}$ as a basis of V , we easily see that for any endomorphism f ,

$$\text{tr}(f) = \frac{f(v) \cdot w + v \cdot f(w)}{v \cdot w}.$$

If we write $\tau_v^n \tau_w^m(v) = av + bw$ and $\tau_v^n \tau_w^m(w) = cv + dw$, we have

$$\begin{aligned} \tau_u^l \tau_v^n \tau_w^m(v) &= \tau_u^l(av + bw) = a(v + l(u \cdot v)u) + b(w + l(u \cdot w)u), \\ \tau_u^l \tau_v^n \tau_w^m(w) &= \tau_u^l(cv + dw) = c(v + l(u \cdot v)u) + d(w + l(u \cdot w)u), \end{aligned}$$

so

$$\frac{\tau_u^l \tau_v^n \tau_w^m(v) \cdot w + v \cdot \tau_u^l \tau_v^n \tau_w^m(w)}{v \cdot w} = a + d + l \left(\frac{b}{v \cdot w} (u \cdot w)^2 + \frac{a - d}{v \cdot w} (u \cdot v)(u \cdot w) - \frac{c}{v \cdot w} (u \cdot v)^2 \right).$$

Finally, we compute $a = 1 - mn(v \cdot w)^2$, $b = -m(v \cdot w)$, $c = n(v \cdot w)$ and $d = 1$. Plugging in these values we get the desired equality. $\square$

3.2 Diophantine equations

As explained in [12, Section 1], a horizontal link $L = \bigcup_{i=1}^3 \gamma_i \subset S^1 \times S^2$ determines a factorization of the form

$$F_L = (\tau_3^{\delta_3}, \tau_2^{\delta_2}, \tau_1^{\delta_1}),$$

where τ_i , $i = 1, 2, 3$, is a positive Dehn twist along γ_i and $\delta_i = \text{fr}_T(\gamma_i) - \text{fr}(\gamma_i) \in \{\pm 1\}$. The 2-handles produce a cobordism from $S^1 \times S^2$ to $S^1 \times S^2$ and, as observed in Section 2, this happens if and only if $m_L(\lambda) \cdot \lambda = 0$. Let $\mu, \lambda \subset T$ be a symplectic basis of oriented, simple closed curves in the standard Heegaard torus $T \subset S^1 \times S^2$. From now on we fix arbitrary orientations of the γ_i 's and we abuse notation by using γ_i to also denote the corresponding homology classes in $H_1(T; \mathbb{Z})$. We have $\gamma_i = p_i \mu + q_i \lambda$, $i = 1, 2$, for some coprime $p_i, q_i \in \mathbb{Z}$. Let us introduce the variables x_1, x_2 and x_3 by setting

$$(3-1) \quad (x_1, x_2, x_3) = (\gamma_2 \cdot \gamma_3, \gamma_1 \cdot \gamma_3, \gamma_1 \cdot \gamma_2).$$

Then, the trace formula of Lemma 3.2 reads

$$(3-2) \quad \delta_1 x_1^2 + \delta_2 x_2^2 + \delta_3 x_3^2 - x_1 x_2 x_3 = \delta_1 \delta_2 \delta_3 (2 - \text{tr}(m_L)).$$

Moreover, observe that $m_L(\lambda) \cdot \lambda = 0$ implies $m_L(\lambda) = \pm \lambda$; therefore $\det(m_L) = 1$ implies $\text{tr}(m_L) = \pm 2$. Also, we have

$$\begin{aligned} m_L(\lambda) \cdot \lambda &= \tau_3^{\delta_3} \tau_2^{\delta_2} \tau_1^{\delta_1}(\lambda) \\ &= \tau_3^{\delta_3} \tau_2^{\delta_2}(\lambda + \delta_1 p_1 \gamma_1) \end{aligned}$$

$$\begin{aligned}
&= \tau_3^{\delta_3}(\lambda + \delta_1 p_1 \gamma_1 + \delta_2(p_2 - \delta_1 p_1 x_3) \gamma_2) \\
&= \lambda + \delta_1 p_1 \gamma_1 + (\delta_2 p_2 - \delta_1 \delta_2 p_1 x_3) \gamma_2 + \delta_3(p_3 - \delta_1 p_1 x_2 - \delta_2 p_2 x_1 + \delta_1 \delta_2 p_1 x_1 x_3) \gamma_3.
\end{aligned}$$

Therefore, the condition $m_L(\lambda) \cdot \lambda = 0$ translates into

$$(3-3) \quad \delta_1 p_1^2 + \delta_2 p_2^2 + \delta_3 p_3^2 - \delta_1 \delta_2 p_1 p_2 x_3 - \delta_1 \delta_3 p_1 p_3 x_2 - \delta_2 \delta_3 p_2 p_3 x_1 + \delta_1 \delta_2 \delta_3 p_1 p_3 x_1 x_3 = 0.$$

In the following sections we study the following sets of solutions of (3-2):

$$S_a^{\delta_1, \delta_2, \delta_3} := \{(\delta_1, \delta_2, \delta_3, x_1, x_2, x_3) \in \{\pm 1\}^3 \times \mathbb{Z}^3 \mid \delta_1 x_1^2 + \delta_2 x_2^2 + \delta_3 x_3^2 - x_1 x_2 x_3 = \delta_1 \delta_2 \delta_3 a\},$$

where $a := 2 - \text{tr}(m_L)$ is either 0 or 4. We shall provide a proof of [Theorem 3.1](#) in the case $(h_1, h_2) = (1, 3)$ by combining the information on the sets $S_a^{\delta_1, \delta_2, \delta_3}$ with the constraint provided by (3-2). Equation (3-3) will be used in [Section 4](#).

3.3 Factorizations and Hurwitz moves

In this section we show that to understand the sets of solutions $S_a^{\delta_1, \delta_2, \delta_3}$ for all values of the δ_i 's it suffices to study the sets $S_a^{1,1,1}$ and $S_a^{-1,1,1}$ with $a \in \{0, 4\}$. The latter sets will be analyzed in [Sections 3.4](#) and [3.7](#).

Let us start with the observation that changing the orientation of one of the curves γ_i changes the signs of two of the elements in the associated triple (x_1, x_2, x_3) , providing another solution of (3-2). We define the *sign changes* as follows:

$$(x_1, x_2, x_3) \mapsto (-x_1, -x_2, x_3), \quad (x_1, x_2, x_3) \mapsto (-x_1, x_2, -x_3), \quad (x_1, x_2, x_3) \mapsto (x_1, -x_2, -x_3).$$

Next, we observe that applying a Hurwitz move we get

$$m_L = \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} \tau_{\gamma_1}^{\delta_1} = \tau_{\tau_{\delta_3}^{\delta_3}(\gamma_2)}^{\delta_2} \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_1}^{\delta_1} = \tau_{-\tau_{\delta_3}^{\delta_3}(\gamma_2)}^{\delta_2} \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_1}^{\delta_1}.$$

To see the effect on a solution (x_1, x_2, x_3) of (3-2) note that we are replacing $(\gamma_3, \gamma_2, \gamma_1)$ with

$$(-\tau_{\gamma_3}^{\delta_3}(\gamma_2), \gamma_3, \gamma_1)$$

and since $\tau_{\gamma_3}^{\delta_3}(\gamma_2) = \gamma_2 + \delta_3(\gamma_3 \cdot \gamma_2)\gamma_3$ the associated triple becomes $(x_1, \delta_3 x_1 x_2 - x_3, x_2)$. Setting $\hat{x}_3 := \delta_3 x_1 x_2 - x_3$, the triple $(\delta_1, \delta_3, \delta_2, x_1, \hat{x}_3, x_2)$ satisfies (3-2). Similarly, the other Hurwitz moves lead to the solutions

$$(\delta_1, \delta_3, \delta_2, x_1, x_3, \hat{x}_2), \quad (\delta_2, \delta_1, \delta_3, x_2, \hat{x}_1, x_3) \quad \text{and} \quad (\delta_2, \delta_1, \delta_3, \hat{x}_2, x_1, x_3),$$

where $\hat{x}_1 := \delta_1 x_2 x_3 - x_1$ and $\hat{x}_2 := \delta_2 x_1 x_3 - x_2$, which also satisfy (3-2). Call *mutations* the changes

$$\mu_1: (\delta_1, \delta_2, \delta_3, x_1, x_2, x_3) \mapsto (\delta_1, \delta_2, \delta_3, \hat{x}_1, x_2, x_3),$$

$$\mu_2: (\delta_1, \delta_2, \delta_3, x_1, x_2, x_3) \mapsto (\delta_1, \delta_2, \delta_3, x_1, \hat{x}_2, x_3),$$

$$\mu_3: (\delta_1, \delta_2, \delta_3, x_1, x_2, x_3) \mapsto (\delta_1, \delta_2, \delta_3, x_1, x_2, \hat{x}_3),$$

and, for any permutation $\sigma \in S_3$, set

$$\sigma(\delta, \mathbf{x}) := (\delta_{\sigma(1)}, \delta_{\sigma(2)}, \delta_{\sigma(3)}, x_{\sigma(1)}, x_{\sigma(2)}, x_{\sigma(3)}).$$

Then we see that the Hurwitz moves induce the transformations

$$(23) \circ \mu_3, \quad (23) \circ \mu_2, \quad (12) \circ \mu_1 \quad \text{and} \quad (12) \circ \mu_2.$$

In particular, each of the three mutations is realized, up to permutations, by at least one Hurwitz move.

Definition 3.3 A solution $(\delta_1, \delta_2, \delta_3, x_1, x_2, x_3)$ of (3-2) is *minimal* (resp. *weakly minimal*) if $|x_i| < |\hat{x}_i|$ (resp. $|x_i| \leq |\hat{x}_i|$) for each $i = 1, 2, 3$. A factorization $(\tau_{\gamma_3}^{\delta_3}, \tau_{\gamma_2}^{\delta_2}, \tau_{\gamma_1}^{\delta_1})$ is *minimal* or *weakly minimal* if so is the corresponding solution of (3-2).

Observe that any solution of (3-2) can be modified into a weakly minimal solution by a (possibly empty) sequence of mutations, each of which makes the quantity $|x_1| + |x_2| + |x_3|$ decrease strictly.

Now suppose that $(\delta, \mathbf{x}) = (\delta_1, \delta_2, \delta_3, x_1, x_2, x_3)$ is a solution of (3-2) and let

$$-(\delta, \mathbf{x}) = (-\delta_1, -\delta_2, -\delta_3, -x_1, -x_2, -x_3).$$

If we represent $S^1 \times S^2$ as a dotted unknot $U \subset S^3$, the union of U with L gives a Kirby diagram for $\hat{X} := B^4 \cup_{\partial-X} X$. As already remarked in [12, Section 4.4], taking the mirror image of such a diagram with respect to a plane intersecting U transversely in two points amounts to replacing the curves $\gamma_i = p_i \mu + q_i \lambda$ with $\bar{\gamma}_i = p_i \mu - q_i \lambda$ and changing the signs of their relative framings. The result is a diagram of a horizontal decomposition of $-X$. Observe that

- if $(\delta, \mathbf{x})$ was associated to the γ_i 's, to the $\bar{\gamma}_i$'s is associated $-(\delta, \mathbf{x})$.

Moreover, using the notation of Section 3.3 it is easy to check that

- since

$$(23) \circ \mu_2 \circ (12) \circ \mu_1 = (123) \quad \text{and} \quad (12) \circ \mu_2 \circ (23) \circ \mu_3 = (132),$$

for each cyclic permutation σ of $(1, 2, 3)$, the transformation $(\delta, \mathbf{x}) \mapsto \sigma(\delta, \mathbf{x})$ is realized by a composition of two Hurwitz moves on the factorization.

Summarizing, the observations above show that there are bijections

$$S_a^{\delta_1, \delta_2, \delta_3} \leftrightarrow S_a^{-\delta_1, -\delta_2, -\delta_3}, \quad S_a^{\delta_1, \delta_2, \delta_3} \leftrightarrow S_a^{\delta_{\sigma(1)}, \delta_{\sigma(2)}, \delta_{\sigma(3)}}$$

with σ any cyclic permutation. Moreover, the above bijections are induced by orientation-reversing diffeomorphisms in the first case and orientation-preserving diffeomorphisms in the second case. Therefore, in order to prove Theorem 3.1 it will suffice to determine the oriented cobordisms underlying the minimal factorizations in the cases $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$ and $(\delta_1, \delta_2, \delta_3) = (-1, 1, 1)$. Since there are two possibilities for $a = 2 - \text{tr}(m_L)$, we have four cases in total, which we treat separately below.

3.4 Minimal solutions when $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$

In this section we determine the weakly minimal solution of the equation

$$(3-4) \quad x^2 + y^2 + z^2 - xyz = a,$$

for a equal to either 0 or 4.

Proposition 3.4 *Let (x, y, z) be a solution of (3-4) with $|x|, |y|$ and $|z|$ pairwise distinct. Then (x, y, z) is not weakly minimal. Indeed, mutation at the greatest element in absolute value makes the quantity $|x| + |y| + |z|$ strictly smaller. Moreover, if $\min(|x|, |y|, |z|) > 1$ then mutation at each of the other two elements makes the quantity $|x| + |y| + |z|$ strictly bigger.*

Proof Up to permuting the variables we may assume

$$|x| > |y| > |z|.$$

Note that, by (3-4), this implies $|z| > 0$ (otherwise $x^2 + y^2 = a \in \{0, 4\}$, which would imply $y = 0 = z$); therefore we have $xyz \neq 0$. We claim that $|\hat{x}| < |x|$. After possibly applying sign changes we may assume that $x, y > 0$, and since $xyz = x^2 + y^2 + z^2 - a > 0$, we have $z > 0$ as well. Also, $x\hat{x} = y^2 + z^2 - a > 0$; therefore $\hat{x} > 0$. Replacing x with y in the equation gives

$$2y^2 + z^2 - y^2z - a < y^2(3 - z).$$

Therefore if $z \geq 3$ then y is between x and $\hat{x}$, and since $y < x$ we must have $\hat{x} < x$, as claimed. On the other hand, $z = 2$ is impossible because it implies

$$(x - y)^2 = a - 4,$$

which has no solutions if $a = 0$ and forces $x = y$ if $a = 4$. Finally, $z = 1$ gives

$$x^2 - xy + y^2 = a - 1 \leq 3,$$

but since $x > y > z > 0$ we have $x > 2$, so

$$x^2 - xy + y^2 = \frac{3}{4}x^2 + \left(\frac{1}{2}x - y\right)^2 \geq \frac{3}{4}x^2 > 3.$$

This proves the first part of the statement. To prove the second part it suffices to show that if

$$|x| > |y| > |z| \geq 2$$

then $|y| < |\hat{y}|$ and $|z| < |\hat{z}|$. We have

$$2|y| < |xy| \leq |xy - z| + |z| < |\hat{z}| + |y|,$$

so $|z| < |y| < |\hat{z}|$. Similarly,

$$2|x| \leq |xz| \leq |xz - y| + |y| < |\hat{y}| + |x|,$$

so $|y| < |x| < |\hat{y}|$. □

Lemma 3.5 *A weakly minimal solution (x, y, z) of (3-4) is equal, up to sign changes and permutations, to*

- $(0, 0, 0)$ or $(3, 3, 3)$ if $a = 0$;
- $(2, 0, 0)$, $(-1, 1, 1)$ or $(2, t, t)$ for some $t \geq 2$ if $a = 4$.

Proof By Proposition 3.4, $|x|$, $|y|$ and $|z|$ are not pairwise distinct. Therefore we can assume, possibly after permuting the variables and changing two signs, that $y = z \geq 0$. Then if $a = 0$ we have

$$x^2 = (x - 2)y^2.$$

If $y = 0$ we get the solution $(0, 0, 0)$, otherwise we must have $x = cy$, with $c^2 = x - 2$ for some $c \in \mathbb{Z}$. Therefore we have $c^2 - cy + 2 = 0$, which implies $c \mid 2$. Moreover, $y > 0$ implies $c > 0$, so that $c \in \{1, 2\}$. These two values of c lead to the solutions $(3, 3, 3)$ and $(6, 3, 3)$, and the conclusion when $a = 0$ follows easily. If $a = 4$ we have

$$x^2 - 4 = (x - 2)y^2,$$

which implies either $x = 2$ or $x = y^2 - 2$. It is easy to check that the solution $(2, y, y)$ can only mutate to itself and to $(y^2 - 2, y, y)$; therefore it is weakly minimal if and only if $|y^2 - 2| \geq 2$, which holds for all $y \neq 1$; on the other hand, a necessary condition for $(y^2 - 2, y, y)$ to be weakly minimal is $|y^2 - 2| \leq 2$, ie $y \in \{0, 1, 2\}$. For $y \in \{0, 2\}$, up to sign changes we get the weakly minimal solutions $(2, 0, 0)$ and $(2, 2, 2)$, while for $y = 1$ we get $(-1, 1, 1)$. $\square$

3.5 Underlying 4-manifolds when $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$ and $a = 0$

The minimal solution $(0, 0, 0)$ is realized only by a triple $(\gamma_1, \gamma_2, \gamma_3)$ with $\gamma_1 = \gamma_2 = \gamma_3$, and since $m_L = \tau_\gamma^3$ must fix λ , we have $\gamma_i = \lambda$ for each i and the factorization $\tau_\lambda \tau_\lambda \tau_\lambda = \tau_\lambda^3$. It is easy to check from the associated Kirby diagram that $\hat{X} \cong 3\overline{\mathbb{CP}^2} \# S^1 \times S^3$.

The following computation shows that the minimal solution $(3, 3, 3)$ of Lemma 3.5 is realized by a triple $(\gamma_1, \gamma_2, \gamma_3)$ with $\gamma_3 = \mu + 2\lambda$, $\gamma_2 = 2\mu + \lambda$ and $\gamma_1 = \mu - \lambda$, corresponding to the factorization $\tau_{\gamma_3} \tau_{\gamma_2} \tau_{\gamma_1} = \tau_\lambda^{-9}$. Since γ_3 is nonseparating there is another curve $\tilde{\gamma} \subset T^2$ with $\gamma_3 \cdot \tilde{\gamma} = 1$. The homology classes of γ_3 and $\tilde{\gamma}$ generate $H_1(T^2; \mathbb{Z})$; therefore we can express γ_1 and γ_2 in these coordinates. From $\gamma_1 \cdot \gamma_3 = \gamma_2 \cdot \gamma_3 = 3$ we get $\gamma_1 = a\gamma_3 - 3\tilde{\gamma}$ and $\gamma_2 = b\gamma_3 - 3\tilde{\gamma}$ for some integers a and b , both coprime with 3. Moreover, $\gamma_1 \cdot \gamma_2 = 3$ implies $b = a + 1$, so that the remainders of a and b modulo 3 are forced to be 1 and 2 respectively. We can then write

$$\gamma_1 = (3c + 1)\gamma_3 - 3\tilde{\gamma} = \gamma_3 - 3\bar{\gamma} \quad \text{and} \quad \gamma_2 = (3c + 2)\gamma_3 - 3\tilde{\gamma} = 2\gamma_3 - 3\bar{\gamma}$$

for some $c \in \mathbb{Z}$, where $\bar{\gamma} := \tilde{\gamma} - c\gamma_3$ is another curve satisfying $\gamma_3 \cdot \bar{\gamma} = 1$. Now it is straightforward to check that $m_L(\bar{\gamma}) = \bar{\gamma}$ and $m_L(\gamma_3) = \gamma_3 + 9\bar{\gamma}$, so $\bar{\gamma} = \pm\lambda$. Without loss of generality, up to reversing $\bar{\gamma}$ and the γ_i 's we may assume $\bar{\gamma} = \lambda$. Then $m_L = \tau_\lambda^{-9}$. In particular, since $\gamma_3 \cdot \lambda = 1$ we can write $\gamma_3 = \mu + k\lambda$ for some $k \in \mathbb{Z}$. Any choice of k determines γ_1 , γ_2 and γ_3 , hence a Kirby diagram for $\hat{X}$.

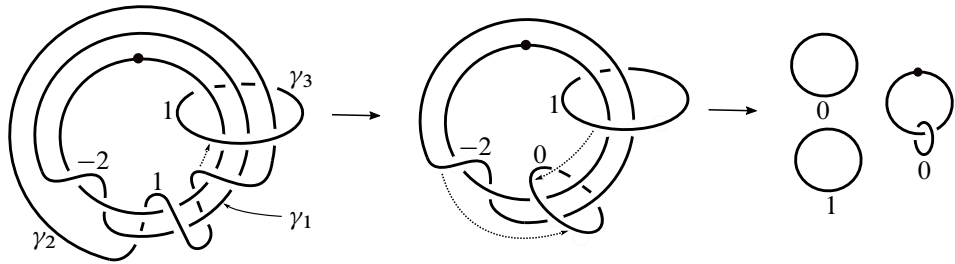


Figure 1: Proof that $\hat{X} \cong \mathbb{C}P^2$ when $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$ and $a = 0$.

Twisting the 1–handle as in Section 2 transforms each γ_i through a power of τ_λ without changing the underlying 4-manifold; therefore it suffices to identify $\hat{X}$ for a single value of k . Choosing $k = 2$ we get

$$\gamma_1 = \mu - \lambda, \quad \gamma_2 = 2\mu + \lambda \quad \text{and} \quad \gamma_3 = \mu + 2\lambda.$$

The resulting Kirby diagram is shown in the first picture from the left in Figure 1, where the presence of a 3- and a 4-handle is understood. The framings $\text{fr}(\gamma_i)$ are obtained as follows. If $\gamma_i = p_i \mu + q_i \lambda$, the Seifert framing induced by the Heegaard torus T is $\text{fr}_T(\gamma_i) = p_i q_i$. Indeed, since this framing is the linking number of γ_i with its push-off in the direction of the negative normal to T , we have

$$\text{fr}_T(\gamma_i) = \text{lk}(\gamma_i, p_i \mu^- + q_i \lambda^-) = \text{lk}(\gamma_i, q_i \lambda^-) = p_i q_i.$$

Then, since $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$, we obtain $\text{fr}(\gamma_i) = \text{fr}_T(\gamma_i) - \delta_i = p_i q_i - 1$ for each $i = 1, 2, 3$. After the handle slide suggested by the dotted arrow one obtains the central picture of Figure 1. Two further handle slides indicated the dotted arrows lead to the picture on the right, where the 0-framed 2-handle cancels with the 3-handle. This shows that $\hat{X} \cong \mathbb{C}P^2$.

3.6 Underlying 4-manifolds when $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$ and $a = 4$

If the solution $(2, 0, 0)$ was realized by curves γ_1, γ_2 and γ_3 then we would have $\gamma_1 \cdot \gamma_2 = \gamma_1 \cdot \gamma_3 = 0$ and therefore $\gamma_1 = \gamma_2 = \gamma_3$ up to signs, which is incompatible with $\gamma_2 \cdot \gamma_3 = 2$. Thus, $(2, 0, 0)$ cannot be realized.

Now consider the solution $(-1, 1, 1)$. Since $\gamma_1 \cdot \gamma_2 = 1$, we may take γ_1 and γ_2 as a symplectic basis and write $\gamma_3 = a\gamma_1 + b\gamma_2$. Then $\gamma_1 \cdot \gamma_3 = 1$ and $\gamma_2 \cdot \gamma_3 = -1$ imply $\gamma_3 = \gamma_1 + \gamma_2$, and the monodromy is

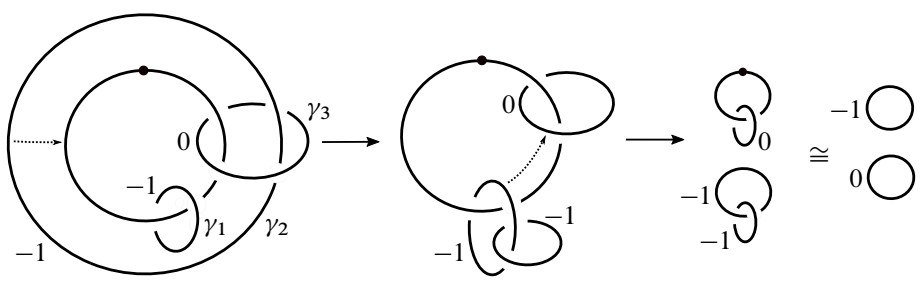


Figure 2: Proof that $\hat{X} \cong \mathbb{C}P^2$ when $(\delta_1, \delta_2, \delta_3) = (1, 1, 1)$ and $a = 4$.

$m_L = \tau_{\gamma_1 + \gamma_2} \tau_{\gamma_2} \tau_{\gamma_1}$. But it is easy to check that $m_L(\gamma_1) = -\gamma_1 - 3\gamma_2$ and $m_L(\gamma_2) = -\gamma_2$. Therefore $\gamma_2 = \pm\lambda$ and $\gamma_1 = \pm\mu + k\lambda$ for some $k \in \mathbb{Z}$. Up to applying a 1-handle twist which does not change the underlying 4-manifold as in [Section 3.5](#), we may assume $\gamma_1 = \pm\mu$, $\gamma_2 = \pm\lambda$ and $\gamma_3 = \pm(\mu + \lambda)$. Then $m_L = -\tau_\lambda^{-3}$ and the resulting Kirby diagram is given by the first picture from the left in [Figure 2](#), where the presence of a 3- and a 4-handle is understood. The indicated handle slides show that $\hat{X} \cong \overline{\mathbb{CP}}^2$.

Finally, consider the solutions $(2, t, t)$ with $t \geq 2$. Suppose that $\tilde{\gamma} \subset T$ is a simple closed curve such that $(\gamma_1, \tilde{\gamma})$ is a symplectic basis. Then $\gamma_2 = a\gamma_1 + t\tilde{\gamma}$ and $\gamma_3 = b\gamma_1 + t\tilde{\gamma}$ for some $a, b \in \mathbb{Z}$ coprime with t , and $2 = \gamma_2 \cdot \gamma_3 = t(a - b)$ implies $t = 2$ and $a = b + 1$. But $b + 1$ and b cannot be both odd, so $(2, t, t)$ is never realized.

3.7 Minimal solutions when $(\delta_1, \delta_2, \delta_3) = (-1, 1, 1)$

In this section we determine the weakly minimal solutions of the equation

$$(3-5) \quad x^2 + xyz - a = y^2 + z^2,$$

where $a \in \{0, 4\}$.

Proposition 3.6 *Let (x, y, z) be a solution of (3-5) with $|x|$, $|y|$ and $|z|$ pairwise distinct. Then (x, y, z) is not weakly minimal. Indeed, mutation at the greatest element in absolute value makes the quantity $|x| + |y| + |z|$ strictly smaller. Moreover, if $\min(|x|, |y|, |z|) > 1$ then mutation at the other two elements makes the quantity $|x| + |y| + |z|$ strictly bigger.*

Proof Note that if (x, y, z) is a solution then so is (x, z, y) . Therefore, up to swapping y and z we may assume $|y| > |z|$. Moreover, it is easy to check that $xyz \neq 0$, because otherwise at least two among x , y and z would be zero. We split the proof into the three cases $|x| > |y| > |z|$, $|y| > |x| > |z|$ and $|y| > |z| > |x|$.

First case Suppose $|x| > |y| > |z|$. We claim that $|\hat{x}| < |x|$. Up to sign changes it suffices to assume $x, y > 0$. This implies $z < 0$, because $z > 0$ gives the contradiction

$$y^2 + z^2 = x^2 + xyz - a > y^2 + 2xz - a \geq y^2 + 2(z + 2)z - a > y^2 + z^2.$$

We have $x(yz + x) = y^2 + z^2 + a > 0$; therefore $\hat{x} = -yz - x < 0$, which implies, since $-yz > 0$ and $-x < 0$, that $|\hat{x}| < |x|$. This shows that mutation at the greatest element in absolute value makes the quantity $|x| + |y| + |z|$ strictly smaller, in particular that (x, y, z) is not weakly minimal. To prove the last part of the statement suppose $|z| > 1$. Then $y > 2$ and

$$(3-6) \quad 2|z| < xy < xy - z + |z| = \hat{z} + |z|,$$

so $|z| < \hat{z}$. Similarly, we have

$$2x < |xz| \leq |xz - y| + |y| < |\hat{y}| + x,$$

so $y < x < |\hat{y}|$. This concludes the proof of the statement in the first case.

Before tackling the other cases we observe that up to sign changes we may assume $y > 0$. Moreover, if $xz < 0$ we have

$$y^2 - xyz + z^2 > x^2 - xyz + z^2 > x^2 \geq x^2 - a,$$

which is a contradiction. Hence, after a further sign change we may assume $x > 0$ and $z > 0$ as well. Therefore, from now on we assume $x, y, z > 0$.

Second case Suppose $|y| > |x| > |z|$. Note that $y\hat{y} = z^2 - x^2 + a \geq 0$ forces $x = 2$, $z = 1$ and $a = 4$, in which case $y\hat{y} = 1$, which is impossible because $y > z \geq 1$. Therefore we have $y\hat{y} < 0$, which implies $\hat{y} = xz - y < 0$. It follows that $0 < xz < y$, hence $|\hat{y}| < |y|$. To prove the last part of the statement suppose $|z| > 1$. Then $|z| < |\hat{z}|$ follows as before from (3-6). Finally, we have

$$2|y| \leq |-yz - x| + |x| < |\hat{x}| + |y|,$$

hence $|x| < |y| < |\hat{x}|$. This proves the statement in the second case.

Third case Suppose $|y| > |z| > |x|$. Since $y\hat{y} = z^2 - x^2 + a > 0$ we have $\hat{y} = xz - y > 0$. If $\hat{y} \geq y$ then $xz \geq 2y$ and, using that $z \geq x + 1 \geq 2$, we get the contradictory inequalities

$$y^2 + z^2 = x^2 + xyz - a > 2y^2 - a \geq y^2 + (|z| + 1)^2 - a > y^2 + z^2.$$

Therefore $|y| = y > \hat{y} = |\hat{y}|$. To prove the last part of the statement observe that

$$2|x| < |yz| \leq |-yz - x| + |x| = |\hat{x}| + |x|,$$

hence $|x| < |\hat{x}|$. Using the extra assumption $|x| \geq 2$ we have

$$2|y| \leq |xy - z| + |z| < |\hat{z}| + |y|,$$

hence $|z| < |y| < |\hat{z}|$. This proves the third case and concludes the proof of the proposition. $\square$

Lemma 3.7 Let (x, y, z) be a weakly minimal solution of (3-5). Then $(|x|, |y|, |z|)$ is of the form

- $(t, t, 0)$ or $(t, 0, t)$ for some $t \geq 0$ if $a = 0$;
- $(2, t, t)$ for some $t \geq 0$ if $a = 4$.

Proof By Proposition 3.6, $|x|$, $|y|$ and $|z|$ are not pairwise distinct; therefore we can assume, possibly after changing two signs, that two variables are equal and nonnegative; also, since the equation is symmetric in y and z , it suffices to discuss the cases $x = y$ and $y = z$. We have the following possibilities:

- When $a = 0$ and $y = z \geq 0$ we have $x^2 = (2 - x)y^2$. If $y = 0$ we get the solution $(0, 0, 0)$, otherwise we must have $x = cy$ with $c^2 = 2 - x = 2 - cy$, so $c \mid 2$; the values of c for which $y > 0$ are 1 and -2 , and the corresponding solutions are $(1, 1, 1)$ and $(-2, 1, 1)$, none of which is weakly minimal.
- When $a = 0$ and $x = y \geq 0$ we have $x^2z = z^2$, so either $z = 0$ or $z = x^2$, and the weakly minimal solutions are those of the form $(t, t, 0)$ for some $t \geq 0$. Because of the symmetry in y and z this also gives the weakly minimal solutions $(t, 0, t)$, $t \geq 0$.

- When $a = 4$ and $y = z \geq 0$ we have $x^2 - 4 = (2 - x)y^2$, so either $x = 2$ or $x = -y^2 - 2$, and the weakly minimal solutions are those of the form $(2, t, t)$ for some $t \geq 0$.
- When $a = 4$ and $x = y \geq 0$ we have $x^2z - 4 = z^2$, which implies $z > 0$ and $z \mid 4$. The only case where x can be an integer is $z = 2$, which gives the weakly minimal solution $(2, 2, 2)$. $\square$

3.8 Underlying 4-manifolds when $(\delta_1, \delta_2, \delta_3) = (-1, 1, 1)$ and $a = 0$

When $t = 0$ the minimal solution $(t, t, 0) = (t, 0, t) = (0, 0, 0)$ of Lemma 3.7 comes from the trivial factorization $\tau_\lambda = \tau_\lambda \tau_\lambda \tau_\lambda^{-1}$, because $0 = \gamma_1 \cdot \gamma_2 = \gamma_2 \cdot \gamma_3 = \gamma_1 \cdot \gamma_3$ and $m_L(\lambda) = \pm\lambda$ implies $\tau_{\gamma_i} = \tau_\lambda$ for each i . It is easy to check that in this case $\hat{X} \cong \mathbb{CP}^2 \# 2\overline{\mathbb{CP}}^2 \# S^1 \times S^3$. When $t \neq 0$ the same kind of analysis shows that the solution $(t, t, 0)$ comes from the factorization $\tau_\lambda = \tau_\lambda \tau_\gamma \tau_\gamma^{-1}$. Similarly, $(t, 0, t)$ comes from a factorization of the form

$$\tau_\lambda = \tau_\gamma \tau_{\gamma'} \tau_\gamma^{-1} = \tau_{\tau_\gamma(\gamma')} \tau_\gamma \tau_\gamma^{-1} = \tau_\lambda \tau_\gamma \tau_\gamma^{-1}.$$

Since Hurwitz moves on the factorization do not change the underlying 4-manifold, it suffices to analyze the former case. We may assume $\gamma = t\mu + q\lambda$ for some integer q coprime with t , with $q = 0$ if $t = 1$. Now we can proceed as in the case $n = 0$ in Section 2. In fact, as in that case the components $\gamma_1 = \gamma_2 = \gamma$ of L cobound an annulus A of the form $\gamma \times [0, 1]$ with framings with respect to A opposite to each other and equal to 1 in absolute value. Sliding γ_1 over γ_2 changes L into the framed link consisting of γ_2 and a 0-framed meridian of γ_2 . Applying [12, Proposition 3.1] (and the preceding discussion in the same paper) it follows that $\hat{X} \cong \overline{\mathbb{CP}}^2 \# B_{t,q} \cup -B_{t,q}$. Since $B_{t,q} \cup -B_{t,q}$ is not simply connected when $t > 1$ while $B_{1,0} \cup -B_{1,0} \cong S^4$, we get $\hat{X} \cong \overline{\mathbb{CP}}^2$ precisely when the factorization is $\tau_\lambda \tau_\mu \tau_\mu^{-1}$.

3.9 Underlying 4-manifolds when $(\delta_1, \delta_2, \delta_3) = (-1, 1, 1)$ and $a = 4$

The argument at the beginning of Section 3.6 shows that $(2, 0, 0)$ cannot be realized, while the argument at the end of the same section shows that $(2, t, t)$ is not realizable if $t \geq 2$. Therefore, we only need to discuss the solution $(2, 1, 1)$. Since $\gamma_1 \cdot \gamma_2 = 1$ we may take γ_1 and γ_2 as a symplectic basis and

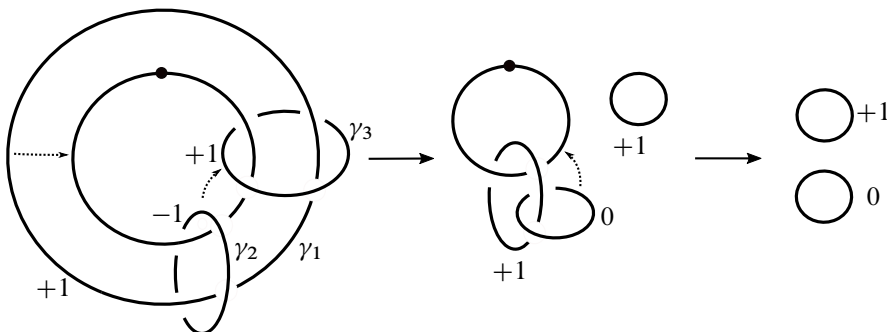


Figure 3: Proof that $\hat{X} \cong \mathbb{CP}^2$ when $(\delta_1, \delta_2, \delta_3) = (-1, 1, 1)$ and $a = 4$.

write $\gamma_3 = a\gamma_1 + b\gamma_2$. Then, $\gamma_1 \cdot \gamma_3 = 1$ and $\gamma_2 \cdot \gamma_3 = 2$ imply $\gamma_3 = \gamma_2 - 2\gamma_1$ and the monodromy is $m_L = \tau_{\gamma_2 - 2\gamma_1} \tau_{\gamma_2} \tau_{\gamma_1}^{-1}$. One can easily check that $m_L(\gamma_1) = -\gamma_1$, which implies $\gamma_1 = \pm\lambda$. Moreover, $\gamma_2 = \mp\mu + k\lambda$ for some $k \in \mathbb{Z}$ and $m_L(\gamma_2) = 5\gamma_1 - \gamma_2$. Arguing as in [Section 3.5](#), we may apply a 1-handle twist which does not change the underlying 4-manifold. Therefore, up to changing the signs of all the curves, we may assume $\gamma_1 = -\lambda$ and $\gamma_2 = \mu$. Thus, $\gamma_3 = \mu + 2\lambda$, $m_L = -\tau_\lambda^{-5}$. As in the previous cases, [Figure 3](#) illustrates some simple Kirby calculus showing $\hat{X} \cong \mathbb{CP}^2$.

Putting together the results of [Sections 3.5, 3.6, 3.8 and 3.9](#) yields [Theorem 3.1](#) and establishes [Theorem 1.1](#) when $\chi(\hat{X}) = 3$.

4 Proof of [Theorem 1.2](#)

Suppose that we have a $(1, 1, 3, 1)$ -decomposition of $\mathbb{CP}^2$ with factorization

$$(4-1) \quad \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} \tau_{\gamma_1}^{\delta_1} = \pm \tau_\lambda^k$$

and associated triple (x_1, x_2, x_3) , defined as in [\(3-1\)](#):

$$(x_1, x_2, x_3) = (\gamma_2 \cdot \gamma_3, \gamma_1 \cdot \gamma_3, \gamma_1 \cdot \gamma_2).$$

In this section we are going to show how to recover the curves $\gamma_i = p_i\mu + q_i\lambda$ from (x_1, x_2, x_3) . The key fact we are going to use is the following.

Lemma 4.1 *A horizontal decomposition of a smooth, closed 4-manifold X of type $(1, 1, 3, 1)$ with factorization [\(4-1\)](#) determines an embedding of the disjoint union $\delta_1 B_{p_1, q_1} \cup \delta_2 B_{p_2, q_2} \cup \delta_3 B_{p_3, q_3}$ into X .*

Proof Note that $S^1 \times D^3$ is obtained from $H_1 \times [0, 1]$ by smoothing corners, where H_1 is a genus-1 handlebody. Hence, the attaching spheres of the three horizontal 2-handles of X are of the form

$$\gamma_i = \gamma \times \{t_i\} \subset \mathbb{T}_i := \partial H_1 \times \{t_i\}$$

for some $t_1, t_2, t_3 \in (0, 1)$ with $t_1 < t_2 < t_3$. For some small $\varepsilon > 0$ the subsets

$$H_1 \times [t_i - \varepsilon, t_i + \varepsilon] \subset H_1 \times [0, 1],$$

$i = 1, 2, 3$, are pairwise disjoint, and smoothing corners yields three pairwise disjoint copies of $S^1 \times D^3$ sitting inside $S^1 \times D^3$. Clearly, the core disks of the three horizontal 2-handles are pairwise disjoint as well; therefore X contains the disjoint union $X_1 \cup X_2 \cup X_3$, where X_i is the smooth 4-manifold obtained by attaching a four-dimensional 2-handle to $S^1 \times D^3$ along $\gamma_i \subset \mathbb{T}_i$ with framing $\text{fr}(\gamma_i) = \text{fr}_{\mathbb{T}_i}(\gamma_i) - \delta_i$. By [\[12, Proposition 3.2\]](#), X_i is orientation-preserving diffeomorphic to $\delta_i B_{p_i, q_i}$. $\square$

Evaluating the left-hand side of [\(4-1\)](#) at γ_1 and taking the products with γ_i gives

$$\tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} \tau_{\gamma_1}^{\delta_1}(\gamma_1) \cdot \gamma_i = \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2}(\gamma_1) \cdot \gamma_i = \tau_{\gamma_2}^{\delta_2}(\gamma_1) \cdot \tau_{\gamma_3}^{-\delta_3}(\gamma_i) = (\gamma_1 - \delta_2 x_3 \gamma_2) \cdot (\gamma_i - \delta_3(\gamma_3 \cdot \gamma_i) \gamma_3)$$

δ	(1, 1, 1)	(1, 1, 1)	(-1, 1, 1)	(-1, 1, 1)
a	0	4	0	4
m_L	τ_λ^{-9}	$-\tau_\lambda^{-3}$	τ_λ	$-\tau_\lambda^{-5}$
minimal factorization	$\tau_{\mu+2\lambda}\tau_{2\mu+\lambda}\tau_{\mu-\lambda}$	$\tau_{\mu+\lambda}\tau_\lambda\tau_\mu$	$\tau_\lambda\tau_\mu\tau_\mu^{-1}$	$\tau_{\mu+2\lambda}\tau_\mu\tau_{-\lambda}^{-1}$
minimal solutions	(3, 3, 3)	(-1, 1, 1)	(1, 1, 0), (1, 0, 1)	(2, 1, 1)
underlying 4-manifold	$\mathbb{CP}^2$	$\overline{\mathbb{CP}}^2$	$\overline{\mathbb{CP}}^2$	$\mathbb{CP}^2$

Table 1

for $i = 1, 2, 3$. When $i = 1$, using that $(x_1, x_2, x_3) \in S_a^{\delta_1, \delta_2, \delta_3}$ where $a = 2 - \text{tr}(m_L) \in \{0, 4\}$ we obtain the equality

$$\tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} \tau_{\gamma_1}^{\delta_1}(\gamma_1) \cdot \gamma_1 = \delta_3 x_2^2 + \delta_2 x_3^2 - \delta_2 \delta_3 x_1 x_2 x_3 = \delta_1 a - \delta_1 \delta_2 \delta_3 x_1^2.$$

On the other hand, evaluating the right-hand side of (4-1) at γ_1 and taking the intersection product with γ_i gives

$$\pm \tau_\lambda^k(\gamma_1) \cdot \gamma_i = \pm(\gamma_1 + k(\lambda \cdot \gamma_1)\lambda) \cdot \gamma_i.$$

Equating the expressions obtained for $i = 1, 2, 3$ we get following equations:

$$\begin{aligned} \delta_1 a - \delta_1 \delta_2 \delta_3 x_1^2 &= \pm k p_1^2, \\ (4-2) \quad x_3 + \delta_3 x_1 x_2 - \delta_2 \delta_3 x_1^2 x_3 &= \pm(x_3 + k p_1 p_2), \\ x_2 - \delta_2 x_1 x_3 &= \pm(x_2 + k p_1 p_3). \end{aligned}$$

Note that the $\pm$ signs on the right-hand sides are plus signs if and only if $a = 0$, since both statements are equivalent to $\text{tr}(m_L) = 2$.

Remark 4.2 If a $(1, 1, 3, 1)$ -decomposition of $\pm \mathbb{CP}^2$ has factorization $(\tau_{\gamma_3}^{\delta_3}, \tau_{\gamma_2}^{\delta_2}, \tau_{\gamma_1}^{\delta_1})$, the sequence of Hurwitz moves

$$\tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} \tau_{\gamma_1}^{\delta_1} \rightarrow \tau_{\gamma_3}^{\delta_3} \tau_{\tau_{\gamma_2}^{\delta_2}(\gamma_1)}^{\delta_1} \tau_{\gamma_2}^{\delta_2} \rightarrow \tau_{\tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2}(\gamma_1)}^{\delta_1} \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} = \tau_{\tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} \tau_{\gamma_1}^{\delta_1}(\gamma_1)}^{\delta_1} \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2} = \tau_{\pm \tau_\lambda^k(\gamma_1)}^{\delta_1} \tau_{\gamma_3}^{\delta_3} \tau_{\gamma_2}^{\delta_2}$$

produces another Kirby diagram of $\pm \mathbb{CP}^2$ and an embedding of the same triple of rational homology balls, because changing a curve γ_i in a factorization by a power of τ_λ corresponds to twisting the 1-handle of $S^1 \times D^3 \cup X_{\gamma_i}$. Note that these moves permute $(\delta_1, \delta_2, \delta_3)$ cyclically. Therefore, in order to prove Theorem 1.2 it suffices to consider, for each possible value of a , a single representative of $\delta = (\delta_1, \delta_2, \delta_3)$ in each orbit of the $\mathbb{Z}/3\mathbb{Z}$ -action by cyclic permutations.

The proof of Theorem 1.2 proceeds as follows. The analysis of Sections 3.5, 3.6, 3.8 and 3.9 together with Remark 4.2 show that a closed 4-manifold with a $(1, 1, 3, 1)$ -decomposition is diffeomorphic to $\mathbb{CP}^2$ or $\overline{\mathbb{CP}}^2$ precisely in the cases of Table 1. In Sections 4.1, 4.2, 4.3 and 4.4 we analyze the cases corresponding to the four columns of Table 1. It turns out that the second column of Table 1 is not relevant for Theorem 1.2, while the first, third and fourth columns correspond, respectively, to the embeddings of Theorem 1.2(1), (2) and (3).

4.1 $\delta = (1, 1, 1)$, $a = 0$ and $m_L = \tau_\lambda^{-9}$

In this case the equations in (4-2) take the form

$$(4-3) \quad \begin{aligned} -x_1^2 &= kp_1^2, \\ x_3 + x_1x_2 - x_1^2x_3 &= x_3 + kp_1p_2, \\ x_2 - x_1x_3 &= x_2 + kp_1p_3. \end{aligned}$$

Moreover, in Section 3.5 we showed that the underlying 4-manifold is $\mathbb{CP}^2$ only when $k = -9$ and (x_1, x_2, x_3) is a nontrivial solution of (3-4) obtained via a sequence of mutations and sign changes from the minimal solution $(3, 3, 3)$. Therefore — as one can prove by induction on the number of mutations — we can write $x_i = 3y_i$, where (y_1, y_2, y_3) are nonzero integers satisfying Markov's equation

$$y_1^2 + y_2^2 + y_3^2 = 3y_1y_2y_3.$$

Relations (4-3) can be written as follows:

$$(4-4) \quad \begin{aligned} y_1^2 &= p_1^2, \\ y_1(y_2 - 3y_1y_3) &= -p_1p_2, \\ y_1y_3 &= p_1p_3. \end{aligned}$$

Solving for p_1 , p_2 and p_3 we obtain $p_1 = \pm y_1$, $p_2 = \pm(3y_1y_3 - y_2) = \pm\hat{y}_2$ and $p_3 = \pm y_3$. Therefore,

$$\gamma_1 = \pm y_1\mu + q_1\lambda, \quad \gamma_2 = \pm\hat{y}_2\mu + q_2\lambda \quad \text{and} \quad \gamma_3 = \pm y_3\mu + q_3\lambda,$$

where, due to relations (3-1), the q_i 's satisfy

$$(4-5) \quad y_1q_2 - \hat{y}_2q_1 = \pm 3y_3, \quad y_1q_3 - y_3q_1 = \pm 3y_2 \quad \text{and} \quad \hat{y}_2q_3 - y_3q_2 = \pm 3y_1.$$

Note that the first and last of (4-5) imply

$$(4-6) \quad q_1 \equiv \mp 3 \frac{y_3}{\hat{y}_2} \pmod{y_1}, \quad q_2 \equiv \mp 3 \frac{y_1}{y_3} \pmod{\hat{y}_2} \quad \text{and} \quad q_3 \equiv \pm 3 \frac{y_1}{\hat{y}_2} \pmod{y_3}.$$

Therefore, arguing as in [12, Proposition 3.2 and Lemma 5.1] we conclude that when a $(1, 1, 3, 1)$ -decomposition induces an embedding in $\mathbb{CP}^2$ of a disjoint union of $B_{y,q}$'s, the union is of the form

$$(4-7) \quad B_{y_1, q_1} \cup B_{\hat{y}_2, q_2} \cup B_{y_3, q_3},$$

where the y_i 's satisfy Markov's equation and the q_i 's satisfy (4-6).

Conversely, given any nonzero Markov triple (y_1, y_2, y_3) , the system of congruences (4-6) has a unique solution in the q_i 's. In fact, when (y_1, y_2, y_3) is a nonzero Markov triple, for either choice of sign there is a triple (q_1, q_2, q_3) satisfying (4-5) and therefore also (4-6). Indeed, since $(y_1, \hat{y}_2, y_3)$ is also a nonzero Markov triple, y_1 , $\hat{y}_2$ and y_3 are pairwise coprime and $y_2\hat{y}_2 = y_1^2 + y_3^2$, which implies $y_3/\hat{y}_2 \equiv y_2/y_3 \pmod{y_1}$. Therefore we can find q_1 , q_2 and q_3 which solve simultaneously the first two of (4-5). But notice that the last equation is obtained dividing by y_1 the difference between the second equation multiplied by $\hat{y}_2$ and the first equation multiplied by y_3 . It follows that (q_1, q_2, q_3) is in fact a

solution the whole system (4-5). Now define $\gamma_1 = y_1\mu + q_1\lambda$, $\gamma_2 = \hat{y}_2\mu + q_2\lambda$ and $\gamma_3 = y_3\mu + q_3\lambda$ and consider the cobordism $X: \partial_- X \rightarrow \partial_+ X$ with a horizontal decomposition of type $(1, 1, 3, 1)$ prescribed by the factorization $(\tau_{\gamma_3}, \tau_{\gamma_2}, \tau_{\gamma_1})$. Then, by (4-5) the triple $(x_1, x_2, x_3) = (\gamma_2 \cdot \gamma_3, \gamma_1 \cdot \gamma_3, \gamma_1 \cdot \gamma_2)$ is a nonzero solution of (3-4) with $a = 0$. Using the fact that (y_1, y_2, y_3) and $(y_1, \hat{y}_2, y_3)$ are Markov triples mutant of each other it is straightforward to check that (3-3) holds replacing (p_1, p_2, p_3) with $(y_1, \hat{y}_2, y_3)$ and (x_1, x_2, x_3) with $(3y_1, 3y_2, 3y_3)$. Since (3-3) is equivalent to $m_L(\lambda) \cdot \lambda = 0$, this implies $m_L(\lambda) = \pm\lambda$ and therefore $X: S^3 \rightarrow S^3$. Moreover, by Lemma 3.5 the triple (x_1, x_2, x_3) is obtained by a sequence of mutations from $(3, 3, 3)$, the arguments of Section 3.5 show that $\hat{X} \cong \mathbb{CP}^2$ and the disjoint union (4-7) embeds in it. Since every nonzero Markov triple is of the form $(y_1, \hat{y}_2, y_3)$, this shows that any disjoint union of the form (4-7) embeds in $\mathbb{CP}^2$. This covers the embeddings of Theorem 1.2(1).

4.2 $\delta = (1, 1, 1)$, $a = 4$ and $m_L = -\tau_\lambda^{-3}$

By the results of Section 3.6, in this case the underlying 4-manifold is $\overline{\mathbb{CP}}^2$. Moreover, the only realizable weakly minimal triple in $S_4^{1,1,1}$ is $(-1, 1, 1)$, and a simple verification shows that the only triples obtainable from $(-1, 1, 1)$ via mutations are $(2, 1, 1)$, $(-1, -2, 1)$ and $(-1, 1, -2)$. On the other hand, as observed in Section 3.3, mirroring the horizontal decompositions produces decompositions of $\mathbb{CP}^2$ with monodromy $m_L = -\tau_\lambda^3$ and it induces an identification $S_4^{-1,-1,-1} \leftrightarrow S_4^{1,1,1}$ given by $(x_1, x_2, x_3) \mapsto (-x_1, -x_2, -x_3)$. Thus, since we are interested in embeddings in $\mathbb{CP}^2$, we only need to consider the triples $(1, -1, -1)$, $(-2, -1, -1)$, $(1, 2, -1)$ and $(1, -1, 2)$. When $\delta = (-1, -1, -1)$, $a = 4$ and $m_L = -\tau_\lambda^3$, the equations in (4-2) become

$$\begin{aligned} x_1^2 - 4 &= -3p_1^2, \\ 2x_3 - x_1x_2 - x_1^2x_3 &= -3p_1p_2, \\ 2x_2 + x_1x_3 &= -3p_1p_3. \end{aligned} \tag{4-8}$$

If $(x_1, x_2, x_3) = (1, -1, -1)$ it is easy to check that the only solutions are $(p_1, p_2, p_3) = \pm(1, 0, 1)$. When $(x_1, x_2, x_3) = (1, 2, -1)$ and $(x_1, x_2, x_3) = (1, -1, 2)$ the only solutions are $(p_1, p_2, p_3) = \pm(1, 1, -1)$ and $(p_1, p_2, p_3) = \pm(1, -1, 0)$, respectively. Finally, when $(x_1, x_2, x_3) = (-2, -1, -1)$, the equations in (4-8) imply $p_1 = 0$ and say nothing about p_2 nor p_3 . In order to extract information about p_2 and p_3 we evaluate both sides of (4-1) on γ_2 and we take the intersection product with γ_3 , obtaining the equation

$$2x_1 - x_2x_3 - x_1x_3^2 = -3p_2p_3.$$

The only solutions when $(x_1, x_2, x_3) = (-2, -1, -1)$ are $(p_2, p_3) = \pm(1, 1)$, and we conclude that $(p_1, p_2, p_3) = \pm(0, 1, 1)$. We observe that in every case we have $\max\{p_i\} \leq 1$; therefore none of the embeddings obtained when $\delta = (1, 1, 1)$ and $a = 4$ are relevant for the statement of Theorem 1.2.

4.3 $\delta = (-1, 1, 1)$, $a = 0$ and $m_L = \tau_\lambda$

The analysis of Section 3.8 shows that (x_1, x_2, x_3) is obtained by a sequence of mutations and sign changes from either $(1, 1, 0)$ or $(1, 0, 1)$. This implies that the x_i 's are pairwise coprime, otherwise by

(3-5) they would have a nontrivial common factor, but such property is preserved by mutation and sign changes while $(1, 1, 0)$ and $(1, 0, 1)$ do not have it. Moreover, the only solution of (3-5) with $x_1 = 0$ is the trivial one $(0, 0, 0)$, which is equivalent only to itself up to mutations and sign changes. Therefore $x_1 \neq 0$.

Since the underlying 4-manifold is $\overline{\mathbb{CP}}^2$ we argue as in Section 4.2. There is a one-to-one correspondence $S_0^{1,-1,-1} \leftrightarrow S_0^{-1,1,1}$ given by the map sending (x_1, x_2, x_3) to $(-x_1, -x_2, -x_3)$, and induced by the operation of taking the mirror image of the Kirby diagram of the underlying 4-manifold. Thus, the triples $(x_1, x_2, x_3) \in S_0^{1,-1,-1}$ whose underlying 4-manifold is $\mathbb{CP}^2$ are those obtained by a sequence of mutations and sign changes from either $(-1, -1, 0)$ or $(-1, 0, -1)$, the x_i 's are pairwise coprime and $x_1 \neq 0$. When $\delta = (1, -1, -1)$, $a = 0$ and $k = -1$, the equations in (4-2) read

$$(4-9) \quad \begin{aligned} x_1^2 &= p_1^2, \\ x_1(x_2 + x_1x_3) &= p_1p_2, \\ -x_1x_3 &= p_1p_3, \end{aligned}$$

and solving for p_1 , p_2 and p_3 we obtain $p_1 = \pm x_1$, $p_2 = \mp \hat{x}_2$ and $p_3 = \mp x_3$. Therefore,

$$\gamma_1 = \pm x_1\mu + q_1\lambda, \quad \gamma_2 = \mp \hat{x}_2\mu + q_2\lambda \quad \text{and} \quad \gamma_3 = \mp x_3\mu + q_3\lambda$$

and, due to relations (3-1), the q_i 's satisfy

$$(4-10) \quad x_1q_2 + \hat{x}_2q_1 = \pm x_3, \quad x_1q_3 + x_3q_1 = \pm x_2 \quad \text{and} \quad x_3q_2 - \hat{x}_2q_3 = \pm x_1.$$

Note that the first and last of (4-10) imply

$$(4-11) \quad \hat{x}_2q_1 \equiv \pm x_3 \pmod{x_1}, \quad x_3q_2 \equiv \pm x_1 \pmod{\hat{x}_2} \quad \text{and} \quad x_2q_3 \equiv \mp x_1 \pmod{x_3}.$$

We can conclude that if a $(1, 1, 3, 1)$ -decomposition with factorization of the form $(\tau_{\gamma_3}^{-1}, \tau_{\gamma_2}^{-1}, \tau_{\gamma_1})$ induces an orientation-preserving embedding in $\mathbb{CP}^2$ of a disjoint union of $\pm B_{x,q}$'s, then such a union is of the form

$$(4-12) \quad B_{x_1,q_1} \cup -B_{x_2,q_2} \cup -B_{x_3,q_3},$$

where the x_i 's satisfy (3-2) with $\delta = (1, -1, -1)$ and $a = 0$ and the q_i 's satisfy (4-11).

We are now going to show that, conversely, if (x_1, x_2, x_3) is a nonzero triple of integers satisfying (3-2) with $\delta = (1, -1, -1)$ and $a = 0$ and the triple (x_1, x_2, x_3) is obtained by a sequence of mutations and sign changes from either $(-1, -1, 0)$ or $(-1, 0, -1)$, then the disjoint union (4-12), where the triple (q_1, q_2, q_3) satisfies (4-11), smoothly embeds in $\mathbb{CP}^2$. We start observing that there is a triple (q_1, q_2, q_3) satisfying (4-10) with a plus sign on the right-hand side. Indeed, since $(x_1, \hat{x}_2, x_3)$ is also a solution of (3-2) with $\delta = (1, -1, -1)$ and $a = 0$, by the argument given at the beginning of this section x_1 , $\hat{x}_2$ and x_3 are pairwise coprime. Moreover, by (3-2) we have $x_2\hat{x}_2 = x_3^2 - x_1^2$, which implies $x_2\hat{x}_2 \equiv x_3^2 \pmod{x_1}$ and, since $x_1 \neq 0$, $(\hat{x}_2, x_3) \neq (0, 0)$. Therefore there is a simultaneous solution $\xi \pmod{x_1}$ of the equations

$$\xi\hat{x}_2 = \pm x_3 \pmod{x_1} \quad \text{and} \quad \xi x_3 = \pm x_2 \pmod{x_1}.$$

This guarantees that we can find q_1, q_2 and q_3 solving simultaneously the first two equations of (4-10) with a plus sign on the right-hand side. Now note that the last equation with a plus sign on the right-hand side is obtained dividing by x_1 the difference between the first equation multiplied by x_3 and the second equation multiplied by $\hat{x}_2$. Since $x_1 \neq 0$, it follows that (q_1, q_2, q_3) is in fact a solution of the whole system. Now define

$$\gamma_1 = x_1\mu + q_1\lambda, \quad \gamma_2 = \hat{x}_2\mu + q_2\lambda \quad \text{and} \quad \gamma_3 = x_3\mu + q_3\lambda$$

and consider the cobordism $X: \partial_- X \rightarrow \partial_+ X$ with a horizontal decomposition of type $(1, 1, 3, 1)$ prescribed by the factorization $(\tau_{\gamma_3}^{-1}, \tau_{\gamma_2}^{-1}, \tau_{\gamma_1})$. Then, by (4-10) the triple

$$(\gamma_2 \cdot \gamma_3, \gamma_1 \cdot \gamma_3, \gamma_1 \cdot \gamma_2) = (x_1, x_2, x_3)$$

is a solution of (3-2) with $\delta = (1, -1, -1)$ and $a = 0$. Using the fact that (x_1, x_2, x_3) and $(x_1, \hat{x}_2, x_3)$ are mutant of each other it is straightforward to check using the equality $x_2\hat{x}_2 = x_3^2 - x_1^2$ that (3-3) holds for $\delta = (1, -1, -1)$ replacing (p_1, p_2, p_3) with $(x_1, \hat{x}_2, x_3)$. Since (3-3) is equivalent to $m_L(\lambda) \cdot \lambda = 0$, this implies $m_L(\lambda) = \pm\lambda$ and therefore $X: S^3 \rightarrow S^3$. But the analysis of Section 3.8 shows that the 4-manifold underlying the triple $(-x_1, -x_2, -x_3) \in S_0^{-1,1,1}$ is $\mathbb{CP}^2$ precisely when $(-x_1, -x_2, -x_3)$ is obtained by a sequence of mutations from either $(1, 1, 0)$ or $(1, 0, 1)$. Since we are assuming that (x_1, x_2, x_3) is obtained by a sequence of mutations and sign changes from either $(-1, -1, 0)$ or $(-1, 0, -1)$, this implies that the disjoint union (4-12) smoothly embeds in $\mathbb{CP}^2$, which is what we needed to show. This covers the embeddings of Theorem 1.2(2).

4.4 $\delta = (-1, 1, 1), a = 4$ and $m_L = -\tau_\lambda^{-5}$

We observed in Section 3.3 and Remark 4.2 that for each cyclic permutation σ of $(1, 2, 3)$ the transformation $(\delta, x) \mapsto \sigma(\delta, x)$ is realized by a composition of two Hurwitz moves on the factorization. Therefore, the permutation (123) induces a bijection $S_4^{-1,1,1} \rightarrow S_4^{1,-1,1}$ which preserves the set of integer triples realized as intersection numbers of triples of curves. By the results of Section 3.9 it follows that a triple $(x_1, x_2, x_3) \in S_4^{1,-1,1}$ is realized by three curves, γ_1, γ_2 and γ_3 , if and only if it is obtained from $(1, 2, 1)$ via mutations and sign changes.

It turns out that such triples, as well as the corresponding curves γ_i , admit an explicit description involving the Fibonacci and the Lucas numbers. Recall that the Lucas sequence $(L_n)_{n \in \mathbb{Z}}$ is defined by the same recursive relation as the Fibonacci sequence, but starting from $L_0 = 2$ and $L_1 = 1$. We will use the following identities:

$$(I1) \quad L_n = F_{n-1} + F_{n+1},$$

$$(I2) \quad L_{-n} = (-1)^n L_n \quad \text{and} \quad F_{-n} = (-1)^{n+1} F_n,$$

$$(I3) \quad L_m L_n = L_{m+n} + (-1)^n L_{m-n},$$

$$(I4) \quad 5F_m F_n = L_{m+n} - (-1)^n L_{m-n},$$

$$(I5) \quad L_m F_n = F_{m+n} - (-1)^n F_{m-n},$$

$$(I6) \quad L_n^2 = 5F_n^2 + 4(-1)^n = L_{2n} + 2(-1)^n,$$

$$(I7) \quad \gcd(F_m, F_n) = F_{\gcd(m,n)},$$

$$(I8) \quad F_{m+n} = F_{m+1}F_n + F_mF_{n-1},$$

$$(I9) \quad F_n^2 = F_{n+r}F_{n-r} + (-1)^{n-r}F_r^2, \quad n \geq r > 0.$$

All of the above identities can be found online, eg on Wikipedia, except perhaps (I3), which follows easily from (I1) and (I5). [Proposition 4.4](#) below shows that the triples (x_1, x_2, x_3) we are interested in have a specific form, described by the following definition.

Definition 4.3 Let $a, b \in \mathbb{Z}$ be odd and coprime and $\varepsilon \in \{\pm 1\}$ and denote by $L(\varepsilon, a, b)$ the triple $(L_a, -\varepsilon L_{a+\varepsilon b}, L_b)$. Define $L \subset \mathbb{Z}^3$ to be the subset of triples of the form $L(\varepsilon, a, b)$ with ε, a and b as above.

Proposition 4.4 The set L is contained in $S_4^{1,-1,1}$ and is invariant under mutations and sign changes. More specifically, let μ_i be the i^{th} mutation map defined in [Section 3.3](#) and $\delta = (1, -1, 1)$. Then the following formulas hold:

$$\mu_i(\delta, L(\varepsilon, a, b)) = \begin{cases} (\delta, L(\varepsilon, -a - 2\varepsilon b, b)) & \text{if } i = 1, \\ (\delta, L(-\varepsilon, a, b)) & \text{if } i = 2, \\ (\delta, L(\varepsilon, a, -b - 2\varepsilon a)) & \text{if } i = 3. \end{cases}$$

Moreover, L coincides with the set of triples obtained from $(1, 2, 1)$ by mutations and sign changes.

Proof By looking at (3-5) for $\delta = (1, -1, 1)$ and $a = 4$, it is easy to check that, given odd coprime integers a and b , the triples $L(1, a, b)$ and $L(-1, a, b)$ are in $S_4^{1,-1,1}$ if and only if $-L_{a+b}$ and L_{a-b} are the roots of the polynomial

$$x^2 + L_a L_b x - (L_a^2 + L_b^2 + 4),$$

or equivalently that

$$-L_{a+b} + L_{a-b} = -L_a L_b \quad \text{and} \quad L_{a+b} L_{a-b} = L_a^2 + L_b^2 + 4.$$

The first equality follows from (I2) and (I3), the second one from (I3) and the identity

$$L_a^2 + L_b^2 + 4 = L_{2a} + L_{2b},$$

which follows from (I6). Now we compute the mutations of a generic triple

$$(x_1, x_2, x_3) = (L_a, -\varepsilon L_{a+\varepsilon b}, L_b).$$

The argument just provided shows that $\hat{x}_2 = \varepsilon L_{a-\varepsilon b}$; therefore $\mu_2(\delta, L(\varepsilon, a, b)) = (\delta, L(-\varepsilon, a, b))$. Moreover, using (I2) and (I3),

$$\hat{x}_1 = -\varepsilon L_{a+\varepsilon b} L_b - L_a = -\varepsilon L_{\varepsilon a+b} L_b - L_a = -\varepsilon (L_{\varepsilon a+2b} - L_{\varepsilon a}) - L_a = L_{-a-2\varepsilon b},$$

so

$$\mu_1(\delta, L(\varepsilon, a, b)) = (\delta, L(\varepsilon, -a - 2\varepsilon b, b)).$$

A similar computation yields $\hat{x}_3 = L_{-b-2\epsilon a}$ and

$$\mu_3(\delta, L(\epsilon, a, b)) = (\delta, L(\epsilon, a, -b - 2\epsilon a)).$$

Finally, the triples that obtained from $L(\epsilon, a, b)$ by sign changes are

- $(-L_a, \epsilon L_{a+\epsilon b}, L_b) = (L_{-a}, \epsilon L_{-a-\epsilon b}, L_b) = L(-\epsilon, -a, b);$
- $(-L_a, -\epsilon L_{a+\epsilon b}, -L_b) = (L_{-a}, -\epsilon L_{-a-\epsilon b}, L_{-b}) = L(\epsilon, -a, -b);$
- $(L_a, \epsilon L_{a+\epsilon b}, -L_b) = (L_a, \epsilon L_{-a-\epsilon b}, L_{-b}) = L(-\epsilon, a, -b).$

This proves the first part of the statement. In order to prove the last sentence we observe that

$$(1, 2, 1) = L(-1, 1, 1).$$

Combining this with the first part of the statement we get that every triple obtained from $(1, 2, 1)$ by mutations and sign changes has the form $L(\epsilon, a, b)$. Conversely, we claim that any triple $L(\epsilon, a, b)$ can be transformed by a (possibly empty) sequence of mutations into $L(\epsilon', a', b')$ with $|a'| = |b'|$. To prove the claim observe that if $|a| \neq |b|$ then the elements of the triple have pairwise distinct absolute values. Therefore, by [Proposition 3.6](#) there exists a mutation which makes the sum of the absolute values strictly decrease, and the claim follows from a standard infinite descent argument. Since a' and b' are coprime, we must have $|a'| = |b'| = 1$, and up to a change of signs we can assume $a' = b' = 1$, so that $L(\epsilon', a', b')$ is either $(1, 2, 1)$ or $(1, -3, 1)$, which are mutants of each other. $\square$

Consider a triple $(x_1, x_2, x_3) = (L_a, -\epsilon L_{a+\epsilon b}, L_b) \in S_4^{1,-1,1}$. Then the equations in (4-2) take the form

$$\begin{aligned} (4-13) \quad & 4 + x_1^2 = 5p_1^2, \\ & 2x_3 + x_1x_2 + x_1^2x_3 = 5p_1p_2, \\ & 2x_2 + x_1x_3 = 5p_1p_3. \end{aligned}$$

Using identity (I6) we get $p_1 = \pm F_a$ from the first equation; and using identity (I4) the third equation can be written as

$$5p_1p_3 = x_2 - \hat{x}_2 = -\epsilon L_{a+\epsilon b} - \epsilon L_{a-\epsilon b} = -\epsilon(L_{a+b} + L_{a-b}) = -5\epsilon F_a F_b,$$

so $p_3 = \mp \epsilon F_b$ (since $F_a \neq 0$). Finally, using (I2), (I3) and (I4) we can write the second equation as

$$\begin{aligned} 5p_1p_2 &= 2x_3 + x_1(x_2 + x_1x_3) = 2x_3 - x_1\hat{x}_2 \\ &= 2L_b - \epsilon L_a L_{a-\epsilon b} \\ &= 2L_b - \epsilon(L_{2a-\epsilon b} + L_{\epsilon b}) = L_b - L_{2\epsilon a-b} = 5F_a F_{b-\epsilon a}, \end{aligned}$$

so $p_2 = \pm F_{b-\epsilon a}$. Therefore,

$$(4-14) \quad \gamma_1 = \pm F_a \mu + q_1 \lambda, \quad \gamma_2 = \pm F_{b-\epsilon a} \mu + q_2 \lambda \quad \text{and} \quad \gamma_3 = \mp \epsilon F_b \mu + q_3 \lambda$$

and, due to relations (3-1), the q_i 's satisfy

$$(4-15) \quad F_a q_2 - F_{b-\epsilon a} q_1 = \pm L_b, \quad F_a q_3 + \epsilon F_b q_1 = \mp \epsilon L_{a+\epsilon b} \quad \text{and} \quad F_{b-\epsilon a} q_3 + \epsilon F_b q_2 = \pm L_a.$$

Observe that $F_{b-\varepsilon a} = 0$ implies $b = \varepsilon a = \pm 1$ and therefore $F_a = F_b = 1$. This leads to a case which falls outside the scope of [Theorem 1.2](#); therefore from now on we assume $F_{b-\varepsilon a} \neq 0$. Note that the first and last of [\(4-15\)](#) imply

$$(4-16) \quad q_1 \equiv \mp \frac{L_b}{F_{b-\varepsilon a}} \bmod F_a, \quad q_2 \equiv \pm \frac{L_b}{F_a} \bmod F_{b-\varepsilon a} \quad \text{and} \quad q_3 \equiv \pm \frac{L_a}{F_{b-\varepsilon a}} \bmod F_b.$$

We conclude that if a $(1, 1, 3, 1)$ -decomposition with monodromy $-\tau_\lambda^{-5}$ induces an orientation-preserving embedding in $\mathbb{CP}^2$ of a disjoint union of $\pm B_{x,q}$'s, then such a union is of the form

$$(4-17) \quad B_{F_a, q_1} \cup -B_{F_{b-\varepsilon a}, q_2} \cup B_{F_b, q_3}$$

for some odd, coprime integers a, b and some $\varepsilon \in \{\pm 1\}$, and where the q_i 's satisfy [\(4-16\)](#).

Now we want to argue that, conversely, if a, b are odd and coprime and $\varepsilon \in \{\pm 1\}$, then the disjoint union [\(4-17\)](#), where the q_i 's satisfy [\(4-16\)](#), smoothly embeds in $\mathbb{CP}^2$. We start observing that, given a, b and ε as above, there is a triple (q_1, q_2, q_3) satisfying [\(4-15\)](#) with the first choice of sign on the right-hand side. Indeed, by (I2) and (I5) we have

$$\varepsilon L_b F_b - F_a L_a = F_{2\varepsilon b} - F_{2a} = -L_{a+\varepsilon b} F_{a-\varepsilon b} = \varepsilon L_{a+\varepsilon b} F_{b-\varepsilon a} \equiv 0 \bmod F_{b-\varepsilon a},$$

so $L_b/F_a \equiv L_a/\varepsilon F_b \bmod F_{b-\varepsilon a}$. This guarantees that we can find integers q_1, q_2 and q_3 satisfying the first and third of [\(4-15\)](#). On the other hand, F_a times the third equation minus εF_b times the first equation gives

$$F_{b-\varepsilon a}(F_a q_3 + \varepsilon F_b q_1) = F_a L_a - \varepsilon F_b L_b \stackrel{(I5)}{=} F_{2a} - F_{2\varepsilon b} = -\varepsilon F_{b-\varepsilon a} L_{a+\varepsilon b},$$

and therefore the second equation holds. Now define γ_i , for $i = 1, 2, 3$, as in [\(4-14\)](#) with the first choice of sign and consider the cobordism X with a horizontal decomposition of type $(1, 1, 3, 1)$ prescribed by the factorization $(\tau_{\gamma_3}, \tau_{\gamma_2}^{-1}, \tau_{\gamma_1})$. We claim that [\(3-3\)](#) holds for $\delta = (1, -1, 1)$ replacing (p_1, p_2, p_3) with $(F_a, F_{b-\varepsilon a}, -\varepsilon F_b)$ and (x_1, x_2, x_3) with $(L_a, -\varepsilon L_{a+\varepsilon b}, L_b)$. Indeed, after these substitutions [\(3-3\)](#) is equivalent to

$$(4-18) \quad F_a^2 - F_{b-\varepsilon a}^2 + F_b^2 + F_a F_{b-\varepsilon a} L_b - F_a F_b L_{a+\varepsilon b} - \varepsilon F_{b-\varepsilon a} F_b L_a + \varepsilon F_a F_b L_a L_b = 0,$$

which turns out to be an identity involving Fibonacci and Lucas numbers. To see why [\(4-18\)](#) holds, use (I2) and (I3) to replace the last term on the left-hand side with $F_a F_b (L_{a+\varepsilon b} - L_{a-\varepsilon b})$, $F_a F_{b-\varepsilon a} L_b$ with $F_a (F_{2b-\varepsilon a} - F_a)$, $F_a F_b L_{a+\varepsilon b}$ with $F_a (F_{2b-\varepsilon a} + F_a)$, and $-F_{b-\varepsilon a} F_b L_{\varepsilon a}$ with $-F_b^2 + F_{2a-\varepsilon b} F_b$. After cancellations, [\(4-18\)](#) becomes

$$F_{2a-\varepsilon b} F_b - F_a^2 - F_{a-\varepsilon b}^2 = 0,$$

which holds by (I9). Since [\(3-3\)](#) is equivalent to $m_L(\lambda) \cdot \lambda = 0$, this implies $m_L(\lambda) = \pm \lambda$ and therefore $X: S^3 \rightarrow S^3$. Moreover, by [\(4-15\)](#) the triple $(\gamma_2 \cdot \gamma_3, \gamma_1 \cdot \gamma_3, \gamma_1 \cdot \gamma_2) = (L_a, -\varepsilon L_{a+\varepsilon b}, L_b)$ belongs to $L \subset S_4^{1,-1,1}$; therefore by the last statement of [Proposition 4.4](#) it is obtained from $(1, 2, 1)$ by mutations and sign changes. Since the identification of $S_4^{1,-1,1}$ with $S_4^{-1,1,1}$ is induced by Hurwitz moves, by the

analysis of [Section 3.9](#) we conclude that $\hat{X} \cong \mathbb{CP}^2$ and the disjoint union (4-17) smoothly embeds in it. This covers the embeddings of [Theorem 1.2\(3\)](#) and concludes the proof of [Theorem 1.2](#).

5 Proof of Theorem 1.4

The organization of the proof requires that we start with [Theorem 1.4\(3\)](#). We need the following result.

Theorem 5.1 (Carmichael's theorem [2]) *Let n be a positive integer, $n \notin \{1, 2, 6, 12\}$. Then F_n admits a primitive factor, ie a prime p such that $n = \min\{k \in \mathbb{N}^+ \mid p \text{ divides } F_k\}$.*

Remark 5.2 An immediate consequence of Carmichael's theorem is that if n is odd and $n > 3$, then F_n always admits an *odd* primitive factor p . Moreover, by (I7), for any integer k we have

$$p \mid F_k \iff p \mid \gcd(F_n, F_k) = F_{\gcd(n,k)} \iff n \mid k.$$

The following statement implies [Theorem 1.4\(3\)](#). We shall prove it after a few preliminaries.

Lemma 5.3 *Given an odd integer $a > 1$, let $S(a) \subset \mathcal{F}_3$ be the subset of elements of the form $B_{F_a,q}$, and let $\psi: \mathbb{N}_{>1} \rightarrow \mathbb{N}$ be the map given by $\psi(a) = |S(a)|$. Then, for each odd prime $a > 3$ we have $\psi(a) = (a-1)/2$. In particular, ψ is unbounded. Moreover, for each odd integer $a > 3$ the map $f: S(a) \rightarrow \mathbb{Z}/F_a\mathbb{Z}$ given by $f(B_{F_a,q}) = q^2 \bmod F_a$ is injective.*

Let $a > 1$ be an odd integer. Given $b \in \mathbb{Z}$ odd and coprime with a and $\varepsilon \in \{\pm 1\}$, by [Theorem 1.2\(3\)](#) there is a smooth, orientation-preserving embedding $B_{F_a,q} \subset \mathbb{CP}^2$, where $q \equiv \pm L_b / F_{b-\varepsilon a} \bmod F_a$. We would like to understand how many different such embeddings we can construct if we keep a fixed. In other words, setting $I(b, \varepsilon) := L_b / F_{b-\varepsilon a} \bmod F_a$, we want to understand for which pairs (b, ε) and (b', ε') we have $I(b, \varepsilon) \equiv \pm I(b', \varepsilon') \bmod F_a$. We are going to do this assuming $a > 3$, because the only $B_{p,q}$'s that we have when $a = 1$ or $a = 3$ are, respectively, $B_{1,0} = B^4$ and $B_{2,1}$, which are both already known to embed in $\mathbb{CP}^2$. In principle we should allow b and ε to vary independently of each other, but we observe that, since a and b are both odd, by (I2) we have

$$I(b, \varepsilon) = L_b / F_{b-\varepsilon a} = (-L_b) / (-F_{b-\varepsilon a}) = L_{-b} / F_{-b-(-\varepsilon)a} = I(-b, -\varepsilon).$$

Therefore, it will suffice to fix $\varepsilon = -1$ and understand for which b and b' we have $I(b, -1) \equiv \pm I(b', -1) \bmod F_a$.

Proposition 5.4 *Let a and b be odd, coprime integers with $a > 3$. Then $I(b, -1) \equiv \pm I(b', -1) \bmod F_a$ if and only if $b \equiv \pm b' \bmod a$.*

Proof Using (I5) we see that the congruences $I(b, -1) \equiv \pm I(b', -1) \bmod F_a$ are equivalent to

$$F_{b+b'+a} - F_{b-b'-a} = L_b F_{b'+a} \equiv \pm L_{b'} F_{b+a} = L_{\pm b'} F_{b+a} = F_{\pm b'+b+a} - F_{\pm b'-b-a} \bmod F_a$$

and using (I8) to

$$(5-1) \quad (F_{b+b'} - F_{\pm b'+b})F_{a-1} + (F_{\pm b'-b} - F_{b-b'})F_{-a-1} \equiv 0 \bmod F_a.$$

Since a is odd, by (I2) we have $F_{-a-1} = -F_{a+1} \equiv -F_{a-1} \pmod{F_a}$, and since F_{a-1} and F_a are coprime, (5-1) is equivalent to

$$(F_{b+b'} - F_{\pm b'+b}) - (F_{\pm b'-b} - F_{b-b'}) \equiv 0 \pmod{F_a}.$$

Thus,

$$(5-2) \quad \begin{aligned} 2F_{b-b'} &\equiv 0 \pmod{F_a} \quad \text{if and only if } I(b, -1) \equiv I(b', -1) \pmod{F_a}, \\ 2F_{b+b'} &\equiv 0 \pmod{F_a} \quad \text{if and only if } I(b, -1) \equiv -I(b', -1) \pmod{F_a}. \end{aligned}$$

Now observe that if $b \equiv \pm b' \pmod{a}$ then by (I7) $\gcd(F_{b \mp b'}, F_a) = F_a$; therefore by (5-2) we have $I(b, -1) \equiv \pm I(b', -1) \pmod{F_a}$. Conversely, by Remark 5.2 the number F_a admits an odd primitive factor p . Equations (5-2) imply that p must divide $F_{b \mp b'}$ if $I(b, -1) \equiv \pm I(b', -1) \pmod{F_a}$, and again by Remark 5.2 this is equivalent to $a \mid b \mp b'$. In other words, $b \equiv \pm b' \pmod{a}$. $\square$

Proof of Lemma 5.3 Given any odd integer b coprime with a we have $I(b, -1) = I(b_0, -1)$ for some odd b_0 such that $b_0 \equiv b \pmod{2a}$ and $0 < b_0 < 2a$. On the other hand, given two odd integers b and b' , both coprime with a and both lying in the interval $(0, 2a)$, we have

$$b \equiv b' \pmod{a} \iff b = b' \quad \text{and} \quad b \equiv -b' \pmod{a} \iff b + b' = 2a.$$

Since an odd integer is coprime with a if and only if it is coprime with $2a$, the interval $(0, 2a)$ contains exactly $\varphi(2a)$ odd integers coprime with a , where φ denotes Euler's totient function, and since the involution $x \mapsto 2a - x$ has no fixed points among such integers, we get $\psi(a) = \frac{1}{2}\varphi(2a) = \frac{1}{2}\varphi(a)$, which takes the value $(a-1)/2$ when a is prime. In particular, ψ is unbounded. This establishes the first part of the statement.

Now observe that $I(b, -1)^2 = L_b^2/F_{b+a}^2 = L_{b'}^2/F_{b'+a}^2 = I(b', -1)^2 \pmod{F_a}$ if and only if

$$(5-3) \quad L_{b'}^2 F_{b+a}^2 - L_b^2 F_{b'+a}^2 = (L_{b'} F_{b+a} + L_b F_{b'+a})(L_{b'} F_{b+a} - L_b F_{b'+a}) \equiv 0 \pmod{F_a}.$$

Moreover, by (I5)

$$L_{b'} F_{b+a} - L_b F_{b'+a} = F_{b-b'-a} - F_{b'-b-a} = L_{-a} F_{b-b'} \equiv 2F_{-a-1} F_{b-b'}$$

and by (I2), (I5) and (I8),

$$\begin{aligned} L_{b'} F_{b+a} + L_b F_{b'+a} &= 2F_{b'+b+a} - F_{b'-b-a} - F_{b-b'-a} \equiv 2F_{-a-b-b'} - F_{-a+1}(F_{b'-b} + F_{b-b'}) \\ &\equiv 2F_{-a+1} F_{-b-b'} \pmod{F_a}. \end{aligned}$$

Therefore (5-3) is equivalent to

$$4F_{b-b'} F_{b+b'} \equiv 0 \pmod{F_a}.$$

Let p be a primitive factor of F_a . As observed in Remark 5.2 p is necessarily odd. Then p divides at least one among $F_{b-b'}$ and $F_{b+b'}$, and by Remark 5.2 we conclude that $b \equiv \pm b' \pmod{a}$ and therefore by Proposition 5.4 $I(b, -1) = \pm I(b', -1) \pmod{F_a}$. $\square$

By Lemma 5.3 the sequence $\{\psi(a)\}_a$ is unbounded. Since, using the notation of Theorem 1.4, we have $|S_{F_a}| \geq \psi(a)$, this implies Theorem 1.4(3). We are now going to prove Theorem 1.4(1) and (2).

Suppose $B_{p,q} \in \mathcal{F}_1$. Without loss of generality we may assume $B_{p,q} = B_{p_1,q_1}$, with (p_1, p_2, p_3) satisfying Markov's equation $p_1^2 + p_2^2 + p_3^2 = 3p_1p_2p_3$. This implies $p_2^2 + p_3^2 \equiv 0 \pmod{p_1}$; therefore $q^2 = q_1^2 \equiv 9p_2^2/p_3^2 \equiv -9 \pmod{p_1}$. Similarly, if $B_{p,q} \in \mathcal{F}_2$ then $B_{p,q} = B_{x_1,q_1}$ and using the equation $x_1^2 + x_1x_2x_3 = x_2^2 + x_3^2$ and the congruence $x_2q_1 \equiv \pm x_3 \pmod{x_1}$ one concludes $q^2 \equiv -1 \pmod{p}$. This proves one direction of both [Theorem 1.4](#)(1) and (2).

To prove the other direction of [Theorem 1.4](#)(1) it suffices to show that any rational ball $B_{p,q}$ appearing in the statement of [Theorem 1.2](#) and satisfying $q^2 \equiv -9 \pmod{p}$ actually belongs to $\mathcal{F}_1$. If $B_{p,q} \in \mathcal{F}_2$, $q^2 \equiv -1 \pmod{p}$ implies $p \mid 8$ and we are forced to have $p \in \{1, 2\}$ because $q^2 \equiv -1 \pmod{p}$ is impossible if $4 \mid p$. Therefore $B_{p,q} \in \{B_{1,0}, B_{2,1}\}$. But $B_{1,0}$ and $B_{2,1}$ belong to $\mathcal{F}_1$ because they both arise from the Markov triple $(2, 1, 1)$. Now suppose that $B_{F_a,q} \in \mathcal{F}_3$. When $a = 3$ the rational ball $B_{F_a,q}$ is equal to $B_{2,1}$ and therefore belongs to $\mathcal{F}_1$. Thus, we may assume $a > 3$. We claim that $q^2 \equiv -9 \pmod{F_a}$ implies $B_{F_a,q} = B_{F_a,F_{a-4}}$. Indeed, observe that, choosing $b = 2 - a$ and $\varepsilon = -1$ we have

$$q \equiv \pm L_{2-a}/F_2 \equiv \pm L_{2-a} \pmod{F_a}.$$

Moreover,

$$L_{2-a} = -L_{a-2} = -(F_{a-3} + F_{a-1}) = -(F_a - F_{a-4}) \equiv F_{a-4} \pmod{F_a}.$$

Therefore, by (I9) we have $q^2 \equiv F_{a-4}^2 \equiv -F_a^2 \equiv -9 \pmod{F_a}$. In view of the last statement of [Lemma 5.3](#), this proves the claim. Since $(1, F_{a-2}, F_a)$ is a Markov triple and $F_{a-4} \equiv 3F_{a-2} \pmod{F_a}$, we have $B_{F_a,F_{a-4}} \in \mathcal{F}_1$, and the other direction of [Theorem 1.4](#)(1) holds.

To prove the other direction of [Theorem 1.4](#)(2), as in the previous case it suffices to show that any rational ball $B_{p,q}$ appearing in the statement of [Theorem 1.2](#) and satisfying $q^2 \equiv -1 \pmod{p}$, actually belongs to $\mathcal{F}_2$. As before, if $B_{p,q} \in \mathcal{F}_1$ then $B_{p,q} \in \{B_{1,0}, B_{2,1}\}$, and both $B_{1,0}$ and $B_{2,1}$ belong to $\mathcal{F}_2$ because they arise, respectively, from the triples $(x_1, x_2, x_3) = (1, 1, 0)$ and $(x_1, x_2, x_3) = (2, 1, 3)$. If $B_{F_a,q} \in \mathcal{F}_3$, as before we may assume $a > 3$. We claim that $q^2 \equiv -1 \pmod{F_a}$ implies $B_{F_a,q} = B_{F_a,F_{a-2}}$. This time we choose $b = 1$ and $\varepsilon = -1$. Then, by (I9), $q \equiv \pm L_1/F_{1+a} \equiv \mp F_{a-1} \equiv \pm F_{a-2} \pmod{F_a}$ and

$$F_{a-2}^2 = F_a F_{a-4} - F_2^2 \equiv -1 \pmod{F_a}.$$

Therefore $q^2 \equiv -1 \pmod{F_a}$ implies $q^2 \equiv F_{a-2}^2 \pmod{F_a}$. As before, in view of [Lemma 5.3](#) this proves the claim. Since by (I9) the triple $(x_1, x_2, x_3) = (F_a, F_{a+1}, 1)$ solves the equation $x_1^2 + x_1x_2x_3 = x_2^2 + x_3^2$,

$$F_{a+1}F_{a-1} \equiv -1 \pmod{F_a} \quad \text{and} \quad F_{a-2} \equiv -F_{a-1} \pmod{F_a},$$

we have $B_{F_a,F_{a-2}} \in \mathcal{F}_2$ and the other direction of [Theorem 1.4](#)(2) holds. This concludes the proof of [Theorem 1.4](#).

References

- [1] **D Auroux**, *Factorizations in $\mathrm{SL}(2, \mathbb{Z})$ and simple examples of inequivalent Stein fillings*, J. Symplectic Geom. 13 (2015) 261–277 [MR](#) [Zbl](#)

- [2] **R D Carmichael**, *On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$* , Ann. of Math. 15 (1913/14) 49–70 [MR](#) [JFM](#)
- [3] **J D Evans, I Smith**, *Markov numbers and Lagrangian cell complexes in the complex projective plane*, Geom. Topol. 22 (2018) 1143–1180 [MR](#) [Zbl](#)
- [4] **R Fintushel, R J Stern**, *Rational blowdowns of smooth 4-manifolds*, J. Differential Geom. 46 (1997) 181–235 [MR](#) [Zbl](#)
- [5] **R Fintushel, R J Stern**, *Double node neighborhoods and families of simply connected 4-manifolds with $b^+ = 1$* , J. Amer. Math. Soc. 19 (2006) 171–180 [MR](#) [Zbl](#)
- [6] **R E Gompf, A I Stipsicz**, *4-manifolds and Kirby calculus*, Graduate Studies in Mathematics 20, Amer. Math. Soc., Providence, RI (1999) [MR](#) [Zbl](#)
- [7] **P Hacking, Y Prokhorov**, *Smoothable del Pezzo surfaces with quotient singularities*, Compos. Math. 146 (2010) 169–192 [MR](#) [Zbl](#)
- [8] **T Khodorovskiy**, *Symplectic rational blow-up and embeddings of rational homology balls*, PhD thesis, Harvard University (2012) Available at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:9572269>
- [9] **T Khodorovskiy**, *Symplectic rational blow-up*, preprint (2013) [arXiv 1303.2581](#)
- [10] **T Khodorovskiy**, *Smooth embeddings of rational homology balls*, Topology Appl. 161 (2014) 386–396 [MR](#) [Zbl](#)
- [11] **P Lisca, A Parma**, *On Stein rational balls smoothly but not symplectically embedded in $\mathbb{C}P^2$* , Bull. Lond. Math. Soc. 54 (2022) 949–960 [MR](#) [Zbl](#)
- [12] **P Lisca, A Parma**, *Horizontal decompositions, I*, Algebr. Geom. Topol. 24 (2024) 3503–3525 [MR](#) [Zbl](#)
- [13] **P Lisca, A Parma**, *On almost complex embeddings of rational homology balls*, from “Frontiers in geometry and topology” (P M N Feehan, L L Ng, P S Ozsváth, editors), Proc. Sympos. Pure Math. 109, Amer. Math. Soc., Providence, RI ([2024] ©2024) 183–193 [MR](#) [Zbl](#)
- [14] **B Owens**, *Smooth, nonsymplectic embeddings of rational balls in the complex projective plane*, Q. J. Math. 71 (2020) 997–1007 [MR](#) [Zbl](#)
- [15] **J Park**, *Simply connected symplectic 4-manifolds with $b_2^+ = 1$ and $c_1^2 = 2$* , Invent. Math. 159 (2005) 657–667 [MR](#) [Zbl](#)
- [16] **A Scorpan**, *The wild world of 4-manifolds*, Amer. Math. Soc., Providence, RI (2005) [MR](#) [Zbl](#)
- [17] **M Symington**, *Symplectic rational blowdowns*, J. Differential Geom. 50 (1998) 505–518 [MR](#) [Zbl](#)

Department of Mathematics, University of Pisa
Pisa, Italy

paolo.lisca@unipi.it, andrea.parma94@gmail.com

Received: 12 March 2023 Revised: 9 March 2024

On the nonorientable four-ball genus of torus knots

FRASER BINNS
SUNGKYUNG KANG
JONATHAN SIMONE
PAULA TRUÖL

The nonorientable four-ball genus of a knot K in S^3 is the minimal first Betti number of nonorientable surfaces in B^4 bounded by K . By amalgamating ideas from involutive knot Floer homology and unoriented knot Floer homology, we give a new lower bound on the smooth nonorientable four-ball genus γ_4 of any knot. This bound is sharp for several families of torus knots, including $T_{4n, (2n \pm 1)^2}$ for even $n \geq 2$, a family Longo showed were counterexamples to Batson’s conjecture. We also prove that, whenever p is an even positive integer and $\frac{1}{2}p$ is not a perfect square, the torus knot $T_{p,q}$ does not bound a locally flat Möbius band for almost all integers q relatively prime to p .

[57K10](#); [57K18](#)

1 Introduction

The *nonorientable (smooth) 4-ball genus* $\gamma_4(K)$ of a knot K in the 3-sphere S^3 is the minimal first Betti number $b_1(F) = \dim H_1(F; \mathbb{Q})$ of any smoothly embedded nonorientable surface F in the 4-ball B^4 with $\partial F = K$. By definition, $\gamma_4(K) \geq 1$ for any knot K in S^3 , and $\gamma_4(K) = 1$ if and only if K bounds a smooth Möbius band in B^4 . Moreover, if K is smoothly slice, ie K bounds a smoothly embedded disk in B^4 , then $\gamma_4(K) = 1$.

First introduced by Murakami and Yasuhara [39], γ_4 has drawn increased attention in recent years. One class of knots that has received considerable attention is the class of torus knots $T_{p,q}$ for relatively prime positive integers p and q . For $p \leq 3$ (or similarly, $q \leq 3$), it is known that $\gamma_4(T_{p,q}) = 1$: $T_{1,q}$ is the unknot for all q ; $T_{2,q}$ bounds an obvious band with q half twists (ie a Möbius band); and a single nonorientable band move turns $T_{3,q}$ into the unknot for each q (see Allen [1, Section 2.1]). In general, however, the values of γ_4 for torus knots are not known.

In [3], Batson constructed a nonorientable surface $F_{p,q}$ in B^4 with $\partial F_{p,q} = T_{p,q}$ via *pinch moves* (defined in Section 2.3). The minimal first Betti number of $F_{p,q}$ is called the *pinch number* of $T_{p,q}$, which we denote by $\vartheta(T_{p,q})$. Batson conjectured that $\gamma_4(T_{p,q}) = \vartheta(T_{p,q})$. This could be seen as the nonorientable analogue of the Milnor conjecture

$$(1-1) \quad g_4(T_{p,q}) = g(T_{p,q}) = \frac{1}{2}(p-1)(q-1),$$

which was first proven by Kronheimer and Mrowka [33]. Here $g_4(K)$ denotes the (orientable) 4-ball genus of a knot K in S^3 — the minimal genus of a compact, oriented surface smoothly embedded in B^4 bounded by K — and $g(K)$ denotes the 3-genus of a knot K in S^3 — the minimal genus of a compact, oriented surface smoothly embedded in S^3 bounded by K .¹

By showing that $\gamma_4(T_{2k,2k-1}) = \vartheta(T_{2k,2k-1}) = k - 1$ for all $k \geq 1$, Batson provided an infinite family of torus knots for which his conjecture is true. Further such infinite families of torus knots were found by Jabuka and Van Cott [29]. However, Batson's conjecture was disproved by Lobb [35], who found that $\gamma_4(T_{4,9}) = 1$ while $\vartheta(T_{4,9}) = 2$. Soon after, Longo [36] generalized this example, showing that $\gamma_4(T_{4n,(2n\pm 1)^2}) \leq 2n - 1$ while $\vartheta(T_{4n,(2n\pm 1)^2}) = 2n$ for all $n \geq 2$. Tairi [46] provided another infinite family of counterexamples, showing in particular that $\gamma_4(T_{4,11}) = 1$ while $\vartheta(T_{4,11}) = 2$.

In this article, we will shed more light on the nonorientable 4-ball genus of torus knots, in both the smooth and topological categories. In particular, we will

- develop a new lower bound on γ_4 using involutive unoriented Floer homology (Theorems 1.15 and 1.16 below, see Section 1.2 for details), which — for so-called L -space knots (and consequently, torus knots) — can be read off from the Alexander polynomial (Theorem 1.1 below);
- develop an obstruction to torus knots bounding (locally flat) Möbius bands using linking forms and number-theoretic arguments (see Section 1.1); and
- use these new tools along with existing tools to either compute or give narrow bounds on γ_4 for some infinite families of torus knots.

Throughout, we will always assume that for any torus knot $T_{p,q}$, p and q are positive, relatively prime integers. The main lower bound on γ_4 that we will use is the following.

Theorem 1.1 *Let K be an L -space knot. Write the Alexander polynomial $\Delta_K(t)$ of K as a monic, symmetric Laurent polynomial, and suppose that its constant term is -1 and that the first nonzero term of positive degree is t^k , ie $\Delta_K(t) = \cdots + t^{-k} - 1 + t^k - \cdots$. Then we have the following lower bound on the nonorientable 4-ball genus: $k - 1 \leq \gamma_4(K)$. In addition, if K bounds a smoothly embedded Möbius band in B^4 , then $k = 1$.*

Remark 1.2 To the best of the authors' knowledge, Theorem 1.1 provides the first Heegaard–Floer theoretical proof that the figure-eight knot does not bound a smooth Möbius band in B^4 (see Example 4.14). This seems worth mentioning since it was not until the advent of involutive theory that Heegaard Floer could be used to show that the figure-eight knot is not slice; see Hendricks and Manolescu [24].

¹In contrast to the orientable setting, where $g_4(T_{p,q}) = g(T_{p,q})$ (see (1-1)), it is in general not true that $\gamma_4(T_{p,q}) = \gamma_3(T_{p,q})$ for any relatively prime positive integers p and q , where $\gamma_3(T_{p,q})$ is the minimal first Betti number of any nonorientable smooth surface F in S^3 with $\partial F = K$ (as defined by Clark [7]). In fact, the difference between γ_3 and γ_4 can be arbitrarily large; see Jabuka and Van Cott [28, Theorem 1.1]. By work of Teragaito [47], it is known that the value of $\gamma_3(T_{p,q})$ is determined by a recursive, arithmetic function of p and q .

Remark 1.3 Notice the difference between the Möbius band case and the case of higher first Betti number. Interestingly, this difference is substantial and cannot be removed since the bound in the case of higher first Betti number is sometimes sharp; see the proof of [Proposition 1.6](#) in [Section 3](#).

Remark 1.4 The integer k in [Theorem 1.1](#) is called the *stretch* of the given L -space knot K , which was defined by Borodzik and Hom in [\[5\]](#). It is shown in [\[5, Theorem 3.10\]](#) that the stretch of a torus knot $T_{p,q}$ is given by $\lfloor \frac{1}{2}(a_k - 1) \rfloor + 1$, where $q > p$ and

$$\frac{q}{p} = [a_0, \dots, a_k] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_k}}}}$$

is a regular continued fraction expansion with $a_k > 1$ and $k \geq 2$.

A quick overview of the proof of [Theorem 1.1](#) is provided in [Section 1.2](#), while the full proof can be found in [Section 4](#). It turns out that [Theorem 1.1](#) provides a sharper bound than many other existing lower bounds involving knot invariants (see [Section 2](#)). In fact, we will use [Theorem 1.1](#) to give a partial affirmative answer to a question of Longo [\[36, Question 4.3\]](#), which asks whether his examples $T_{4n, (2n \pm 1)^2}$ satisfy $\gamma_4(T_{4n, (2n \pm 1)^2}) = 2n - 1$.

Proposition 1.5 *If $n \geq 2$ is even, then $\gamma_4(T_{4n, (2n \pm 1)^2}) = 2n - 1 = \vartheta(T_{4n, (2n \pm 1)^2}) - 1$.*

In fact, we prove the following more general result.

Proposition 1.6 *Let $n \geq 2$ be even and $k \geq 0$. Then*

$$\gamma_4(T_{4n+2k, (4n+2k)(n \pm 1) + 1}) = 2n + k - 1 = \vartheta(T_{4n+2k, (4n+2k)(n \pm 1) + 1}) - 1.$$

Using [Theorem 1.1](#) and other existing results (see [Section 2](#)), we can also find narrow bounds on γ_4 for certain infinite families of torus knots $T_{p,q}$ with fixed p .

Theorem 1.7 *Fix $p > 3$. If $q \equiv p - 1, p + 1$, or $2p - 1 \pmod{2p}$, then $\vartheta(T_{p,q}) - 1 \leq \gamma_4(T_{p,q}) \leq \vartheta(T_{p,q})$.*

In particular, we have the following:

(a) *Let p be odd.*

If $q \equiv p - 1 \pmod{2p}$, then $\gamma_4(T(p, q)) \in \{\frac{1}{2}(p - 3), \frac{1}{2}(p - 1)\}$.

If $q \equiv p + 1$ or $2p - 1 \pmod{2p}$, then $\gamma_4(T_{p,q}) = \frac{1}{2}(p - 1)$.

(b) *Let p be even.*

If $q > p$ and $q \equiv p - 1, p + 1$ or $2p - 1 \pmod{2p}$, then $\gamma_4(T_{p,q}) \in \{\frac{1}{2}(p - 2), \frac{1}{2}p\}$.

Remark 1.8 The torus knots in [Proposition 1.6](#) belong to the family considered in [Theorem 1.7\(b\)](#). In particular, [Theorem 1.7\(b\)](#) tells us that $\gamma_4(T_{4n+2k, (4n+2k)(n \pm 1) + 1}) \in \{2n + k - 1, 2n + k\}$. To prove the

first equality in [Proposition 1.6](#), we will show that performing k pinch moves on $T_{4n+2k, (4n+2k)(n\pm 1)+1}$ yields $T_{4n, (2n\pm 1)^2}$ and then apply [Proposition 1.5](#).

Remark 1.9 The lower bound on γ_4 given in the first part of [Theorem 1.7\(a\)](#) is obtained using [Theorem 1.1](#). We will see that other existing lower bounds involving knot invariants (see [Section 2](#)) do not suffice to show this. The second statement of [Theorem 1.7\(a\)](#) follows from existing work (see [Corollary 2.11](#)). Similarly, the middle case of [Theorem 1.7\(b\)](#) is proved using [Theorem 1.1](#), whereas the others follow from existing work (see [Corollary 2.11](#)).

Remark 1.10 Consider the family of torus knots in [Theorem 1.7\(b\)](#) with $q = p + 1$. These are of the form $T_{2n, 2n+1}$. Feller and Golla [[13](#), Proposition 5.7] showed that these torus knots do not bound a smooth Möbius band in B^4 if $n \equiv 3 \pmod 4$, $2n + 1$ is square-free, and $n + 1$ is not a square. [Theorem 1.7\(b\)](#) generalizes this result.

Remark 1.11 Notice that [Theorem 1.7](#) includes the case $q = p - 1$ when p is odd, but not when p is even. The latter case was solved by Batson [[3](#)]; he showed that $\gamma_4(T_{p, p-1}) = \frac{1}{2}(p - 2) = \vartheta(T_{p, p-1})$ for even p .

We now turn our attention to the family of torus knots $T_{4, q}$ for odd $q > 4$. We note that the bound, 2, that [Theorem 1.7\(b\)](#) yields is no better than the one obtained by purely diagrammatic methods in [Lemma 2.9](#). However, we are able to get more mileage out of [Theorem 1.1](#) than was used in [Theorem 1.7\(b\)](#) since we need only obstruct the existence of Möbius bands (cf [Remark 1.3](#)).

Theorem 1.12 *If $q \equiv 5$ or $7 \pmod 8$, then $\gamma_4(T_{4, q}) = 2$. Moreover, $T_{4, q}$ does not bound a Möbius band for almost all $q \equiv 1$ or $3 \pmod 8$.*

In [Theorem 1.12](#), since q is relatively prime to 4, there are four cases to consider: $q \equiv 1, 3, 5, 7 \pmod 8$. The cases $q \equiv 5, 7 \pmod 8$ are easy to deal with: we first show that $\gamma_4(T_{4, q}) \leq \vartheta(T_{4, q}) = 2$; and then we use [Theorem 1.1](#) to show that $\gamma_4(T_{4, q}) = 2$. When $q \equiv 1, 3 \pmod 8$, we again have that $\gamma_4(T_{4, q}) \leq \vartheta(T_{4, q}) = 2$, but the lower bound on γ_4 given by [Theorem 1.1](#) does not apply. Moreover, other existing lower bounds involving knot invariants (see [Sections 2](#) and [3.1](#)) do not provide any additional information. At this point, one might wonder whether [Theorem 1.12](#) can be upgraded to $\gamma_4(T_{4, q}) = 2$ for all $q \equiv 1, 3 \pmod 8$, confirming Batson's conjecture for the torus knots $T_{4, q}$. It is known, however, that this is not the case — as mentioned above, we have $\gamma_4(T_{4, 9}) = 1$ [[35](#)] and $\gamma_4(T_{4, 11}) = 1$ [[46](#)].

The issue with applying [Theorem 1.1](#) and the other existing lower bounds is that these lower bounds all use knot invariants (eg the Arf invariant, signature, upsilon invariant, etc) that are periodic modulo $2p$ for $T_{p, q}$. Since $\gamma_4(T_{p, 1}) = \gamma_4(T_{p, 3}) = 1$ for any p , this periodic behavior implies that all of these lower bounds on $\gamma_4(T_{p, q})$ are at most 1 for all $q \equiv 1$ or $3 \pmod{2p}$. In particular we cannot use these lower bounds to obstruct $T_{4, q}$ from bounding a Möbius band for $q \equiv 1$ or $3 \pmod 8$. Instead, to prove the second statement of [Theorem 1.12](#), we consider the linking form of the double cover of S^3 branched over $T_{4, q}$

(and more generally branched over $T_{p,q}$ for p even) and apply [Theorem 1.13](#) below. We in fact prove something stronger, namely that these torus knots do not even bound *locally flat* Möbius bands. This gives us information about the *topological nonorientable 4-ball genus* γ_4^{top} of $T_{4,q}$. It is worth noting that this obstruction does not work if p and q are both odd; in this case, the double cover of S^3 branched over $T_{p,q}$, which is the Brieskorn sphere $\Sigma(2, p, q)$, is an integral homology sphere (as opposed to a rational homology sphere if p or q is even) and thus the linking form obstruction vanishes.

1.1 Obstructing Möbius bands in the topological category

For a knot K in S^3 , let $\gamma_4^{\text{top}}(K)$ denote the minimal first Betti number of any locally flat nonorientable surface in B^4 bounding K ; note that $\gamma_4^{\text{top}}(K) \leq \gamma_4(K)$ for any knot K . Gilmer and Livingston showed in [\[18, Theorem D\]](#) that there exist knots K with $\gamma_4^{\text{top}}(K) = 1$, but $\gamma_4(K) > 1$. They found these knots by taking the connected sum of a topologically slice knot with a knot K satisfying $\gamma_4^{\text{top}}(K) = \gamma_4(K) = 1$. In fact, the nonorientable 4-ball genus of topologically slice knots can be arbitrarily large; see Feller, Park and Ray [\[15\]](#).

In [\[13, Theorem 1.3\]](#), Feller and Golla showed that $\gamma_4^{\text{top}}(T_{2n-1,2n}) \leq n-2 < n-1 = \gamma_4(T_{2n-1,2n})$ for all $n \geq 5$; the latter equality is due to Batson [\[3\]](#). Moreover, they provided new infinite families of knots not bounding locally flat Möbius bands and were particularly interested in the question of whether there exists a torus knot $T_{p,q}$ with $\gamma_4^{\text{top}}(T_{p,q}) = 1$ and $\gamma_4(T_{p,q}) > 1$. To the best of the authors' knowledge, this question still remains open.

In [Section 5](#), using a bound of Murakami and Yasuhara [\[39, Theorem 2.5\]](#) we prove that torus knots which bound locally flat Möbius bands are very rare: when p is even and $\frac{1}{2}p$ is not a square, the set of odd integers q satisfying $\gamma_4^{\text{top}}(T_{p,q}) = 1$ has “zero density”. In particular, we prove the following theorem.

Theorem 1.13 *Let p be a positive even integer such that $\frac{1}{2}p$ is not a perfect square. Given a positive integer N , consider the set $S_{N,p} = \{n \in \mathbb{Z} \mid 1 \leq n \leq N, \gcd(p, n) = 1\}$. Then*

$$\lim_{N \rightarrow \infty} \frac{\#\{q \in S_{N,p} \mid T_{p,q} \text{ bounds a locally flat Möbius band in } B^4\}}{\#S_{N,p}} = 0.$$

Remark 1.14 The proof of [Theorem 1.13](#) actually shows the following: if p is even and $\frac{1}{2}p$ is not a perfect square, then there exists an integer r , relatively prime to p , such that if q admits a prime factor s with odd exponent in the prime factorization of q and $s \equiv r \pmod{2p}$, then $T_{p,q}$ does not bound a locally flat Möbius band in B^4 (see [Proposition 5.12](#)). For example, if $q \equiv 1$ or $3 \pmod{8}$ and q has a prime factor $s \equiv 5 \pmod{8}$ such that the exponent of s in the prime factorization of q is odd, then $T_{4,q}$ does not bound a locally flat Möbius band (see [Example 5.11](#) for details). Notice that the smallest values of q satisfying this condition are $q = 35, 65, 91, 105$. Thus $T_{4,35}$, $T_{4,65}$, $T_{4,91}$, and $T_{4,105}$ do not bound locally flat Möbius bands; moreover, for any $q \equiv 1$ or $3 \pmod{8}$ with $q < 105$ and $q \neq 35, 65, 91$, it remains unknown whether the torus knot $T_{4,q}$ bounds a locally flat Möbius band.

1.2 New lower bound on the smooth nonorientable 4-ball genus

In the smooth category, our new lower bound (see [Theorem 1.1](#)) comes from a construction which amalgamates involutive Heegaard Floer homology [24] and unoriented knot Floer homology [12]. Given a knot K in S^3 , we construct a module over the ring $\mathbb{F}_2[U, Q]/(Q^2)$, which we call the *involutive unoriented knot Floer homology* $\text{HFKI}'(K)$. Using the structure of $\text{HFKI}'(K)$, we can define two concordance invariants $\bar{v}$ and $\underline{v}$, which are the values at 1 of the *involutive Upsilon invariants* $\bar{\Upsilon}_t$ and $\underline{\Upsilon}_t$ defined by Hogancamp and Livingston [25].

Although HFKI' itself does not seem to satisfy the strong functoriality properties that Heegaard Floer theory and its variants satisfy, we can still show that $\bar{v}$ and $\underline{v}$ behave nicely under nonorientable cobordisms. In particular, we have the following results, which can be seen as analogues of the oriented involutive genus bound given by Juhász and Zemke in [32].

Theorem 1.15 *Let K be a knot in S^3 which bounds a smooth Möbius band in B^4 . Then $\bar{v}(K) - \underline{v}(K) \leq 1$.*

Theorem 1.16 *For any knot K in S^3 , we have*

$$\bar{v}(K) - v(K) \leq \gamma_4(K) + 1,$$

$$v(K) - \underline{v}(K) \leq \gamma_4(K) + 1,$$

$$\bar{v}(K) - \underline{v}(K) \leq \gamma_4(K) + 2.$$

Here, $v(K)$ is the *upsilon invariant* defined by Ozsváth, Stipsicz and Szabó [41].

In [Section 4](#), we will prove [Theorems 1.15](#) and [1.16](#) and use them to derive [Theorem 1.1](#).

1.3 Questions

Recall that a torus knot of the form $T_{4,q}$ bounds a Möbius band when $q = 1, 3, 9, 11$. [Theorem 1.12](#) eliminates the cases $q \equiv 5, 7 \pmod{8}$ in the smooth category and tells us that the set of all q such that $T_{4,q}$ bounds a locally flat Möbius band has zero density, but there can still be infinitely many q such that $T_{4,q}$ bounds a Möbius band in either smooth or topological category (see [Remark 1.14](#)). Thus the following question naturally arises.

Question 1 Does there exist an integer $q \notin \{1, 3, 9, 11\}$ such that $\gamma_4(T_{4,q}) = 1$? If such an integer q exists, then are there infinitely many such integers q ?

Moreover, all known counterexamples to Batson's conjecture are of the form $T_{p,q}$ where p is even. No example with p and q both odd is currently known. We thus ask the following question.

Question 2 Do there exist relatively prime odd integers p and q such that $\gamma_4(T_{p,q}) < \vartheta(T_{p,q})$?

In the topological category, [Theorem 1.13](#) tells us that for any even p with $\frac{1}{2}p$ not a perfect square, $T_{p,q}$ does not bound a locally flat Möbius band for almost all q . However, this obstruction does not imply that

the set of q such that $T_{p,q}$ bounds a locally flat Möbius band is finite. In addition, the case when $\frac{1}{2}p$ is a perfect square remains mysterious. Hence one can ask the following question; note that, when p is odd, $\gamma_4(T_{p,q}) = \vartheta(T_{p,q}) = 1$ whenever $q \equiv \pm 2 \pmod{p}$ [29, Proposition 1.6].

Question 3 Is there an even integer $p \geq 4$ such that $T_{p,q}$ bounds a locally flat Möbius band in B^4 for infinitely many integers q ?

Organization

This article is organized as follows. In [Section 2](#), we recall some previously known results regarding lower and upper bounds on the nonorientable 4-ball genus before proving [Proposition 1.6](#) and [Theorems 1.7](#) and [1.12](#) in [Section 3](#). In [Section 4](#), we develop a theory of involutive unoriented knot Floer homology and prove [Theorems 1.15](#), [1.16](#), and [1.1](#). Lastly, in [Section 5](#), we compute the linking form of $\Sigma(2, p, q)$, which is the branched double cover of S^3 along $T_{p,q}$, when p is even, and use it to prove [Theorem 1.13](#).

Acknowledgements

This work is the product of a research group formed under the auspices of the American Institute for Mathematics (AIM) in their virtual research community on 4-dimensional topology. We are grateful to AIM, and especially to the program organizers Miriam Kuzbary, Maggie Miller, Juanita Pinzón-Caicedo and Hannah Schwartz. The authors thank Peter Feller, Marco Golla, Stanislav Jabuka and the referees for their useful comments on a draft of this paper. Kang thanks Seungwon Kim for helpful conversations and acknowledges support by the Institute for Basic Science (IBS-R003-D1). Truöl was supported by the Swiss National Science Foundation grant 181199.

2 Background on the nonorientable 4-ball genus

In this section, we give an overview of the lower and upper bounds on the nonorientable 4-ball genus provided in the literature, including a review of pinch moves, the pinch number, and related computations.

2.1 Obstructions for the existence of smooth nonorientable surfaces

Even before the notion of nonorientable 4-ball genus was born, Viro [48] proved that the figure-eight knot cannot bound a smoothly embedded Möbius band in B^4 using Witt classes of intersection forms of branched covers of B^4 branched over nonorientable surfaces.

In [51], Yasuhara formulated an obstruction using the classical signature $\sigma(K)$ and the Arf invariant $\text{Arf}(K)$ of a knot K which can be stated as follows.

Lemma 2.1 [51, Proposition 5.1] *Let K be a knot in S^3 that bounds a smoothly embedded Möbius band in B^4 . Then $\sigma(K) + 4 \text{Arf}(K) \equiv 0$ or $\pm 2 \pmod{8}$.*

Remark 2.2 By [34], the Arf invariant satisfies

$$\text{Arf}(K) = \begin{cases} 0 & \text{if } \Delta_K(-1) \equiv \pm 1 \pmod{8}, \\ 1 & \text{if } \Delta_K(-1) \equiv \pm 3 \pmod{8}. \end{cases}$$

For torus knots $T_{p,q}$ with relatively prime positive integers p and q , it is given as follows. Without loss of generality, we assume that q is odd. The determinant of $T_{p,q}$ is then given as

$$\det(T_{p,q}) = |\Delta_{T_{p,q}}(-1)| = \begin{cases} 1 & \text{if } p \text{ is odd,} \\ q & \text{if } p \text{ is even,} \end{cases}$$

so

$$\text{Arf}(T_{p,q}) = \begin{cases} 0 & \text{if } p \text{ is odd or } q \equiv \pm 1 \pmod{8}, \\ 1 & \text{if } p \text{ is even and } q \equiv \pm 3 \pmod{8}. \end{cases}$$

Along the same lines as Lemma 2.1, Gilmer and Livingston showed the following.

Lemma 2.3 [18, Theorem 10] *Let K be a knot in S^3 that bounds a smoothly embedded punctured Klein bottle F in B^4 . If $\Sigma_2(B^4, F)$, the **double branched cover of B^4 branched over F** , has a positive definite intersection form, then $\sigma(K) + 4 \text{Arf}(K) \equiv 0, 2 \text{ or } 4 \pmod{8}$. Moreover, if $\Sigma_2(B^4, F)$ is negative definite, then $\sigma(K) + 4 \text{Arf}(K) \equiv 0, 4 \text{ or } 6 \pmod{8}$.*

It was an open question until work of Batson [3] whether $\gamma_4(K) \leq 3$ for all knots K in S^3 . Batson found a lower bound on γ_4 using Heegaard Floer homology and used it to show that $\gamma_4(T_{2k,2k-1}) = k - 1$ for all $k \geq 1$. His lower bound was the following.

Lemma 2.4 [3] *For any knot K in S^3 , we have $\frac{1}{2}\sigma(K) - d(S_{-1}^3(K)) \leq \gamma_4(K)$, where $d(S_{-1}^3(K))$ denotes the correction term by Ozsváth and Szabó [42] of the (-1) -surgery of S^3 along K , equipped with the unique Spin^c structure.*

Later, Ozsváth, Stipsicz and Szabó defined an additive concordance invariant $\nu(K)$ for any knot K in S^3 using the Heegaard Floer knot complex [41] and proved the following bound.

Lemma 2.5 [41, Theorem 1.2] *For any knot K in S^3 , we have $|\nu(K) - \frac{1}{2}\sigma(K)| \leq \gamma_4(K)$.*

Using this lower bound, they reproved Batson's result that the nonorientable 4-ball genus of a knot can be arbitrarily large by showing that $\gamma_4(\#_n T_{3,4}) = n$ for all $n \geq 1$ [41, Corollary 1.4].

There are now several lower bounds on γ_4 that look similar to that given in Lemma 2.5. First, Golla and Marengon [19] showed that $\gamma_4(K) \geq \frac{1}{2}\sigma(K) - \min_{m \geq 0} \{m + 2V_m(\bar{K})\}$ for any knot K in S^3 , where $\bar{K}$ denotes the mirror of K and $\{V_i\}_i$ are the concordance invariants explored in [40] (and originally defined in [44]). This bound agrees with the bound in Lemma 2.5 for alternating knots and L -space knots, in particular for torus knots. In addition, Daemi and Scaduto [9] and Ballinger [2] defined invariants h_s and t in gauge theory and Khovanov homology, respectively, from which similar inequalities can be obtained. Again, for many knots, including alternating knots and torus knots, these bounds are not better than the one from Lemma 2.5. Shortly after, Allen showed the following.

Lemma 2.6 [1, Proposition 6.5] *Let K be a knot in S^3 with $\sigma(K) < 2\nu(K)$ such that the **double branched cover of S^3 along K** , denoted by $\Sigma_2(S^3, K)$, is an integer homology 3-sphere and $\delta(K) < 0$, where $\delta(K)$ is a concordance invariant defined by Manolescu and Owens [37]. Then*

$$\nu(K) - \frac{1}{2}\sigma(K) + 1 \leq \gamma_4(K).$$

It is well known that the double cover of S^3 branched along the torus knot $T_{p,q}$ is the Brieskorn sphere $\Sigma(2, p, q)$, which is a homology sphere if and only if p and q are both odd. Therefore, Allen's lower bound does not apply to the infinite families of torus knots in Theorems 1.7 and 1.12.

Finally, given a knot K , one can obstruct the existence of a Möbius band bounded by K (or possibly nonorientable surfaces with higher first Betti number) by computing Frøyshov's h -invariant of $\Sigma_2(S^3, K)$ and appealing to [16, Theorem 3] or by using Donaldson's diagonalization theorem as in [27]. Because the latter method has been used recently to help calculate the nonorientable 4-ball genus of all knots with 8 or 9 crossings in [27] and of all knots with 10 crossings in [17], we will highlight this strategy.

The idea is as follows. Suppose K bounds a Möbius band M and that the double cover of B^4 branched along M — $\Sigma_2(B^4, M)$ — is positive (resp. negative) definite. Suppose moreover that the double cover of S^3 branched along K — $\Sigma_2(S^3, K)$ — bounds a negative (resp. positive) definite 4-manifold X . Then the closed 4-manifold $X \cup_{\Sigma_2(S^3, K)} (-\Sigma_2(B^4, M))$ is negative (resp. positive) definite. By Donaldson's diagonalization theorem [11], there exists a lattice embedding $(H_2(X), Q_X) \rightarrow (\mathbb{Z}^{\text{rank}(H_2(X))+1}, -I)$. If one can show that there is no such embedding, then it follows that K does not bound a Möbius band. See [27] for details. The downside to this obstruction is that it relies on finding a suitable definite 4-manifold bounded by $\Sigma_2(S^3, K)$. The more complicated the knot in question, the more nontrivial it can be to find such a definite 4-manifold and, indeed, such definite 4-manifolds need not exist.

2.2 Obstructions for the existence of topologically locally flat nonorientable surfaces

Given a knot K in S^3 , recall that $\gamma_4^{\text{top}}(K)$ is the minimal first Betti number of locally flat nonorientable surfaces in B^4 bounded by K . Murakami and Yasuhara [39] proved the following lower bound in this category.

Lemma 2.7 [39, Theorem 2.5] *Suppose that a knot K in S^3 bounds a locally flat null-homologous nonorientable surface of first Betti number g in a compact 4-manifold W bounding S^3 with $H_1(W; \mathbb{Z}) = 0$. Then the linking form λ on $H_1(\Sigma_2(S^3, K); \mathbb{Z})$ splits into a direct sum $(G_1, \lambda_1) \oplus (G_2, \lambda_2)$, where the following conditions are satisfied:*

- *There exists a square matrix V with integer entries, whose size is $2b_2(W) + g$ and whose determinant is $\pm|G_1|$, such that λ_1 is presented by $-V^{-1}$.*
- *λ_2 is metabolic.*

Using this lower bound, Gilmer and Livingston [18] were able to prove that there exist knots with $\gamma_4^{\text{top}} \geq 3$. Feller and Golla [13, Proposition 1.4] used it to show that if $p > 0$ with $p \equiv 5 \pmod{8}$, then $T_{p,p \pm 1}$ does

not bound a locally flat Möbius band in B^4 . Moreover, by [13, Proposition 5.6], for each odd prime p and each choice of sign, there are infinitely many positive integers k such that the knot $T_{2p,2kp\pm 1}$ does not bound a locally flat Möbius band in B^4 . To the best of the authors' knowledge, it is unknown whether γ_4^{top} can be arbitrarily large.

2.3 Pinch moves and the pinch number

Consider the torus as a square with opposite edges identified. Then the torus knot $T_{p,q}$ can be represented by the collection of parallel strands in the square with slope p/q . A *pinch move* is a nonorientable band move obtained by attaching an untwisted band between two adjacent strands of $T_{p,q}$. The resulting knot still lies in the torus and is thus a torus knot, which in fact does not depend on the adjacent strands to which we choose to attach the band. A formula for the resulting torus knot was first provided in [3] and proved in [29].

Lemma 2.8 [3; 29, Lemma 2.1] *Let p and q be relatively prime positive integers and consider a diagram of the torus knot $T_{p,q}$ on the flat torus. After applying a pinch move to $T_{p,q}$, the resulting torus knot (up to orientation) is $T(|p - 2t|, |q - 2h|)$, where t and h are the integers uniquely determined by the requirements*

$$\begin{aligned} t &\equiv -q^{-1} \pmod{p} \quad \text{and} \quad t \in \{0, \dots, p-1\}, \\ h &\equiv p^{-1} \pmod{q} \quad \text{and} \quad h \in \{0, \dots, q-1\}. \end{aligned}$$

Suppose performing a pinch move on $T_{p,q}$ yields $T_{r,s}$. This pinch move can either be *positive* or *negative*; it is called positive if $rq - sp > 0$ and negative if $rq - sp < 0$ (cf [29, Lemma 2.4]). Note that since $T_{p,q} = T_{q,p}$ and $T_{r,s} = T_{s,r}$, performing a pinch move on $T_{q,p}$ yields $T_{s,r}$. Further note that the sign of a pinch move depends on the order of p and q ; that is, the sign of the pinch move performed on $T_{p,q}$ yielding $T_{r,s}$ is opposite of the sign of the pinch move performed on $T_{q,p}$ yielding $T_{s,r}$. Finally, the *pinch number* of $T_{p,q}$, which we denote by $\vartheta(T_{p,q})$, is the minimal number of pinch moves needed to turn $T_{p,q}$ into the unknot (which is equivalent to the first Betti number of the nonorientable surface $F_{p,q}$ constructed by Batson [3] discussed in the introduction). Using Lemma 2.8, we can prove the following result, which we will use later.

Lemma 2.9 *Let $p \geq 2$ and $k \geq 1$. Performing a pinch move on the torus knot $T_{p,kp\pm 1}$ yields the torus knot $T_{p-2,k(p-2)\pm 1}$.² Consequently,*

$$\vartheta(T_{p,kp\pm 1}) = \begin{cases} \frac{1}{2}(p-1) & \text{if } p \text{ is odd,} \\ \frac{1}{2}p & \text{if } p \text{ is even and } kp \pm 1 \neq p-1, \\ \frac{1}{2}(p-2) & \text{if } p \text{ is even and } kp \pm 1 = p-1. \end{cases}$$

²When $p = 2$ and $k = 1$, this formula gives $T_{0,-1}$. However (as the proof shows), Lemma 2.8 yields $T_{0,1}$. Since $T_{0,1}$ and $T_{0,-1}$ are isotopic to the unknot, we find it convenient to eliminate the use of absolute values, writing $T_{p-2,k(p-2)\pm 1}$ instead of $T_{|p-2|,|k(p-2)\pm 1|}$.

Moreover, the pinch move performed on $T_{p,kp+1}$ is negative, and the pinch move performed on $T_{p,kp-1}$ is positive if $p > 2$ and negative if $p = 2$.

Proof We perform a pinch move on $T_{p,kp\pm 1}$ with p and k as in the statement of the lemma. Set $t = p - 1$ and $h = k(p - 1) + 1$. Then

$$\begin{aligned} t(kp + 1) &= kp^2 + p - kp - 1 \equiv -1 \pmod{p}, \\ hp &= p(kp + 1) - kp \equiv 1 \pmod{kp + 1}. \end{aligned}$$

Now since $|p - 2t| = p - 2$ and $|kp + 1 - 2h| = k(p - 2) + 1$, Lemma 2.8 shows that performing a pinch move on $T_{p,kp+1}$ yields $T_{p-2,k(p-2)+1}$. Moreover, this pinch move is clearly negative.

Next let $t = 1$ and $h = k$. Then

$$\begin{aligned} t(kp - 1) &\equiv -1 \pmod{p}, \\ hp &= kp \equiv 1 \pmod{kp - 1}. \end{aligned}$$

Note that $|p - 2t| = p - 2$ for all p ; $|kp - 1 - 2h| = k(p - 2) - 1$ for all $p \geq 3$; and $|kp - 1 - 2h| = 1$ when $p = 2$. Thus Lemma 2.8 shows that performing a pinch move on $T_{p,kp-1}$ yields $T_{p-2,k(p-2)-1}$ when $p \geq 3$ and $T_{0,1}$ when $p = 2$. Moreover, this pinch move is clearly positive if $p > 2$ and negative if $p = 2$.

Now, if p is odd, then by performing $\frac{1}{2}(p - 1)$ pinch moves to $T_{p,kp\pm 1}$, we obtain the torus knot $T_{1,k\pm 1}$, which bounds a disk; thus $\vartheta(T_{p,kp\pm 1}) = \frac{1}{2}(p - 1)$. If p is even and $kp \pm 1 \neq p - 1$, then by performing $\frac{1}{2}(p - 2)$ pinch moves to $T_{p,kp\pm 1}$, we obtain the torus knot $T_{2,2k\pm 1}$, which bounds a Möbius band; thus $\vartheta(T_{p,kp\pm 1}) = \frac{1}{2}p$. If p is even and $kp \pm 1 = p - 1$, then by performing $\frac{1}{2}(p - 2)$ pinch moves to $T_{p,p-1}$, we obtain the torus knot $T_{2,1}$, which bounds a disk; thus $\vartheta(T_{p,p-2}) = \frac{1}{2}(p - 2)$. $\square$

It was conjectured in [3] that $\gamma_4(T_{p,q}) = \vartheta(T_{p,q})$. This conjecture is clearly true in some basic cases, as it is not hard to observe that $\gamma_4(T_{2,q}) = 1 = \vartheta(T_{2,q})$ for all odd q and $\gamma_4(T_{3,q}) = 1 = \vartheta(T_{3,q})$ for all q relatively prime to 3. Jabuka and van Cott [29] studied Batson's conjecture and showed that it is true for several infinite families of torus knots, using pinch moves and the lower bound from Lemma 2.5. Their main result is the following.

Theorem 2.10 [29, Theorem 1.10 and Corollary 1.11] *Let $p, q > 1$ be relatively prime integers and suppose that q is odd. Suppose further that there is a sequence of $n \geq 1$ pinch moves $T_{p_i,q_i} \rightarrow T_{p_{i-1},q_{i-1}}$ starting at $T_{p,q} = T_{p_n,q_n}$ and ending at the unknot $T_{p_0,1}$ with $p_0 \geq 0$ for integers $p_i, q_i \geq 1$, $i \in \{1, \dots, n\}$. For all $i \in \{1, \dots, n\}$, let $\varepsilon_i \in \{\pm 1\}$ be $+1$ if the pinch move $T_{p_i,q_i} \rightarrow T_{p_{i-1},q_{i-1}}$ is positive and -1 otherwise. For all $i \in \{1, \dots, n - 1\}$, $m_i = (p_{i+1} + \varepsilon_i \varepsilon_{i+1} p_{i-1}) / p_i$ is an even integer.*

(a) *Let p be odd and $p > q$. If $q_1 \equiv \varepsilon_1 \pmod{4}$ and $m_i \equiv 2 \pmod{4}$ for all $i \in \{1, \dots, n - 1\}$, then $\gamma_4(T_{p,q}) = n$.*

(b) *Let p be even. If $\varepsilon_i = 1$ for all $i \in \{1, \dots, n\}$, then $\gamma_4(T_{p,q}) = n$.*

In both cases, $\gamma_4(T_{p,q}) = \vartheta(T_{p,q}) = n$.

Theorem 2.10 heavily relies on the lower bound $|\nu - \frac{1}{2}\sigma|$ on γ_4 provided by Ozsváth, Stipsicz and Szabó (see [Lemma 2.5](#)). In fact, Jabuka and Van Cott give formulas for this lower bound for all torus knots $T_{p,q}$ in terms of the integers p_i, q_i, ε_i and m_i from [Theorem 2.10](#); see [\[29, Theorems 1.7 and 1.8\]](#). Using these formulas along with [Lemma 2.9](#) and [Theorem 2.10](#), one can prove the following.

Corollary 2.11 (a) *If p is odd and $q \equiv p + 1$ or $2p - 1 \pmod{2p}$, then $\gamma_4(T_{p,q}) = \frac{1}{2}(p - 1)$.*
 (b) *If p is even and $q \equiv p - 1$ or $2p - 1 \pmod{2p}$, then $\gamma_4(T_{p,q}) \in \{\frac{1}{2}(p - 2), \frac{1}{2}p\}$.*

As mentioned in the introduction, Lobb [\[35\]](#) disproved Batson's conjecture by showing that $\gamma_4(T_{4,9}) = 1$, while $\vartheta(T_{4,9}) = 2$. This was achieved by finding a nonorientable band move from $T_{4,9}$ to the Stevedore knot, which is slice. Generalizing his approach, an infinite family of counterexamples was given by Longo [\[36\]](#).

Theorem 2.12 [\[36\]](#) *For $n \geq 2$, we have $\vartheta(T_{4n,(2n\pm 1)^2}) = 2n$, while $\gamma_4(T_{4n,(2n\pm 1)^2}) \leq 2n - 1$.*

In her masters thesis, Tairi [\[46\]](#) gave another infinite family of counterexamples to Batson's conjecture.

Theorem 2.13 [\[46\]](#) (a) *$\vartheta(T_{4,11}) = 2$, while $\gamma_4(T_{4,11}) = 1$.*

(b) *Let $n \geq 0$ and $m \geq 2$. Then $\vartheta(T_{4n+2m+2,10n+6m+5}) = m + 1$, while $\gamma_4(T_{4n+2m+2,10n+6m+5}) = m$.*

As in Lobb's proof that $T_{4,9}$ bounds a Möbius band, Tairi's proof that $T_{4,11}$ bounds a Möbius band relies on finding a nonorientable band move from $T_{4,11}$ to the Stevedore knot. To show that the torus knot $T_{4n+2m+2,10n+6m+5}$ bounds a nonorientable surface F with $g(F) = m$, Tairi shows that performing $m - 1$ pinch moves turns $T_{4n+2m+2,10n+6m+5}$ into $T_{4,11}$, which satisfies $\gamma_4(T_{4,11}) = 1$.

Using this idea, one can start with any torus knot whose nonorientable 4-ball genus is strictly less than its pinch number and create an infinite family of such torus knots by reversing the pinch move operation. For each of Longo's torus knots in [Theorem 2.12](#), we can then produce an infinite family of such torus knots.

Proposition 2.14 *For all $n \geq 2$ and $k \geq 0$,*

$$\vartheta(T_{4n+2k,(4n+2k)(n\pm 1)+1}) = 2n + k, \quad \text{while } \gamma_4(T_{4n+2k,(4n+2k)(n\pm 1)+1}) \leq 2n + k - 1.$$

Proof By repeatedly applying [Lemma 2.9](#), performing k pinch moves on $T_{4n+2k,(4n+2k)(n\pm 1)+1}$ yields the torus knot $T_{4n,4n(n\pm 1)+1} = T_{4n,(2n\pm 1)^2}$. By [Theorem 2.12](#), $\vartheta(T_{4n,(2n\pm 1)^2}) = 2n$ and $\gamma_4(T_{4n,(2n\pm 1)^2}) \leq 2n - 1$. Consequently,

$$\vartheta(T_{4n+2k,(4n+2k)(n\pm 1)+1}) = 2n + k \quad \text{and} \quad \gamma_4(T_{4n+2k,(4n+2k)(n\pm 1)+1}) \leq 2n + k - 1. \quad \square$$

In [Section 3](#), we will see that when n is even, $\gamma_4(T_{4n+2k,(4n+2k)(n\pm 1)+1}) = 2n + k - 1$. In particular, for $k = 0$, this shows that half of Longo's examples in [Theorem 2.12](#) satisfy

$$\gamma_4(T_{4n+2k,(4n+2k)(n\pm 1)+1}) = \vartheta(T_{4n+2k,(4n+2k)(n\pm 1)+1}) - 1.$$

3 Calculations: Proofs of Proposition 1.6 and Theorems 1.7 and 1.12

In this section, we prove Proposition 1.6 and Theorems 1.7 and 1.12. In particular, we use Lemma 2.9 to find upper bounds on γ_4 for these families and Theorem 1.1 to find lower bounds.

3.1 The family $T_{4,q}$

By Lemma 2.9, $\vartheta(T_{4,q}) = 2$ for all $q \geq 5$. Thus we have an obvious upper bound $\gamma_4(T_{4,q}) \leq 2$ for all q . Now, from the formula

$$\begin{aligned} t^{3n} \Delta_{T_{4,2n+1}}(t) &= \frac{(t^{8n+4} - 1)(t - 1)}{(t^4 - 1)(t^{2n+1} - 1)} \\ &= \frac{(t^{2n+1} + 1)(t^{4n+2} + 1)}{(t + 1)(t^2 + 1)} = \left(\sum_{k=0}^{2n} (-t)^k \right) \left(\sum_{k=0}^{2n} (-t^2)^k \right), \end{aligned}$$

we see that

$$\Delta_{T_{4,2n+1}}(t) = \sum_{k=0}^{6n} a_k t^{k-3n},$$

where

$$\begin{aligned} a_0 &= 1, & a_{6n} &= 1, & a_{3n+1} &= 0, \\ a_{3n} &= \begin{cases} (-1)^{3n/2} \frac{1+(-1)^n}{2} & \text{if } n \text{ is even,} \\ (-1)^{(3n+1)/2} \frac{1+(-1)^{(n-1)/2}}{2} & \text{if } n \text{ is odd,} \end{cases} \\ a_{3n+2} &= \begin{cases} (-1)^{(3n+2)/2} \frac{1+(-1)^n}{2} & \text{if } n \text{ is even,} \\ (-1)^{(3n+3)/2} \frac{1+(-1)^{(n-1)/2}}{2} & \text{if } n \text{ is odd.} \end{cases} \end{aligned}$$

Thus,

$$\Delta_{T_{4,2n+1}}(t) = \begin{cases} t^{-3n} - \dots - 1 + t^2 + \dots + t^{3n} & \text{if } n \equiv 2 \text{ or } 3 \pmod{4}, \\ t^{-3n} - \dots + 1 - t^2 + \dots + t^{3n} & \text{if } n \equiv 0 \text{ or } 1 \pmod{4}. \end{cases}$$

Hence, applying Theorem 1.1 gives $\gamma_4(T_{4,q}) = 2$ for all $q \equiv 5, 7 \pmod{8}$.

On the other hand, Theorem 1.1 cannot be applied to $T_{4,q}$ when $q \equiv 1$ or $3 \pmod{8}$. However, the obstruction from the linking form on the first homology of the double cover of $T_{4,q}$ from Lemma 2.7 does obstruct $T_{4,q}$ from bounding a Möbius band for infinitely many $q \equiv 1$ or $3 \pmod{8}$. In fact, Theorem 1.13 (proved in Section 4) implies that

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \gamma_4^{\text{top}}(T_{4,8k+1})}{n} = \lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n \gamma_4^{\text{top}}(T_{4,8k+3})}{n} = 2.$$

This means that Batson's conjecture is also true for torus knots of the form $T_{4,8k+1}$ and $T_{4,8k+3}$ for almost all $k \geq 0$. As mentioned in Remark 1.14, we can explicitly find values of $q \equiv 1$ or $3 \pmod{8}$ such that $T_{4,q}$ does not bound a locally flat Möbius band. See Example 5.11 for details.

It is worth noting here that [Theorem 1.1](#) gives a sharper lower bound than the other known lower bounds detailed in [Section 2](#) when $q \equiv 5$ or $7 \pmod{8}$. Using the recursive formulas for $\sigma(T_{p,q})$ [\[22\]](#), $\nu(T_{p,q})$ [\[14\]](#), and $d(S_{-1}^3(T_{-p,q}))$ [\[3\]](#), and the formula in [Remark 2.2](#), we see that

$$\begin{aligned} \sigma(T_{4,q}) &= \begin{cases} -2(q-1) & \text{if } q \equiv 1 \pmod{4}, \\ -2q & \text{if } q \equiv 3 \pmod{4}, \end{cases} & \nu(T(4,q)) &= -(q-1) \quad \text{for all odd } q, \\ d(S_{-1}^3(T_{-4,q})) &= \begin{cases} q-1 & \text{if } q \equiv 1 \pmod{8}, \\ q-1 & \text{if } q \equiv 3 \pmod{8}, \\ q+1 & \text{if } q \equiv 5 \pmod{8}, \\ q+1 & \text{if } q \equiv 7 \pmod{8}, \end{cases} & \text{Arf}(T_{4,q}) &= \begin{cases} 0 & \text{if } q \equiv 1, 7 \pmod{8}, \\ 1 & \text{if } q \equiv 3, 5 \pmod{8}. \end{cases} \end{aligned}$$

It follows that

$$\begin{aligned} (\sigma + 4 \text{Arf})(T_{4,q}) \pmod{8} &= \begin{cases} 0 & \text{if } q \equiv 1 \pmod{8}, \\ 2 & \text{if } q \equiv 3 \pmod{8}, \\ 4 & \text{if } q \equiv 5 \pmod{8}, \\ 6 & \text{if } q \equiv 7 \pmod{8}, \end{cases} \\ |(\nu - \tfrac{1}{2}\sigma)(T(4,q))| &= \begin{cases} 0 & \text{if } q \equiv 1 \pmod{4}, \\ 1 & \text{if } q \equiv 3 \pmod{4}, \end{cases} \\ \tfrac{1}{2}\sigma(T_{-4,q}) - d(S_{-1}^3(T_{-4,q})) &= \begin{cases} 0 & \text{if } q \equiv 1 \pmod{8}, \\ 1 & \text{if } q \equiv 3 \pmod{8}, \\ -2 & \text{if } q \equiv 5 \pmod{8}, \\ -1 & \text{if } q \equiv 7 \pmod{8}. \end{cases} \end{aligned}$$

We can see that [Lemma 2.1](#) only applies when $q \equiv 5 \pmod{8}$ and the lower bound $|\frac{1}{2}\sigma - \nu|$ from [Lemma 2.5](#) does not give as sharp a lower bound as [Theorem 1.1](#) when $q \equiv 5, 7 \pmod{8}$. Consequently, the lower bounds due to Golla and Marengon, Ballinger, and Daemi and Scaduto mentioned in [Section 2](#) are not as sharp either. Notice that Batson's lower bound (from [Lemma 2.4](#)) shows that

$$\tfrac{1}{2}\sigma(T_{-p,q}) - d(S_{-1}^3(T_{-p,q})) \leq \gamma_4(T_{-p,q}) = \gamma_4(T_{p,q})$$

(since $T_{-p,q}$ bounds a nonorientable surface of first Betti number g if and only if $T_{p,q}$ bounds a nonorientable surface of first Betti number g). Therefore, Batson's lower bound is also not as sharp as [Theorem 1.1](#) when $q \equiv 5, 7 \pmod{8}$. Meanwhile, Allen's bound from [Lemma 2.6](#) does not apply since the double cover of S^3 branched over $T_{4,q}$ is not an integral homology sphere. Finally, the obstruction from [Lemma 2.7](#) applies to the cases $q \equiv 5$ or $7 \pmod{8}$ in the same way as it applies to the cases $q \equiv 1$ or $3 \pmod{8}$; however, it does not obstruct every such torus knot from bounding a Möbius band (see [Example 5.11](#) for details).

We finally mention that the techniques of [\[27\]](#) can be used to show that $T_{4,q}$ does not bound a Möbius band for all $q \equiv 5$ or $7 \pmod{8}$, although it requires considerable more work than simply applying [Theorem 1.1](#). To use this obstruction, one must first locate certain definite 4-manifolds bounded by the double branched cover along $T_{4,q}$ and then obstruct the existence of certain lattice embeddings given by Donaldson's diagonalization theorem [\[11\]](#).

3.2 A detour: Semigroups and the Alexander polynomial

To apply [Theorem 1.1](#), one must compute the Alexander polynomial. Such computations are often tedious, especially when considering infinite families of knots. To help with these calculations we exploit a relationship between the Alexander polynomial of torus knots (or in general, algebraic knots) and certain subsemigroups of $\mathbb{Z}_{\geq 0}$.

Recall that a knot K is an *L-space knot* if it admits a positive Dehn surgery which is an *L-space*; see [\[43, Definition 1.1\]](#) for the definition of an *L-space*. It is proven in [\[43, Theorem 1.2\]](#) that when K is an *L-space knot*, the Alexander polynomial of K takes the form

$$(3-1) \quad \Delta_K(t) = (-1)^m + \sum_{i=1}^m (-1)^{m-i} (t^{n_i} + t^{-n_i})$$

for some sequence of positive integers $0 < n_1 < \cdots < n_m = g(K)$.

The *formal semigroup* of an *L-space knot* K , defined in [\[50, Section 2.1\]](#) (see also [\[4\]](#)), is the subset $S_K \subset \mathbb{Z}_{\geq 0}$ satisfying the following equation in $\mathbb{Z}[[t, t^{-1}]]$:

$$(3-2) \quad t^{\deg \Delta_K} \cdot \Delta_K(t) = (1-t) \cdot \left(\sum_{i \in S_K} t^i \right).$$

Recall that a knot K is an *algebraic knot* if it arises as a knot associated to a plane curve singularity in $\mathbb{C}^2$. It is known that algebraic knots are *L-space knots* [\[23\]](#) and that the formal semigroup S_K of an algebraic knot K is an actual semigroup and it agrees with the traditionally defined semigroup for algebraic knots [\[49\]](#). Torus knots are algebraic knots because they are precisely the knots associated to the singularity $z^p - w^q = 0$ for $z, w \in \mathbb{C}$. Therefore, each torus knot gives rise to a semigroup. Moreover, it is well known that the semigroup for $T_{p,q}$ is $S_{T_{p,q}} = p\mathbb{Z}_{\geq 0} + q\mathbb{Z}_{\geq 0}$ (see eg [\[6\]](#)) and that the Alexander polynomial of $T_{p,q}$ is the Laurent polynomial given by

$$\Delta_{T_{p,q}}(t) = t^{-(p-1)(q-1)/2} \cdot \frac{(t-1)(t^{pq}-1)}{(t^p-1)(t^q-1)},$$

which has degree $\frac{1}{2}(p-1)(q-1)$.

Example 3.1 Consider the torus knot $K = T_{3,5}$. Its Alexander polynomial is given by

$$\Delta_K(x) \doteq x^{-4} \cdot \frac{(x-1)(x^{15}-1)}{(x^3-1)(x^5-1)} = x^{-4} - x^{-3} + x^{-1} - 1 + x - x^3 + x^4.$$

On the other hand, the semigroup generated by 3 and 5 is given by

$$\langle 3, 5 \rangle = \{0, 3, 5, 6\} \cup \mathbb{Z}_{\geq 8}.$$

This matches the equality given by (3-2):

$$\begin{aligned} x^4(x^{-4} - x^{-3} + x^{-1} - 1 + x - x^3 + x^4) &= 1 - x + x^3 - x^4 + x^5 - x^7 + x^8 \\ &= (1-x) \left(1 + x^3 + x^5 + x^6 + \sum_{i=8}^{\infty} x^i \right). \end{aligned}$$

Since algebraic knots are L -space knots, their Alexander polynomials are as in (3-1). Coupling this with (3-2), we have the following lemma, which was also stated (and proven) in a slightly different way in [5].

Lemma 3.2 [5, Proposition 3.8] *Let K be an algebraic knot whose Alexander polynomial has leading terms t^d and t^{-d} . Then, for any integer i , the coefficient of the term t^i of $\Delta_K(t)$ is 1 if $d + i - 1 \notin S_K$ and $d + i \in S_K$, -1 if $d + i - 1 \in S_K$ and $d + i \notin S_K$, and 0 if $d + i - 1$ and $d + i$ are either both contained in S_K or both not contained in S_K .*

3.3 The family $T_{p,(2k+1)p-1}$, where p is odd

We now prove the first statement of Theorem 1.7(a) (the second statement of Theorem 1.7(a) is due to Corollary 2.11). Suppose p is odd and let $q \equiv p - 1 \pmod{2p}$. Set $q = 2kp + (p - 1) = (2k + 1)p - 1$. Notice that by Lemma 2.9, $\vartheta(T_{p,(2k+1)p-1}) = \frac{1}{2}(p - 1)$ and so $\gamma_4(T_{p,(2k+1)p-1}) \leq \frac{1}{2}(p - 1)$.

Consider the semigroup S_K for the torus knot $K = T_{p,(2k+1)p-1}$, which is generated by p and $(2k + 1)p - 1$. The degree of the Alexander polynomial Δ_K of K is $d = \frac{1}{2}(p - 1)((2k + 1)p - 2)$. We first claim that $d + i \notin S_K$ for all $0 \leq i \leq \frac{1}{2}(p - 3)$. Suppose otherwise. Then there exist nonnegative integers a and b satisfying

$$ap + b((2k + 1)p - 1) = \frac{(p - 1)((2k + 1)p - 2)}{2} + i.$$

Solving for a , we find

$$a = \frac{(p - 1)(2k + 1)p - 2p - 2b(2k + 1)p}{2p} + \frac{1 + i + b}{p}.$$

Since p is odd and $a \in \mathbb{Z}$, we necessarily have that $1 + i + b = \alpha p$ for some integer α . Since $b \geq 0$, we have that $b = \alpha p - i - 1 \geq 0$, implying that $\alpha \geq 1$. Therefore,

$$\begin{aligned} a &= \frac{(p - 1)(2k + 1)p - 2p - 2(\alpha p - i - 1)(2k + 1)p + 2\alpha p}{2p} \\ &= \frac{-2\alpha p((2k + 1)p - 1) + (2k + 1)p(p - 1 + 2(i + 1)) - 2p}{2p} \\ &\leq \frac{(2k + 1)p(-2p + 2i + p + 1)}{2p} \leq \frac{-2p(2k + 1)}{2p} < 0 \end{aligned}$$

using $i \leq \frac{1}{2}(p - 3)$ in the second-to-last inequality, which contradicts the assumption that $a \geq 0$. Thus $d + i \notin S_K$ for all $0 \leq i \leq \frac{1}{2}(p - 3)$. It follows from Lemma 3.2 that since the constant term of the Alexander polynomial is nonzero and $d \notin S_K$, the constant term of the Alexander polynomial is -1 . Moreover, since $d + i \notin S_K$ for all $0 \leq i \leq \frac{1}{2}(p - 3)$, we can further deduce by Lemma 3.2 that the first nontrivial term of positive degree in Δ_K has exponent at least $\frac{1}{2}(p - 1)$. Finally, setting $a = 0$ and $b = \frac{1}{2}(p - 1)$, we have that $ap + b((2k + 1)p - 1) = \frac{1}{2}((2k + 1)p - 1)(p - 1) = d + \frac{1}{2}(p - 1)$. Therefore, $d + \frac{1}{2}(p - 1) \in S_K$. By Lemma 3.2, the first nontrivial term of positive degree in Δ_K has exponent $\frac{1}{2}(p - 1)$. By Theorem 1.1, $\gamma_4(T_{p,(2k+1)p-1}) \geq \frac{1}{2}(p - 3)$. Thus we have shown $\gamma_4(T_{p,(2k+1)p-1}) \in \{\frac{1}{2}(p - 3), \frac{1}{2}(p - 1)\}$.

3.4 The family $T_{p,(2k+1)p+1}$, where p is even

Finally, we prove the middle case of [Theorem 1.7\(b\)](#) (the other cases of [Theorem 1.7\(b\)](#) are due to [Corollary 2.11](#)). Let $q \equiv p + 1 \pmod{2p}$. Set $q = 2kp + p + 1 = (2k + 1)p + 1$. By [Lemma 2.9](#), $\vartheta(T_{p,(2k+1)p+1}) = \frac{1}{2}p$ and thus $\gamma_4(T_{p,(2k+1)p+1}) \leq \frac{1}{2}p$.

As above, let S_K denote the semigroup for $T_{p,(2k+1)p+1}$, which is generated by p and $(2k + 1)p + 1$. We claim that $d + i \notin S_K$ for all $0 \leq i \leq \frac{1}{2}(p - 2)$, where $d = \frac{1}{2}(p - 1)(2k + 1)p$ is the degree of the Alexander polynomial. Indeed, if there exist nonnegative integers a and b satisfying

$$ap + b((2k + 1)p + 1) = d + i,$$

then

$$a = \frac{(p - 1)(2k + 1)}{2} - b(2k + 1) + \frac{i - b}{p}.$$

Since p is even and $a \in \mathbb{Z}$, $\frac{i-b}{p/2}$ must be odd. Thus $i - b = \alpha \frac{p}{2}$, where α is an odd integer. Consequently,

$$0 \leq b = i - \alpha \frac{p}{2} \leq -\frac{\alpha - 1}{2}p - 1$$

(using $i \leq \frac{1}{2}(p - 2)$) and so $\alpha \leq -1$ (using that α is odd). Thus,

$$\begin{aligned} a &= \frac{(p - 1)(2k + 1)}{2} - b(2k + 1) + \frac{i - b}{p} \\ &= \frac{2k + 1}{2} \underbrace{\left(p - 1 - 2 \left(i - \alpha \frac{p}{2} \right) \right)}_{= p(\alpha + 1) - (2i + 1)} + \frac{\alpha}{2} \leq -\frac{(2i + 1)(2k + 1)}{2} < 0, \end{aligned}$$

contradicting the assumption that $a \geq 0$. Moreover, setting $a = \frac{1}{2}(p - 1)(2k + 1) + 1$ and $b = 0$, we see that $ap + b((2k + 1)p + 1) = d + \frac{1}{2}p \in S_K$. Now arguing as in [Section 3.3](#), we have that the Alexander polynomial of K has constant term -1 and the first nontrivial term of positive degree has exponent $\frac{1}{2}p$. Thus by [Theorem 1.1](#), $\gamma_4(T_{p,(2k+1)p-1}) \geq \frac{1}{2}(p - 2)$.

3.5 The family $T_{4n+2k,(4n+2k)(n \pm 1)+1}$

Finally, we finish the proof of [Proposition 1.6](#). Notice that since n is even, the torus knots in this family are contained in the set of torus knots considered in [Section 3.4](#) (cf [Remark 1.8](#)). Thus we have $\gamma_4(T_{4n+2k,(4n+2k)(n \pm 1)+1}) \in \{2n + k - 1, 2n + k\}$. Now, by [Proposition 2.14](#),

$$\gamma_4(T_{4n+2k,(4n+2k)(n \pm 1)+1}) = 2n + k - 1.$$

Remark 3.3 By [Lemma 2.9](#), all of the pinch moves used for the $T_{p,(2k+1)p+1}$ family in [Section 3.4](#) are negative; since p is even, [\[29, Theorem 1.8\]](#) tells us that the lower bound given by [Lemma 2.5](#) is 0. Similarly, by [Lemma 2.9](#), all of the pinch moves used for the $T_{p,(2k+1)p-1}$ family in [Section 3.3](#) are positive. However, since $(2k + 1)p - 1$ is even, to apply [\[29, Theorem 1.8\]](#), we reverse the role of p and $(2k + 1)p - 1$ so that all pinch moves are negative. By doing so we find that the lower bound

given by Lemma 2.5 is 0. Finally, one can see that the lower bound in Lemma 2.4 is, in general, not as sharp as Theorem 1.1 for these families (eg for $T_{8,9}$, the lower bound given by Lemma 2.4 is -4). Thus Theorem 1.1 gives a much sharper lower bound than the existing lower bounds discussed in Section 2.

4 Involutive obstruction and genus bounds from the unoriented link Floer TQFT

In this section, we will prove Theorems 1.15, 1.16, and 1.1. We begin by briefly recalling Fan's disoriented link Floer TQFT, as defined in [12]. Although Fan's TQFT accounts for nonorientable cobordisms between links in S^3 , there are a number of restrictions of which one should be aware.

Definition 4.1 A *disoriented link* (in S^3) is a quadruple $\mathcal{L} = (L, \mathbf{p}, \mathbf{q}, \mathbf{l})$ where L is a link in S^3 , $\mathbf{p}$ and $\mathbf{q}$ are finite subsets of L satisfying $\mathbf{p} \cap \mathbf{q} = \emptyset$, such that points in $\mathbf{p}$ and points in $\mathbf{q}$ appear alternately on each component of L , and $L \setminus (\mathbf{p} \cup \mathbf{q})$ consists of $2n$ arcs $\mathbf{l} = \{l_1, \dots, l_{2n}\}$ for some integer n .

Remark 4.2 For simplicity, we will occasionally refer to a disoriented link $\mathcal{L} = (L, \mathbf{p}, \mathbf{q}, \mathbf{l})$ and its underlying link L interchangeably.

Given a disoriented link $\mathcal{L} = (L, \mathbf{p}, \mathbf{q}, \mathbf{l})$, where $\mathbf{p} = \{p_1, \dots, p_n\}$ and $\mathbf{q} = \{q_1, \dots, q_n\}$, choose any orientation on L . Viewing the points $\mathbf{p}$ and $\mathbf{q}$ on L as z -basepoints and w -basepoints respectively gives the link Floer chain complex $\text{CFL}(S^3, \mathcal{L})$ a structure of a curved complex of modules over the ring $\mathbb{F}_2[U_1, \dots, U_n, V_1, \dots, V_n]$, as shown in [53]. The curvature of this complex is given by the formula

$$\partial_{\text{CFL}(S^3, \mathcal{L})}^2 = \sum_{K \in \pi_0(L)} U_{p_{K,1}} V_{q_{K,1}} + V_{q_{K,1}} U_{p_{K,2}} + \dots + V_{q_{K,t}} U_{p_{K,1}},$$

where the points of $\mathbf{p}$ and $\mathbf{q}$ on a component K are labeled $p_{K,1}, q_{K,1}, \dots, p_{K,t}, q_{K,t}$ as one travels along L with the chosen orientation. Consider the truncation of $\text{CFL}(S^3, \mathcal{L})$ by identifying all of the variables U_i and V_i with a single formal variable U , and denote the resulting complex by $\text{CFL}'(S^3, \mathcal{L})$. Then the curvature vanishes and thus $\text{CFL}'(S^3, \mathcal{L})$ is a chain complex over the ring $\mathbb{F}_2[U]$.

Note that $\text{CFL}(S^3, \mathcal{L})$ is endowed with two gradings $\mathbf{gr}_{\mathbf{p}}$ and $\mathbf{gr}_{\mathbf{q}}$. The truncation is not compatible with this bigrading, and we only get a single grading, called the (normalized) δ -grading, which is given by $\delta = \frac{1}{2}(\sigma(L) + \mathbf{gr}_{\mathbf{p}} + \mathbf{gr}_{\mathbf{q}})$. Note that unlike in the setting of $\text{CFL}(S^3, \mathcal{L})$, U lowers the normalized δ grading by 1. The resulting δ -graded chain complex $\text{CFL}'(S^3, \mathcal{L})$ is called the *unoriented link Floer chain complex* of $\mathcal{L}$ and its homology is called the *unoriented link Floer homology*. When $\mathcal{L}$ is a knot, we write CFK' instead of CFL' and similarly CFK for CFL .

Definition 4.3 A *disoriented link cobordism* (in $S^3 \times I$) between disoriented links $\mathcal{L}_1 = (L_1, \mathbf{p}_1, \mathbf{q}_1, \mathbf{l}_1)$ and $\mathcal{L}_2 = (L_2, \mathbf{p}_2, \mathbf{q}_2, \mathbf{l}_2)$ is a pair $\mathcal{F} = (F, a)$ where F is a smoothly embedded surface in $S^3 \times I$ such that $\partial F = (L_1 \times \{0\}) \cup (L_2 \times \{1\})$ and a is an oriented 1-manifold which is properly embedded in $(F, \partial F)$ such that $F \setminus a$ is orientable and $\partial a = \mathbf{q}_1 - \mathbf{p}_1 + \mathbf{p}_2 - \mathbf{q}_2$.

The main theorem of [12] is as follows. Note that, while the theorem is originally stated on the level of homology, one can observe directly from its proof that it actually holds on chain level up to (a possibly nonunique) chain homotopy. Slightly more specifically, the map F_S is constructed by decomposing the disoriented link cobordism into a composition of four types of elementary disoriented link cobordisms [12, Chapter 2] to each component of which one assigns a δ -grading preserving $\mathbb{F}_2[U]$ -module map. Different decompositions into elementary cobordisms are related by a collection of moves (see [12, Section 7.1]), each of which preserves the filtered chain homotopy class — see Step 6 in the proof of [12, Theorem 1.3.1].

Theorem 4.4 [12, Theorem 1.3.1] *Let S be a disoriented link cobordism between (disoriented) links L_1 and L_2 . Then there is a δ -grading-preserving $\mathbb{F}_2[U]$ -module map*

$$F_S: \text{CFL}'(S^3, L_1) \rightarrow \text{CFL}'(S^3, L_2),$$

called the cobordism map induced by S , whose chain homotopy class depends only on the isotopy class of S , as a disoriented link cobordism. This map is functorial in the following sense: if S_1 and S_2 are disoriented link cobordisms between L_1 and L_2 and between L_2 and L_3 , respectively, then we have functoriality up to chain homotopy:

$$F_{S_2 \circ S_1} \sim F_{S_2} \circ F_{S_1}.$$

Remark 4.5 For oriented link cobordisms, the notion of disoriented cobordism is equivalent to the notion of decorated cobordism in [30; 53]. This equivalence is a little subtle. The decoration in a decorated cobordism with underlying surface Σ is an embedded 1-manifold in Σ whose boundary is required to be disjoint from the basepoints on the link. On the other hand, we have that the decoration in a disoriented cobordism is required to coincide with the basepoints on the link. Nevertheless, there is a notion of compatibility between disoriented and decorated orientable link cobordisms [20, Remark 4.14] and it is known that the cobordism maps from Zemke's oriented link Floer TQFT induce the cobordism maps from Fan's unoriented link Floer TQFT under truncation for appropriate compatible cobordisms [20, Remark 4.15].

For later use, we will give some trivial examples of disoriented cobordisms for which Fan's cobordism maps can be easily computed.

Example 4.6 Consider the unknot U , made into a disoriented knot $\mathcal{U}$ by choosing one p point and one q point. Performing a nonorientable band move on U with normal Euler number ± 1 gives two nonorientable cobordisms $\mathcal{S}_\pm$ from $\mathcal{U}$ to $\mathcal{U}$. From the δ -grading shift formula (see [12, Proposition 7.2.5]), we see that the grading shifts of $F_{\mathcal{S}_+}$ and $F_{\mathcal{S}_-}$ are 0 and -1 , respectively. Now, from the fact that the cobordism maps $F_{\mathcal{S}_\pm}$ induced by $\mathcal{S}_\pm$ become quasi-isomorphisms after localizing by U^{-1} (which follows directly from [20, Proposition 5.4]), we get

$$F_{\mathcal{S}_+} = \text{id} \quad \text{and} \quad F_{\mathcal{S}_-} = U \cdot \text{id}.$$

In this section, we will use not only (disoriented) cobordisms in $S^3 \times I$, but also cobordisms in B^4 . They can be seen as disoriented cobordisms from a nontrivial link to the empty link, or from the empty link to a nontrivial link. To define cobordism maps for such cobordisms, given a disoriented cobordism (S, a) in B^4 , we choose any p in the interior of a and its sufficiently small neighborhood $N(p)$ in B^4 , so that $S \cap N(p)$ is a boundary-parallel disk in B^4 and $a \cap N(p)$ is a properly embedded arc on $S \cap N(p)$. We then consider $(S \setminus N(p), a \setminus N(p))$, which is now a disoriented cobordism in $S^3 \times I$, from some disoriented link L to the unknot. Then the map $F_{(S,a)}$ induced by (S, a) can be naturally defined as the disoriented cobordism map $F_{(S \setminus N(p), a \setminus N(p))}$ induced by the punctured cobordism $(S \setminus N(p), a \setminus N(p))$ by Theorem 4.4. We can similarly define the opposite-direction cobordism map $F_{(-S,a)}$. Note that this definition is compatible with the birth/death maps of unknots in Zemke's oriented link Floer TQFT.

Proposition 4.7 *For any disoriented cobordism (S, a) in B^4 , the chain homotopy class of the map $F_{(S,a)}$ depends only the isotopy class of (S, a) . The same statement also holds for $(-S, a)$.*

Proof Choose any two points p and q in the interior of a and choose sufficiently small neighborhoods $N(p)$ and $N(q)$ of p and q , respectively, so that $N(p) \cap N(q) = \emptyset$ and $S \cap N(p)$. We claim that $F_{(S \setminus N(p), a \setminus N(p))} = F_{(S \setminus N(q), a \setminus N(q))}$; the desired result follows directly from this claim. To prove the claim, we first observe that the case when S is orientable is already covered by Remark 4.5. In the general case when S can be nonorientable, choose any smoothly embedded ball $B \subset B^4$ such that $N(p) \cup N(q) \subset B$, $S \cap B$ is orientable, and ∂B intersects transversely with both S and a . Since $(S \setminus B, a \setminus B)$ and $(S \cap (B \setminus N(p)), a \cap (B \setminus N(p)))$ are disoriented cobordisms in $S^3 \times I$, we have

$$F_{(S \setminus N(p), a \setminus N(p))} \simeq F_{(S \setminus B, a \setminus B)} \circ F_{(S \cap (B \setminus N(p)), a \cap (B \setminus N(p)))} \simeq F_{(S \setminus B, a \setminus B)} \circ F_{(S \cap B, a \cap B)}.$$

Similarly, we also have $F_{(S \setminus N(q), a \setminus N(q))} \simeq F_{(S \cap B, a \cap B)}$. The claim follows. $\square$

We now introduce involutive knot Floer homology. Given an oriented knot K in S^3 , we can represent it as a doubly pointed Heegaard diagram $H = (\Sigma, \alpha, \beta, z, w)$, and suppose that K lies on the Heegaard surface Σ . We have a canonical chain skew-isomorphism

$$\eta: \text{CFK}(H) \rightarrow \text{CFK}(\bar{H}),$$

where $\bar{H}$ denotes the Heegaard diagram $(-\Sigma, \beta, \alpha, w, z)$. Here, the prefix “skew” means that we are considering maps which intertwine the actions of U and V on the domain with the actions of V and U on the codomain.

We then consider the half-twist map

$$\tau: \text{CFK}(\bar{H}) \rightarrow \text{CFK}(\tau(\bar{H}))$$

which corresponds to the self-diffeomorphism $(S^3, K, z, w) \rightarrow (S^3, K, w, z)$ induced by rotating a neighborhood of K (in Σ) halfway around K in the direction of the orientation induced by K . Here, $\tau(\bar{H})$ denotes the Heegaard diagram $(-\Sigma, \tau(\beta), \tau(\alpha), z, w)$, where $\tau(\beta)$ and $\tau(\alpha)$ are the images of β

and α under the given “half-twist” self-diffeomorphism. Now, the doubly pointed Heegaard diagrams $\tau(\bar{H})$ and H encode the same oriented knot, so the naturality of link Floer homology [31] tells us that there is a chain homotopy equivalence

$$n: \text{CFK}(\tau(\bar{H})) \rightarrow \text{CFK}(H),$$

which is well defined up to chain homotopy. Composing these three maps gives a homotopy self skew-equivalence

$$\iota_K = n \circ \tau \circ \eta,$$

whose chain homotopy class depends only on the oriented isotopy class of K . It is known [52] that ι_K is not an involution in general, but rather an order 4 map which satisfies

$$\iota_K^2 \sim 1 + \Phi\Psi,$$

where Φ and Ψ are the basepoint actions, which can be easily computed as the formal partial derivatives of the chain differential with respect to the formal variables U and V , respectively. Note that, since the truncation from CFK to CFK' identifies U and V into a single variable U , we see that ι_K induces a δ -grading-preserving $\mathbb{F}_2[U]$ -equivariant chain homotopy self-equivalence

$$\iota_K: \text{CFK}'(S^3, K) \rightarrow \text{CFK}'(S^3, K),$$

whose chain homotopy class depends on the oriented isotopy class of K . It is clear that, if we choose the opposite orientation of K , the map ι_K is replaced by its homotopy inverse ι_K^{-1} .

Since ι_K is now a U -equivariant chain homotopy equivalence of $\mathbb{F}[U]$ -modules rather than a skew-equivariant chain homotopy equivalence of $\mathbb{F}[U, V]$ -modules (which is the case in the oriented setting), we can define new knot invariants as follows.

Definition 4.8 Given an unoriented knot K , choose either orientation on K and define the *involutive unoriented knot Floer chain complex* $\text{CFKI}'(K)$ as the mapping cone of

$$1 + \iota_K: \text{CFK}'(S^3, K) \rightarrow \text{CFK}'(S^3, K),$$

which is a chain complex of $\mathbb{F}_2[U, Q]/(Q^2)$ -modules. This complex is endowed with a grading which is uniquely determined from the δ -grading on $\text{CFK}'(S^3, K)$ by the rule that the grading of Q is -1 . The homology of $\text{CFKI}'(K)$, called the *involutive unoriented knot Floer homology*, is denoted $\text{HFKI}'(K)$.

We note our grading conventions for the mapping cone differ from the conventions used in Hendricks and Manolescu’s definition of involutive Heegaard Floer homology [24].

Remark 4.9 If we denote the differential on $\text{CFK}(S^3, K)$ as ∂_K , then the differential ∂ on $\text{CFKI}'(K)$ is given by $\partial(a + Qb) = \partial_K(a) + Q(a + \iota_K(a) + \partial_K(b))$.

It is easy to see that the above definition is independent of the choice of orientation on K .

Lemma 4.10 *Let C and D be chain complexes and $g: C \rightarrow D$ be a chain homotopy equivalence. Then for any chain map $f: C \rightarrow C$, the mapping cones of f and $g \circ f$ are chain homotopy equivalent.*

Proof The differentials on the mapping cones of f and $g \circ f$ are given by $\partial_{M(f)}(a, b) = (\partial a, f(a) + \partial b)$ and $\partial_{M(g \circ f)}(a, b) = (\partial a, g(f(a)) + \partial b)$, respectively. Consider the maps $F: M(f) \rightarrow M(g \circ f)$ and $G: M(g \circ f) \rightarrow M(f)$, defined as $F(a, b) = (a, g(b))$ and $G(a, b) = (a, h(b))$, where h is a homotopy inverse of g . Then we have

$$(\partial_{M(g \circ f)} F + F \partial_{M(f)})(a, b) = (\partial a, (g \circ f)(a) + \partial g(b)) + (\partial a, g(f(a) + \partial b)) = 0,$$

so F is a chain map, and similarly G is also a chain map. Now, if we denote a null-homotopy of $h \circ g - \text{id}$ as H , then we have

$$G \circ F - \text{id} = (0, h \circ g - \text{id}) = \partial_{M(f)} \circ (0, H) - (0, H) \circ \partial_{M(f)},$$

so $G \circ F \sim \text{id}$ and similarly $F \circ G \sim \text{id}$. Therefore F is a chain homotopy equivalence. $\square$

Theorem 4.11 *Given a knot K , the chain homotopy equivalence class of $\text{CFKI}'(K)$ as a chain complex of $\mathbb{F}_2[U, Q]/(Q^2)$ -modules depends only on the isotopy class of K .*

Proof Choose an orientation on K , which then defines ι_K on $\text{CFKI}'(K)$. We have to prove that the mapping cone of $1 + \iota_K$ is chain homotopy equivalent to the mapping cone of $1 + \iota_K^{-1}$. But since

$$1 + \iota_K^{-1} = \iota_K^{-1} \circ (1 + \iota_K)$$

and ι_K is a chain homotopy self-equivalence of $\text{CFKI}'(K)$, [Lemma 4.10](#) implies that the two mapping cones are indeed chain homotopy equivalent. $\square$

The following lemma tells us about the structure of HFKI' . For simplicity, we denote the localization by introducing a formal inverse of U as U^{-1} .

Lemma 4.12 $U^{-1}\text{HFKI}'(K) \simeq \mathbb{F}_2[U^{\pm 1}, Q]/(Q^2).$

Proof Since $U^{-1}\text{CFK}'(S^3, K)$ is the truncation of $(U, V)^{-1}\text{CFK}(S^3, K)$ by the identification $U = V$ and $(U, V)^{-1}\text{CFK}(S^3, K) \simeq (U, V)^{-1}\mathbb{F}_2[U, V]$, we know that

$$U^{-1}\text{CFK}'(S^3, K) \simeq U^{-1}\mathbb{F}_2[U] \simeq \mathbb{F}_2[U^{\pm 1}].$$

Thus $U^{-1}\iota_K$, the localization of ι_K at U , is chain homotopic to the identity map. Now the lemma follows from the fact that a localization of a mapping cone is chain homotopy equivalent to the mapping cone of a localized map. $\square$

The above lemma, combined with [Theorem 4.11](#), allows us to make the following definition.

Definition 4.13 Given a knot K , we define two knot invariants $\bar{\nu}(K)$ and $\underline{\nu}(K)$ as follows:

$$\underline{\nu}(K) = \max\{r \mid \exists x \in \text{HFKI}'(K), \mathbf{gr}_\delta(x) = r, \forall n, U^n x \notin \text{Im}(Q)\},$$

$$\bar{\nu}(K) = \max\{r \mid \exists x \in \text{HFKI}'(K), \mathbf{gr}_\delta(x) = r, \forall n, U^n x \neq 0; \exists m \geq 0 \text{ s.t. } U^m x \in \text{Im}(Q)\} + 1,$$

where $\mathbf{gr}_\delta(x)$ denotes the (normalized) δ -grading of a homogeneous element x .

We note that the grading shifts in the above definition differ from the corresponding grading shifts in the definitions of $\underline{d}$ and $\bar{d}$ given by Hendricks and Manolescu [24, Section 5] on account of the different grading convention for mapping cones we used in our definition of involutive unoriented knot Floer homology.

As mentioned in the introduction, the invariants $\underline{\nu}$ and $\bar{\nu}$ coincide with the values at $t = 1$ of the involutive Upsilon invariants $\underline{\Upsilon}_t$ and $\bar{\Upsilon}_t$ defined by Hogancamp and Livingston. It thus follows from [25, Proposition 11 and Theorem 12] that $\underline{\nu}$ and $\bar{\nu}$ are (smooth) concordance invariants and

$$\bar{\nu}(K) \geq \nu(K) \geq \underline{\nu}(K).$$

Example 4.14 Consider the figure-eight knot $K = 4_1$. Its unoriented knot Floer complex has generators a, b, c, d , and x , whose gradings are all 0, and where the differential is given by

$$\partial a = U(b + c), \quad \partial b = \partial c = Ud, \quad \partial x = \partial d = 0.$$

The involution ι_K , which was computed in [24, Section 8.2], is given by

$$\iota_K(a) = a + x, \quad \iota_K(b) = c, \quad \iota_K(c) = b, \quad \iota_K(d) = d, \quad \iota_K(x) = x + d.$$

We will calculate $\bar{\nu}(K)$ and $\underline{\nu}(K)$ by explicitly computing the cycles of $\text{CFKI}'(K)$. Recall that the differential of $\text{CFKI}'(K)$ is given by $\partial(s + Qt) = \partial s + Q(s + \iota_K(s) + \partial t)$. Hence an element $a + Qb \in \text{CFKI}'(K)$ is a cycle if and only if a is a cycle in $\text{CFK}'(S^3, K)$ and $(1 + \iota_K)(a) = \partial b$.

The space of cycles in $\text{CFK}'(S^3, K)$ is generated over $\mathbb{F}_2[U]$ by $x, b + c$, and d , whose images under the action of $1 + \iota_K$ are given by $d, 0$, and 0 , respectively. Hence the space of cycles in $\text{CFKI}'(K)$ is generated over $\mathbb{F}_2[U, Q]/(Q^2)$ by the cycles

$$Ux + Qb, \quad b + c, \quad d, \quad Qx.$$

Let ∂' denote the differential on $\text{CFK}'(K)$. Since $\partial'(b) = Ud + Q(b + c)$, Ud is in the image of Q in $\text{HFKI}'(K)$. Likewise $\partial'(a + Qx) = U(b + c) + Q(x)$, so that $U(b + c)$ is in the image of Q in $\text{HFKI}'(K)$.

On the other hand, if $Ux + Qb$ contained in the image of Q in $\text{HFKI}'(K)$ for some $n \geq 0$, then we should have

$$U^n(Ux + Qb) = Qw + \partial_1(y + Qz)$$

for elements $w, y, z \in \text{CFK}'(S^3, K)$. It follows that $\partial y = U^{n+1}(x)$, a contradiction. Hence the homology class of $U^n(Ux + Qb)$ is not contained in the image of Q for any $n \geq 0$. Finally observe that if $U^n(b + c) = 0$ in homology then there exist elements $z, y \in \text{CFK}'(K)$ and $n \geq 0$ such that $\partial z = U^n(b + c)$ and $(1 + \iota_K)(x) + \partial y = 0$. It follows that $z = U^{n-1}a$, whence $U^{n-1}a + \partial y = 0$, a contradiction.

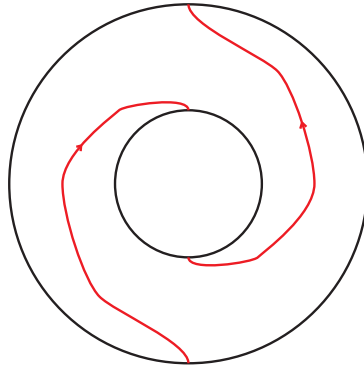


Figure 1: The cobordism which induces the half twist map τ_K . Note that the underlying surface is $K \times I$.

Now, using the definition of $\bar{v}$ and $\underline{v}$, we have

$$\bar{v}(K) = \mathbf{gr}_\delta(b + c) + 1 = 1 \quad \text{and} \quad \underline{v}(K) = \mathbf{gr}_\delta(Ux + Qb) = -1.$$

Together with [Theorem 1.15](#), which will be proven in this section, we recover the fact that the figure-eight knot does not bound a smooth Möbius band in B^4 , which was proven in the topological category in [\[48\]](#). Note that, since the figure-eight knot is rationally slice, we have $v(K) = \sigma(K) = 0$.

In general, via a straightforward computation using model complexes as given in [\[24, Section 8.3\]](#), it is easy to see that the involutive upsilon invariants of a thin knot K (which includes all quasialternating knots) are given by

$$\begin{aligned} \bar{v}(K) = v(K) = \underline{v}(K) = 0 & \quad \text{if } \sigma(K) + 4 \operatorname{Arf}(K) \equiv 0 \pmod{8}, \\ \bar{v}(K) = 1, \quad v(K) = 0, \quad \underline{v}(K) = -1 & \quad \text{if } \sigma(K) + 4 \operatorname{Arf}(K) \equiv 4 \pmod{8}. \end{aligned}$$

We thus see that a thin knot K does not bound a smooth Möbius band in B^4 if $\sigma(K) + 4 \operatorname{Arf}(K) \equiv 4 \pmod{8}$, which is the same as Yasuhara's obstruction from [Lemma 2.1](#).

Now we move on to developing involutive obstructions and bounds on the first Betti number for nonorientable slice-surfaces bounding a given knot. Recall that, given a knot K with two prescribed points $p, q \in K$, the map ι_K on the unoriented knot Floer complex is defined by composing the canonical conjugation map

$$\eta_K: \operatorname{CFK}'(K, p, q) \rightarrow \operatorname{CFK}'(K, q, p)$$

together with the “half twist” map (which depends on an orientation of K)

$$\tau_K: \operatorname{CFK}'(K, q, p) \rightarrow \operatorname{CFK}'(K, p, q).$$

Note that the half twist map can be seen as a cobordism map induced by the “half twist cobordism”, as shown in [Figure 1](#). In particular, it is not a skew-isomorphism anymore; it is a genuine isomorphism, as we are now working with CFK' , where the actions of U and V coincide.

For more generality, we extend the definition of canonical conjugation map to disoriented links. Given any disoriented link $\mathcal{L} = (L, \mathbf{p}, \mathbf{q}, l)$, we also have a canonically defined conjugation map

$$\eta_{\mathcal{L}}: \text{CFL}'(\mathcal{L}) \rightarrow \text{CFL}'(\bar{\mathcal{L}}),$$

where $\bar{\mathcal{L}} = (L, \mathbf{q}, \mathbf{p}, l)$.

Lemma 4.15 *Let $\mathcal{S} = (S, a)$ be a disoriented cobordism between disoriented links $\mathcal{L}_1$ and $\mathcal{L}_2$. Then*

$$F_{(S, -a)} \circ \eta_{\mathcal{L}_1} \sim \eta_{\mathcal{L}_2} \circ F_{(S, a)},$$

where $-a$ denotes a endowed with the reverse orientation.

Proof Recall that the cobordism maps in Fan's unoriented link Floer TQFT are computed by decomposing a given cobordism into elementary disoriented link cobordisms and composing the maps corresponding to each of these pieces. Among all possible pieces, we only have to consider nonorientable saddle moves, since this lemma is verified in all other cases in the proof of [52, Theorem 1.3] (for compatible decorated link cobordisms). Since the cobordism map for a saddle move is defined via a count of pseudoholomorphic triangles on certain simple Heegaard triple-diagrams — see the proof of [12, Theorem 5.1.4] — they homotopy-commute with conjugation maps. $\square$

We now have to check that the half twist map τ_K also homotopy-commutes with (disoriented) cobordism maps. When the given cobordism is a Möbius band, then the commutativity is very easy to observe.

Lemma 4.16 *Let $\mathcal{S} = (S, a)$ be a disoriented cobordism from a disoriented knot $K = (K, \mathbf{p}, \mathbf{q}, l)$ to the empty link, where $\mathbf{p}$ and $\mathbf{q}$ both consist of a single point. Suppose that S is a smoothly embedded Möbius band and a is a simple arc. Then, for any choice of orientation on K , we have $F_{(S, -a)} \circ \tau_K \sim F_{(S, a)}$, where $-a$ denotes a endowed with the reverse orientation. Moreover, if $\bar{S}$ denotes the cobordism from the empty link to K given by flipping S upside-down, then we have $\tau_K \circ F_{(\bar{S}, a)} \sim F_{(\bar{S}, -a)}$.*

Proof This follows directly from Figure 2. $\square$

With Lemma 4.16 at hand, we can now prove Theorem 1.15.

Proof of Theorem 1.15 Let M be a smooth Möbius band in B^4 bounding K . Choose a properly embedded simple arc $a \subset M$ such that $M \setminus a$ is a rectangle. Then (M, a) is a disoriented cobordism from K to the empty link and $(-M, a)$ is a disoriented cobordism from the empty link to K . By choosing an interior point $p \in a \cap \text{int}(M)$ and its small neighborhood $N(p) \subset B^4$, we can consider two disoriented cobordisms $(M \setminus N(p), a \setminus N(p))$ and $(-M \setminus N(p), a \setminus N(p))$, which are cobordisms from K to the unknot and the unknot to K , respectively; note that $F_{(M \setminus N(p), a \setminus N(p))} = F_{(M, a)}$ and $F_{(-M \setminus N(p), a \setminus N(p))} = F_{(-M, a)}$. Then, using [20, Proposition 5.4] on the punctured cobordism $(M \setminus N(p), a \setminus N(p))$, we see that there exists an orientable disoriented cobordism T with first Betti number 1 from K to K , shown in Figure 3, such that

$$U^c \cdot F_T \circ F_{(-M, a)} \circ F_{(M, a)} \sim U^{b-m+1} \cdot \text{id},$$

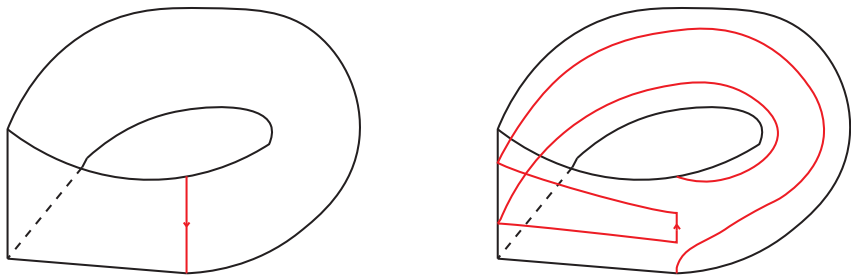


Figure 2: Left, the cobordism $\mathcal{G}$. Right, the cobordism representing $\tau_K \circ (\bar{S}, -a)$. Note that the cobordism on the right is isotopic to the one on the left.

where m , b , and c are the numbers of local minima, saddles, and local maxima of M ; note that $b - m - c = 1$ since M is a Möbius band. Since the underlying space of T is an orientable cobordism, F_T is homotopic to the $U = V$ truncation of the cobordism map induced by T (viewed as a compatible decorated link cobordism) with some orientation in Zemke’s oriented link Floer TQFT. Theorem C of [53] implies that $U^{-1}F_T \sim U \cdot \text{id}$ — where here $U^{-1}F_T$ is the localization of F_T at U — so we see that the localized map $U^{-1}(F_{(-M,a)} \circ F_{(M,a)})$ is a quasi-isomorphism. Since $U^{-1}\text{HFK}'(S^3, K) \simeq \mathbb{F}_2[U^{\pm 1}]$, we see that the localized maps $U^{-1}F_{(-M,a)}$ and $U^{-1}F_{(M,a)}$ are both chain homotopic to the identity. Also, since $U^{-1}F_T \sim U \cdot \text{id}$, we know that F_T shifts the δ -grading by -1 . Hence, if we denote the δ -grading shifts of $F_{(-M,a)}$ and $F_{(M,a)}$ by $\text{gr}_\delta(F_{(-M,a)})$ and $\text{gr}_\delta(F_{(M,a)})$, then we have

$$\text{gr}_\delta(F_{(-M,a)}) + \text{gr}_\delta(F_{(M,a)}) = -1.$$

Choose either orientation on K . Lemmas 4.15 and 4.16 imply that $F_{(M,a)}$ and $F_{(-M,a)}$ satisfy

$$F_{(M,a)} \circ \iota_K \sim F_{(M,a)}, \quad \iota_K \circ F_{(-M,a)} \sim F_{(-M,a)}.$$

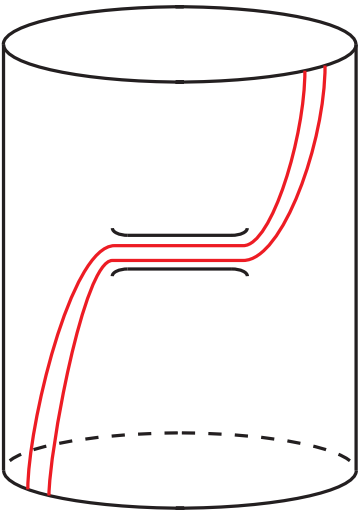


Figure 3: The orientable cobordism T of first Betti number 1.

Hence, by choosing relevant chain homotopies, we get maps between involutive homology,

$$F: \text{HFKI}'(K) \rightarrow \mathbb{F}_2[U, Q]/(Q^2),$$

$$G: \mathbb{F}_2[U, Q]/(Q^2) \rightarrow \text{HFKI}'(K),$$

where the grading shifts of F and G are given by $\mathbf{gr}_\delta(F_{(M,a)})$ and $\mathbf{gr}_\delta(F_{(-M,a)})$, respectively. Note that F and G become quasi-isomorphisms when localized by U^{-1} .

Now, from the definition of $\bar{v}$, we know that there exists an element $x \in \text{HFKI}'(K)$ such that the following conditions are satisfied.

- The δ -grading of x is $\bar{v}(K) - 1$.
- For any $n \geq 0$, we have $U^n x \neq 0$.
- $U^m x = Qy$ for some $m \geq 0$ and $y \in \text{HFKI}'(K)$.

Consider the element $F(x) \in \mathbb{F}_2[U, Q]/(Q^2)$. Since the localized map $U^{-1}F$ is chain homotopic to U times the identity map and the homology class of x generates the tower given by the image of Q inside the (U) -localized homology, we know that $F(x) \neq 0$. Also, we have $U^i F(x) = QF(y)$, so $F(x) = QU^i$ for some $i \geq 0$. Hence the δ -grading of $F(x)$ is at most -1 , so we get

$$\bar{v}(K) \leq -\mathbf{gr}_\delta(F_{(M,a)}).$$

We now consider the element $1 \in \text{CFKI}'(\text{unknot}) = \mathbb{F}_2[U, Q]/(Q^2)$. Then $G(1)$ is a cycle in $\text{CFKI}'(K)$. Suppose that there exists some $m \geq 0$ such that the homology class of $U^m G(1)$ is a multiple of Q . Then there exist a cycle $d \in \text{CFK}'(S^3, K)$ and chains $e, f \in \text{CFK}'(S^3, K)$ such that

$$G(U^m) = U^m G(1) = Qd + \partial(e + Qf) = \partial e + Q(d + e + \iota_K(e) + \partial f).$$

Since $G(U^m)$ is of the form $F_{(-M,a)}(U^m) + Qc$ for some $c \in \text{CFK}'(S^3, K)$, we must have

$$F_{(-M,a)}(U^m) = \partial e,$$

ie the homology class $[F_{(-M,a)}(U^m)] \in \text{HFK}'(K)$ vanishes. But this contradicts the fact that the localized map $U^{-1}F_{(-M,a)}$ is chain homotopic to identity. Hence we deduce that $U^m G(1) \notin \text{Im}(Q)$ for any $m \geq 0$, which implies that

$$\underline{v}(K) \geq \mathbf{gr}_\delta(F_{(-M,a)}).$$

Therefore we deduce that

$$\bar{v}(K) - \underline{v}(K) \leq -(\mathbf{gr}_\delta(F_{(-M,a)}) + \mathbf{gr}_\delta(F_{(M,a)})) = 1. \quad \square$$

We move on to prove an involutive bound on the nonorientable genus. For this we require the following lemmas.

Lemma 4.17 [12, Proposition 6.6.2] *Let $S_i = (\Sigma, a_i)$, $i = 1, 2, 3$, be disoriented link cobordisms on the same underlying unoriented surface Σ . Suppose that S_1, S_2 , and S_3 are related by a **bypass move**, which is defined as in Figure 4. Then we have $F_{S_1} + F_{S_2} + F_{S_3} \sim 0$.*

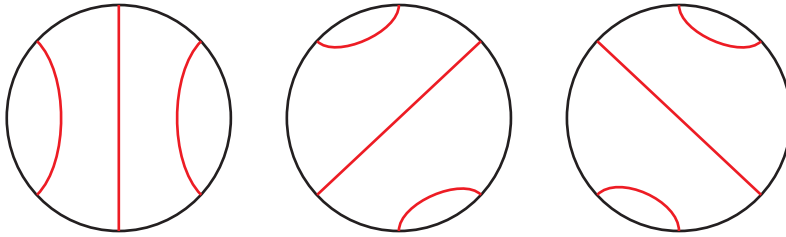


Figure 4: We say that three disoriented cobordisms are related by a *bypass move* if they differ only inside a disk as drawn. See [12, Definition 7.1.5] for more details.

Lemma 4.18 *Let $S = (S, a)$ be a disoriented link cobordism from a disoriented link L to the empty link. Suppose that the one-manifold a contains a closed component c which is **two-sided**, ie admits an orientable neighborhood in S . Then $U \cdot F_S \sim 0$ and $U \cdot F_{\bar{S}} \sim 0$, where $\bar{S}$ is given by flipping S upside-down.*

Proof Choose a disk D , smoothly embedded in the interior of B^4 , which intersects S transversely in its interior and satisfies $\partial D = c$. Choose a small neighborhood N_D of D . By perturbing a , we may assume that $S_1 = (S \setminus N_D, a \setminus N_D)$ and $S_2 = (S \cap N_D, a \cap N_D)$ are disoriented cobordisms and $c \subset N_D$. Note that, since $S = S_2 \circ S_1$, we have $F_S = F_{S_2} \circ F_{S_1}$.

Since c is two-sided and the interior of D intersects S in finitely many points, $S \cap N_D$ is orientable. Thus, once again replacing the disoriented cobordism structure on S_1 and S_2 with the structure of compatible decorated link cobordisms, by [32, Lemma 5.2] and Remark 4.5, we deduce that $U \cdot F_{S_2} \sim 0$. Therefore $U \cdot F_S = (U \cdot F_{S_2}) \circ F_{S_1} \sim 0$. The proof for $\bar{S}$ is the same. $\square$

Now, we can prove that the disoriented cobordism maps induced by surfaces of higher first Betti number homotopy-commute with τ_K up to multiplication by U .

Lemma 4.19 *Let $\mathcal{S} = (S, a)$ be a disoriented cobordism from a disoriented knot $\mathcal{K} = (K, p, q, l)$ to the empty link, where p and q consist of a single point. Suppose that S is a smoothly embedded nonorientable surface and a is a simple arc. Then, for any choice of orientation on K ,*

$$U \cdot (F_{(S,a)} \circ \tau_K) \sim U \cdot F_{(S,-a)},$$

where $-a$ denotes a endowed with the reverse orientation. Also, if $\bar{S}$ denotes the cobordism from the empty link to K given by flipping S upside-down, then $U \cdot (\tau_K \circ F_{(\bar{S},a)}) \sim U \cdot F_{(\bar{S},-a)}$.

Proof We will only provide a proof for (S, a) ; the proof for $(\bar{S}, a)$ is the same.

We first assume that the first Betti number of S is odd. Then we can consider S , as an abstract surface with boundary, as a connected sum of a Möbius band and a closed orientable surface, and a as an arc on S which is shown on the left in Figure 5. Choose an orientation on K and denote by C the disoriented cobordism from $\mathcal{K}$ to (K, q, p, l) given by a half-twist along the given orientation on K ; note that the

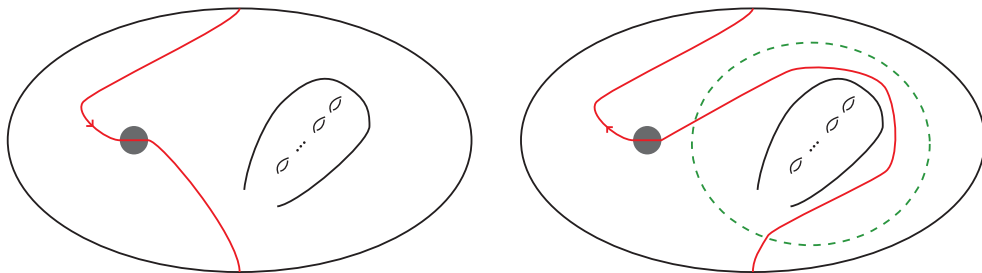


Figure 5: Left, the cobordism $\mathcal{S}$. Right, the cobordism $\mathcal{S}'$. The region enclosed by the dashed curve is S_o . The gray circles denote crosscaps, ie Möbius bands attached along their boundaries.

cobordism map induced by C is τ_K . Then $\mathcal{S} \circ C$ is isotopic to the disoriented cobordism $\mathcal{S}' = (S, b)$, where b is the arc shown on the right in Figure 5.

The oriented arcs $-a$ and b differ in an oriented subsurface S_o of S . Denote the first Betti number of S_o by n . Then one can consider a sequence of properly embedded (oriented) arcs $-a = a_0, a_1, \dots, a_n = b$, as drawn in Figure 6, such that

- $a_0, \dots, a_n$ differ only in S_o ;
- for each i , $S_o \setminus a_i$ consists of a surface A_i of genus i and a surface B_i of genus $n - i$, where A_i lies on the right of a_i ;
- $A_0 \subset A_1 \subset \dots \subset A_n$.

Note that, since a_i is oriented and divides S_o into two components, we have a well-defined notion of a component lying on the left side and the right side of a_i . Since we have Lemma 4.18 and a bypass relation for Fan's TQFT, we can directly apply the arguments used in the proof of [32, Lemma 5.3 and Proposition 5.5] to show that $U \cdot F_{(S, a_{i-1})} \sim U \cdot F_{(S, a_i)}$ for all i . Hence we get $U \cdot (F_{\mathcal{S}} \circ \tau_K) \sim U \cdot F_{(S, -a)}$.

We now consider the remaining case when the first Betti number of S is even. Then, by Example 4.6, we see that taking a connected sum of $\mathcal{S}$ with a unknotted $\mathbb{RP}^2$ with normal Euler number 2 gives a disoriented cobordism of odd first Betti number which has the same induced cobordism map as $\mathcal{S}$. This reduces the problem to the case of odd first Betti number, so we get $U \cdot (F_{\mathcal{S}} \circ \tau_K) \sim U \cdot F_{(S, -a)}$, as desired. $\square$

We are now able to prove Theorem 1.16. Our proof is similar to that of equation 18 in [32, Theorem 1.7].

Proof of Theorem 1.16 Let S be a smoothly embedded nonorientable surface of first Betti number g in B^4 which bounds K . Suppose that S consists of M local maxima, b (possibly nonorientable) saddles, and $m + 1$ local minima. Then, by applying [20, Proposition 5.5] as in the proof of Theorem 1.15, there exists an orientable disoriented cobordism T of first Betti number 3, given as shown in Figure 3, and a properly embedded oriented arc $a \subset S$, such that $S \setminus a$ is orientable and

$$U^M \cdot (F_T \circ F_{(-S, a)} \circ F_{(S, a)}) \sim U^{b-m+1} \cdot \text{id}.$$

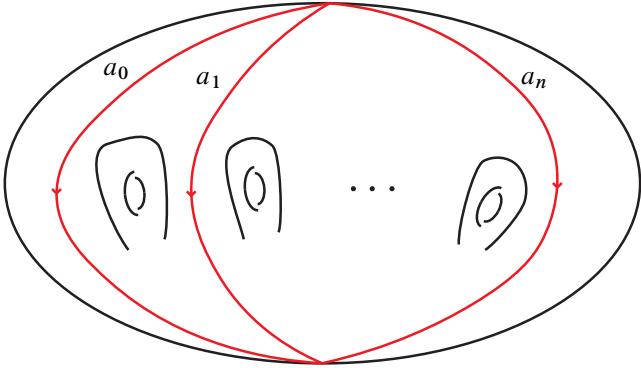


Figure 6: The arcs $a_0, a_1, \dots, a_n$ on the orientable surface S_δ .

Following the proof of [Theorem 1.15](#) tells us that the localized maps $U^{-1}F_{(-S,a)}$ and $U^{-1}F_{(S,a)}$ are chain homotopic to the identity map, and if we denote the δ -grading shifts of $F_{(-S,a)}$ and $F_{(S,a)}$ by $\mathbf{gr}_\delta(F_{(-S,a)})$ and $\mathbf{gr}_\delta(F_{(S,a)})$, then we have

$$\mathbf{gr}_\delta(F_{(-S,a)}) + \mathbf{gr}_\delta(F_{(S,a)}) = M + m - b = g.$$

However, since $F_{(S,a)}$ and $F_{(-S,a)}$ may not homotopy-commute with ι_K , they do not induce maps between $\mathrm{HFKI}'(K)$ and $\mathrm{HFKI}'(\mathrm{unknot}) = \mathbb{F}_2[U, Q]/(Q^2)$. Fortunately, from [Lemmas 4.15](#) and [4.19](#), we see that

$$U \cdot (\iota_K \circ F_{(S,a)}) \sim U \cdot F_{(S,a)}, \quad U \cdot (\iota_K \circ F_{(-S,a)}) \sim U \cdot F_{(-S,a)},$$

so $U \cdot F_{(S,a)}$ and $U \cdot F_{(-S,a)}$ induce maps between involutive homology,

$$\begin{aligned} F &: \mathrm{HFKI}'(K) \rightarrow \mathbb{F}_2[U, Q]/(Q^2), \\ G &: \mathbb{F}_2[U, Q]/(Q^2) \rightarrow \mathrm{HFKI}'(K), \end{aligned}$$

where the grading shifts of F and G are denoted by $\mathbf{gr}_\delta(F_{(S,a)}) - 1$ and $\mathbf{gr}_\delta(F_{(-S,a)}) - 1$, respectively.

Now, following the proof of [Theorem 1.15](#) gives

$$\bar{v}(K) \leq 1 - \mathbf{gr}_\delta(F_{(S,a)}), \quad \underline{v}(K) \geq \mathbf{gr}_\delta(F_{(-S,a)}) - 1.$$

Furthermore, using the same argument for $v(K)$ implies

$$\mathbf{gr}_\delta(F_{(-S,a)}) \leq v(K) \leq -\mathbf{gr}_\delta(F_{(S,a)}).$$

The theorem then follows directly from the above inequalities. □

We now focus on the case of L -space knots. The knot Floer homology of L -space knots is very simple: if K is an L -space knot, then $\mathrm{CFK}(S^3, K)$ is a staircase complex which is completely determined by the Alexander polynomial of K , as shown in [\[43, Theorem 1.2\]](#). Recall that the Alexander polynomial of K always takes the form

$$\Delta_K(t) = (-1)^m + \sum_{i=1}^m (-1)^{m-i} (t^{n_i} + t^{-n_i})$$

for some sequence of positive integers $0 < n_1 < \cdots < n_m = g(K)$. The chain complex $\text{CFK}(S^3, K)$ is chain homotopy equivalent to a free chain complex generated by elements $x_0, x_1^1, x_1^2, \dots, x_m^1, x_m^2$. Here, x_0 lies on the bigrading $(0, 0)$ and either

- (1) the bigradings of x_{m-i}^1 and x_{m-i-1}^1 differ by $(n_{m-i} - n_{m-i-1}, 0)$ if i is even and $(0, n_{m-i} - n_{m-i-1})$ if i is odd; or
- (2) the bigradings of x_{m-i}^1 and x_{m-i-1}^1 differ by $(n_{m-i} - n_{m-i-1}, 0)$ if i is odd and $(0, n_{m-i} - n_{m-i-1})$ if i is even.

Moreover, if x_s^1 lies on the bigrading (i, j) , then x_s^2 lies on the bigrading (j, i) . The differential on the staircase complex is given as in [24, Figure 12].

The action of ι_K on such a staircase complex is very easy to describe: we have $\iota_K(x_0) = x_0$, $\iota_K(x_s^1) = x_s^2$, and $\iota_K(x_s^2) = x_s^1$. In particular, we have $\iota_K^2 \sim \text{id}$. See [24, Section 7] for more details.

Lemma 4.20 *Let K be an L -space knot. If the Alexander polynomial of K is of the form*

$$\Delta_K(t) = -1 + \sum_{i=1}^m (-1)^{m-i} (t^{n_i} + t^{-n_i})$$

for some sequence of positive integers $0 < n_1 < \cdots < n_m$, then

$$\bar{v}(K) = v(K) \geq \underline{v}(K) + n_1.$$

Proof The unoriented knot Floer chain complex of L -space knots can be written as a graded root, as described in [20, Section 7]. The involution ι_K then acts on the graded root as a reflection along the central axis, so we can consider $(\text{CFK}'(S^3, K), \iota_K)$ as a symmetric graded root, which is defined in [10, Definition 2.11]. See Figure 7 for a pictorial description of a symmetric graded root induced by the action of ι_K on $\text{CFK}'(S^3, K)$.

It is shown in [10, Section 6] that the invariants d and $\bar{d}$ can be read off from a symmetric graded root. As v and $\bar{v}$ are defined analogously to d and $\bar{d}$, Dai and Manolescu's arguments likewise show that $v(K)$ can be read off from the given symmetric graded root as the grading of the highest leaf, and $\bar{v}(K) = v(K)$. Furthermore, $\underline{v}(K)$ is the grading of the highest node which is left invariant by the action of ι_K . Since the constant term of the Alexander polynomial of K is -1 , we have $\partial x_0 = U^{n_1}(x_1^1 + x_1^2)$ in $\text{CFK}'(K)$, so $\underline{v}(K)$ is represented by the cycle $U^{n_1}x_1^1 + Qx_0$ in $\text{CFKI}'(K)$. Also, since Qx_1 is a cycle which satisfies

$$U^{n_1} \cdot Qx_1 = Q \cdot (U^{n_1}x_1^1 + Qx_0),$$

we see that U^{n_1} times the homology class of Qx_1 is contained in the image of Q and cannot be annihilated by any powers of U . Therefore we get

$$\bar{v}(K) - \underline{v}(K) \geq (\text{gr}(Qx_1) + 1) - \text{gr}(U^{n_1}x_1^1 + Qx_0) = n_1. \quad \square$$

We are finally ready to prove Theorem 1.1.

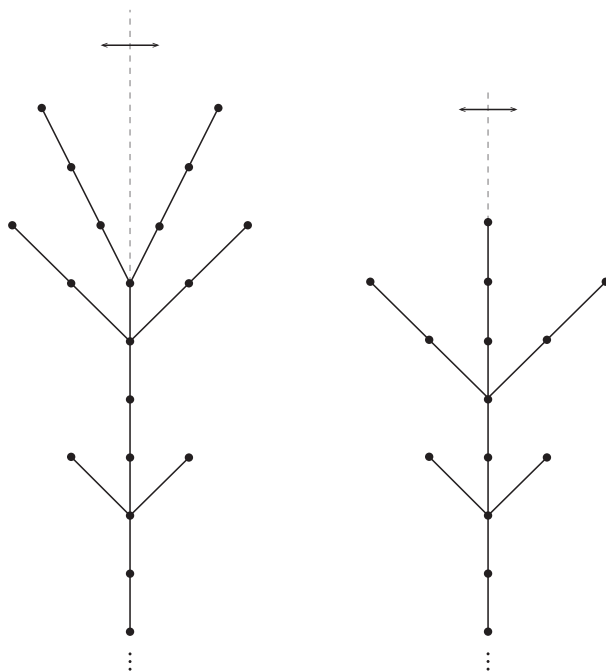


Figure 7: Left, the symmetric graded root representing $\text{CFK}'(S^3, T_{6,7})$. Right, the symmetric graded root representing $\text{CFK}'(S^3, T_{5,6})$. Note that the action of ι_K is given by reflection along the central axis, given by the gray dashed line. Note moreover that $\Delta_K(0) \neq -1$ for these two knots, so [Lemma 4.20](#) does not apply.

Proof of Theorem 1.1 Since torus knots admit lens space surgeries [38], and lens spaces are L -spaces [43], torus knots are L -space knots. The theorem follows directly from Theorems 1.15 and 1.16 and [Lemma 4.20](#). $\square$

5 Asymptotically obstructing locally flat Möbius bands

In this section we prove [Theorem 1.13](#) by studying the linking forms of double branched covers of torus knots. It is well known that the double branched cover over $T_{p,q}$ is the Brieskorn sphere $\Sigma_2(S^3, T_{p,q}) = \Sigma(2, p, q)$. Some number theory and a result of [51] allow us to show that the linking form obstructs certain torus knots $T_{p,q}$ — subject to restrictions on p and q — from bounding locally flat Möbius bands. A subsequent argument shows that the restrictions on p and q are appropriately generic.

5.1 The linking form of $\Sigma(2, p, q)$, where p is even and q is odd

We begin by computing the linking forms of the Brieskorn sphere $\Sigma(2, p, q)$, for any even p and odd q such that $\gcd(p, q) = 1$. Recall that the linking form of a rational homology 3-sphere M is a $\mathbb{Q}/\mathbb{Z}$ -valued bilinear form

$$\lambda_M: H_1(M; \mathbb{Z}) \times H_1(M; \mathbb{Z}) \rightarrow \mathbb{Q}/\mathbb{Z}.$$

To compute the linking form $\lambda_{\Sigma(2,p,q)}$ on $\Sigma(2, p, q)$, we start with a Goeritz matrix of $T_{p,q}$; recall that, in [21, Section 3], it is shown that the linking form of the double cover of S^3 branched over a knot K is presented by $-G^{-1} \bmod 1$, where G is a Goeritz matrix of K .

Consider the checkerboard surface F for $T_{p,q}$, where p is even, given in Figure 8. Note that F is built from disks (0-handles) and bands (1-handles). Following [21], the Goeritz matrix G of $T_{p,q}$ given by F is the $(\frac{1}{2}(pq - 2q + 2) \times \frac{1}{2}(pq - 2q + 2))$ -matrix

$$G = \begin{pmatrix} q & -\mathbf{1}_q^h & 0 & 0 & \cdots & 0 & 0 \\ -\mathbf{1}_q^v & A & -I_q & 0 & \cdots & 0 & 0 \\ 0 & -I_q & A & -I_q & \ddots & 0 & 0 \\ 0 & 0 & -I_q & A & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & A & -I_q \\ 0 & 0 & 0 & 0 & \cdots & -I_q & A \end{pmatrix},$$

where I_q is the $q \times q$ identity matrix, $\mathbf{1}_q^h$ and $\mathbf{1}_q^v$ are the horizontal and vertical vectors of size q whose entries are all 1, respectively, and A is the $(q \times q)$ -matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Topologically, G can be viewed as a matrix for the intersection form of the double cover of B^4 branched over F , where we consider F as a properly embedded surface by pushing its interior into B^4 . Denote this 4-manifold by X . Then $\partial X = \Sigma(2, p, q)$ and X is simply connected. Instead of computing the full inverse of G , we will focus on two entries of G^{-1} , namely the $(1, \frac{1}{2}(pq - 2q + 2))$ -entry and $(\frac{1}{2}(pq - 2q + 2), \frac{1}{2}(pq - 2q + 2))$ -entry. It will turn out that these two entries are enough to recover the linking form. Note that $\det(G) = \det(T_{p,q}) = |\Delta_{T_{p,q}}(-1)| = q$ if p is even and q is odd. Since all entries of G are integers, we thus see that all entries of G^{-1} are integer multiples of $1/q$.

Lemma 5.1 *The $(1, \frac{1}{2}(pq - 2q + 2))$ -entry of G^{-1} is $1/q$.*

Proof Consider the horizontal vector

$$v = \left(\frac{p}{2}, \underbrace{\frac{p}{2} - 1, \dots, \frac{p}{2} - 1}_q, \underbrace{\frac{p}{2} - 2, \dots, \frac{p}{2} - 2}_q, \dots, \underbrace{2, \dots, 2}_q, \underbrace{1, \dots, 1}_q \right)$$

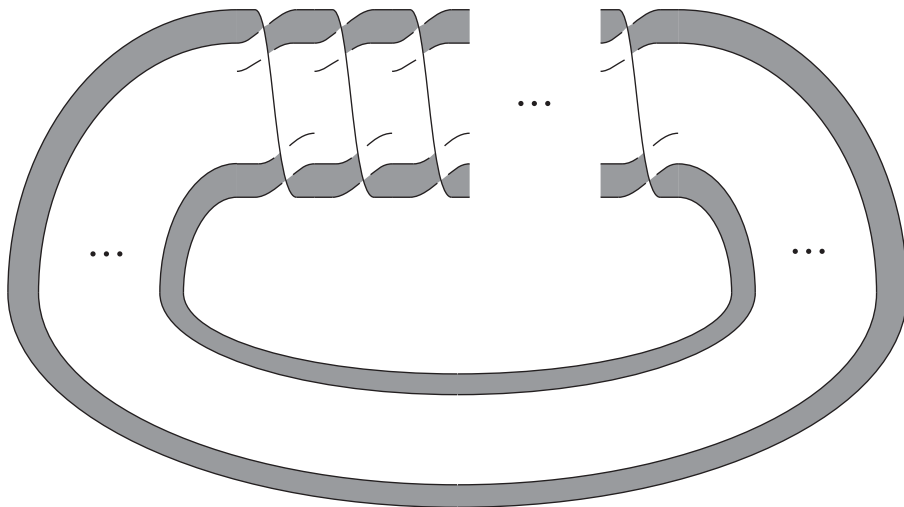


Figure 8: A checkerboard surface F for the torus knot $T_{p,q}$, where p is even.

of length $\frac{1}{2}(pq - 2q + 2)$. We claim that $vG = (q, 0, \dots, 0)$. Denote the k^{th} entry of vG by $(vG)_k$. When $k = 1$,

$$(vG)_1 = \frac{1}{2}p \cdot q + q \cdot \left(\frac{1}{2}p - 1\right) \cdot (-1) = q.$$

When $2 \leq k \leq q + 1$,

$$(vG)_k = \frac{1}{2}p \cdot (-1) + 2 \cdot \left(\frac{1}{2}p - 1\right) \cdot 1 + \left(\frac{1}{2}p - 2\right) \cdot (-1) = 0.$$

For $k \geq q + 2$, it is easy to check that we also have $(vG)_k = 0$. Thus $vG = (q, 0, \dots, 0)$ and so v is the top row of qG^{-1} . Therefore the $(1, \frac{1}{2}(pq - 2q + 2))$ -entry of G^{-1} is given by $(1/q) \cdot 1 = 1/q$. $\square$

Lemma 5.2 The $(\frac{1}{2}(pq - 2q + 2), \frac{1}{2}(pq - 2q + 2))$ -entry of G^{-1} is of the form m/q , where m is an integer satisfying $\frac{1}{2}mp \equiv 1 \pmod{q}$.

Proof Let $v = (v_0, v_1^1, \dots, v_1^q, v_2^1, \dots, v_2^q, \dots, v_{\frac{p}{2}-1}^1, \dots, v_{\frac{p}{2}-1}^q)$ be the i^{th} row of the matrix qG^{-1} . For simplicity, our notation of v_k^r here is chosen to be cyclic on $r \pmod{q}$, ie $v_k^{r+q} = v_k^r$. Furthermore, we define $v_0^r = v_0$ for any r .

We will show that for any fixed $k \in \{1, \dots, \frac{1}{2}p - 1\}$, the integers $v_k^1, \dots, v_k^q$ are congruent mod q . Since v is the i^{th} row of qG^{-1} , vG is of the form qe_i , where e_i is the i^{th} standard basis vector. Thus every entry of vG is either 0 or q . Hence the following conditions hold:

- For any indices k and r such that $0 \leq k \leq \frac{1}{2}p - 3$, we have $v_{k+1}^r + v_{k+1}^{r+2} \equiv v_k^{r+1} + v_{k+2}^{r+1} \pmod{q}$.
- For any index r , we have $v_{\frac{p}{2}-2}^{r+1} \equiv v_{\frac{p}{2}-1}^r + v_{\frac{p}{2}-1}^{r+2} \pmod{q}$.

We claim that for any r and k with $0 \leq k \leq \frac{p}{2} - 1$, we have

$$v_k^r + v_k^{r+2} + \dots + v_k^{r+p-2k} = v_{k-1}^{r+1} + v_{k-1}^{r+3} + \dots + v_{k-1}^{r+p-2k-1} \pmod{q}.$$

Using the second condition above as the base case, inductively assume that

$$v_{k+1}^r + v_{k+1}^{r+2} + \cdots + v_{k+1}^{r+p-2k-2} \equiv v_k^{r+1} + v_k^{r+3} + \cdots + v_k^{r+p-2k-3} \pmod{q}.$$

If $\frac{1}{2}(p-2k)$ is even, then by the first condition, the quantity

$$v_k^r + v_k^{r+2} + v_k^{r+4} + v_k^{r+6} + \cdots + v_k^{r+p-2k-2} + v_k^{r+p-2k}$$

is equivalent modulo q to

$$v_{k-1}^{r+1} + v_{k-1}^{r+5} + \cdots + v_{k-1}^{r+p+2k-7} + v_{k-1}^{r+p+2k-1} + v_{k+1}^{r+1} + v_{k+1}^{r+5} + \cdots + v_{k+1}^{r+p+2k-7} + v_{k+1}^{r+p+2k-1}.$$

Applying the inductive hypothesis to the latter half of this expression, we see that this is equivalent modulo q to

$$v_{k-1}^{r+1} + v_{k-1}^{r+5} + \cdots + v_{k-1}^{r+p+2k-7} + v_{k-1}^{r+p+2k-1} + v_k^{r+2} + v_k^{r+4} + \cdots \\ \cdots + v_k^{r+p-2k-2} - v_{k+1}^{r+3} - v_{k+1}^{r+7} - \cdots - v_{k+1}^{r+p+2k-7} - v_{k+1}^{r+p+2k-1}.$$

Finally, applying the first condition to the terms $v_k^{r+2} + v_k^{r+4} + \cdots + v_k^{r+p-2k-2}$, the above expression is equivalent modulo q to

$$v_{k-1}^{r+1} + v_{k-1}^{r+3} + \cdots + v_{k-1}^{r+p-2k-1},$$

as desired.

On the other hand, if $\frac{1}{2}(p-2k)$ is odd, then by the first condition, the quantity

$$v_k^r + v_k^{r+2} + v_k^{r+4} + v_k^{r+6} + \cdots + v_k^{r+p-2k-4} + v_k^{r+p-2k-4} + v_k^{r+p-2k}$$

is equivalent modulo q to

$$v_{k-1}^{r+1} + v_{k-1}^{r+5} + \cdots + v_{k-1}^{r+p+2k-7} + v_{k-1}^{r+p+2k-3} + v_{k+1}^{r+1} + v_{k+1}^{r+5} + \cdots \\ \cdots + v_{k+1}^{r+p+2k-7} + v_{k+1}^{r+p+2k-3} + v_k^{r+p-2k}.$$

Applying the inductive hypothesis to the latter half of the portion of the expression preceding v_k^{r+p-2k} , we see that this is equivalent modulo q to

$$v_{k-1}^{r+1} + v_{k-1}^{r+5} + \cdots + v_{k-1}^{r+p+2k-7} + v_{k-1}^{r+p+2k-3} + v_k^{r+2} + v_k^{r+4} + \cdots \\ \cdots + v_k^{r+p-2k} - v_{k+1}^{r+3} - v_{k+1}^{r+7} - \cdots - v_{k+1}^{r+p+2k-3}.$$

Finally, applying the first condition to the terms $v_k^{r+2} + v_k^{r+4} + \cdots + v_k^{r+p-2k}$, the above expression is equivalent modulo q to

$$v_{k-1}^{r+1} + v_{k-1}^{r+3} + \cdots + v_{k-1}^{r+p-2k-1},$$

as desired.

Taking $k = 1$ gives

$$v_1^r + v_1^{r+2} + \cdots + v_1^{r+p-2} = v_0^{r+1} + v_0^{r+3} + \cdots + v_0^{r+p-3} = \left(\frac{1}{2}p - 1\right)v_0 \pmod{q}.$$

So, for any r , we have $v_1^r - v_1^{r+p} = (v_1^r + v_1^{r+2} + \cdots + v_1^{r+p-2}) - (v_1^{r+2} + v_1^{r+4} + \cdots + v_1^{r+p}) = 0 \pmod q$, ie $v_1^r = v_1^{r+p} \pmod q$. But since q is relatively prime to p , there exists an integer α such that $\alpha p = 1 \pmod q$, so we get $v_1^r = v_1^{r+\alpha p} = v_1^{r+1} \pmod q$ for all r . Therefore we deduce that $v_1^1 = \cdots = v_1^q \pmod q$. It now follows by the first condition and induction on k , that for a fixed k , the integers $v_k^1, \dots, v_k^q$ are congruent modulo q .

We next claim that for all k and r , $v_k^r \equiv (\frac{1}{2}p - k)v_{\frac{p}{2}-1}^q \pmod q$. We will prove this via induction and appealing to the fact that for all k , $v_k^1, \dots, v_k^q$ are equivalent mod q . The claim is clearly true when $k = \frac{1}{2}p - 1$ and $k = \frac{1}{2}p - 2$. Assume inductively that the claim is true for all $j \geq k + 1$. Then

$$\begin{aligned} v_k^r &\equiv v_{k+1}^{r-1} + v_{k+1}^{r+1} - v_{k+2}^r \equiv 2v_{k+1}^r - v_{k+2}^r \equiv 2(\tfrac{1}{2}p - k - 1)v_{\frac{p}{2}-1}^q - (\tfrac{1}{2}p - k - 2)v_{\frac{p}{2}-1}^q \\ &\equiv (\tfrac{1}{2}p - k)v_{\frac{p}{2}-1}^q \pmod q. \end{aligned}$$

In particular, when $k = 0$, we have that $v_0 \equiv \frac{1}{2}pv_{\frac{p}{2}-1}^q \pmod q$. Thus, letting v be the last row of qG^{-1} , we have

$$(qG^{-1})_{\frac{pq-2q+2}{2}, 1} \equiv \frac{1}{2}p \cdot (qG^{-1})_{\frac{pq-2q+2}{2}, \frac{pq-2q+2}{2}} \pmod q.$$

Since G^{-1} is symmetric, we have that

$$(qG^{-1})_{1, \frac{pq-2q+2}{2}} \equiv \frac{1}{2}p \cdot (qG^{-1})_{\frac{pq-2q+2}{2}, \frac{pq-2q+2}{2}} \pmod q.$$

By Lemma 5.1, the $(1, \frac{1}{2}(pq - 2q + 2))$ -entry of G^{-1} is $1/q$. Therefore,

$$(qG^{-1})_{\frac{pq-2q+2}{2}, 1} = (qG^{-1})_{1, \frac{pq-2q+2}{2}} = 1$$

and so $\frac{1}{2}p \cdot (qG^{-1})_{\frac{pq-2q+2}{2}, \frac{pq-2q+2}{2}} \equiv 1 \pmod q$. Thus we now see that $(qG^{-1})_{\frac{pq-2q+2}{2}, \frac{pq-2q+2}{2}}$ is congruent to the modular inverse of $\frac{1}{2}p \pmod q$, as desired. $\square$

To calculate $\lambda_{\Sigma(2,p,q)}$ explicitly, we need to understand the domain on which the pairing is defined, namely $H_1(\Sigma_2(S^3, T_{p,q}); \mathbb{Z})$. It turns out that this group is always cyclic, so $\lambda_{\Sigma(2,p,q)}$ is determined by a single rational number.

Lemma 5.3 $H_1(\Sigma_2(S^3, T_{p,q}); \mathbb{Z}) \cong \mathbb{Z}/q\mathbb{Z}$.

Proof The Alexander module $A_{T_{p,q}}$ of $T_{p,q}$ is given by $\mathbb{Z}[t, t^{-1}]/(\Delta(t))$, where $\Delta(t)$ denotes the Alexander polynomial of $T_{p,q}$; this can be shown easily using Fox calculus. It is known that taking its quotient by the ideal generated by $t + 1$ gives the first homology group H_1 of $\Sigma_2(S^3, T_{p,q})$; see [45, Chapter 4] for a general formula for H_1 of branched cyclic covers over knots. Therefore we get

$$H_1(\Sigma_2(S^3, T_{p,q}); \mathbb{Z}) \simeq A_{T_{p,q}}/(t+1) \simeq \mathbb{Z}[t, t^{-1}]/(t+1, \Delta_{T_{p,q}}(t)) \simeq \mathbb{Z}/\Delta_{T_{p,q}}(-1)\mathbb{Z} \simeq \mathbb{Z}/\det(T_{p,q})\mathbb{Z}.$$

Since $\det(T_{p,q}) = q$, we deduce that $H_1(\Sigma_2(S^3, T_{p,q}); \mathbb{Z}) \simeq \mathbb{Z}/q\mathbb{Z}$. $\square$

Now we are ready to compute $\lambda_{\Sigma(2,p,q)}$.

Proposition 5.4 For any relatively prime positive integers p and q with p even, we have

$$\lambda_{\Sigma(2,p,q)}(x, x) \equiv -\frac{p}{2q} \pmod{1}$$

for some generator x of $H_1(\Sigma(2, p, q); \mathbb{Z})$.

Proof As above, let X denote the double cover of B^4 branched over the surface F after it is pushed into B^4 . Then $\partial X = \Sigma(2, p, q)$ and X is simply connected. Consequently, we have the long exact sequence

$$0 \rightarrow H_2(X) \xrightarrow{G} H_2(X, \Sigma(2, p, q)) \xrightarrow{f} H_1(\Sigma(2, p, q)) \rightarrow 0.$$

Given $y \in H_1(\Sigma(2, p, q))$, to compute $\lambda_{\Sigma(2,p,q)}(y, y)$ using G , we must first lift y to $H_2(X, \Sigma(2, p, q))$ via the map f . Note that we can view $H_1(\Sigma(2, p, q))$ as the cokernel of G and f as the quotient map.

Since the $(1, \frac{1}{2}(pq - 2q + 2))$ -entry of G^{-1} is $1/q$ by [Lemma 5.1](#), the element

$$(0, \dots, 0, 1) \in H_2(X, \Sigma(2, p, q))$$

is a lift of $1 \in H_1(\Sigma(2, p, q)) \cong \mathbb{Z}/q\mathbb{Z}$. To see this, let k be any integer satisfying $f(0, \dots, 0, k) = 0$. Then we have $(0, \dots, 0, k)^T - Gv = 0$ for some vector $v = (v_1, \dots, v_{\frac{pq-2q+2}{2}})^T$, where all v_i are integers. But then $v = G^{-1}(0, \dots, 0, k)^T$, so by comparing the first entries of both sides, we get $v_1 = k/q$. Since $v_1 \in \mathbb{Z}$, k must be a multiple of q . Thus we see that $f(0, \dots, 0, 1)$ generates $H_1(\Sigma(2, p, q))$, as desired.

Now, recall that the linking form $\lambda_{\Sigma(2,p,q)}$ is presented by $-G^{-1} \pmod{1}$. Thus the value of

$$-\lambda_{\Sigma(2,p,q)}(1, 1) = (0, \dots, 0, 1)G^{-1}(0, \dots, 0, 1)^T$$

is given by the $(\frac{1}{2}(pq - 2q + 2), \frac{1}{2}(pq - 2q + 2))$ -entry of G^{-1} , which by [Lemma 5.2](#) is given by m/q where m is some integer satisfying $\frac{1}{2}mp \equiv 1 \pmod{q}$. So we get

$$\lambda_{\Sigma(2,p,q)}\left(\frac{p}{2}, \frac{p}{2}\right) = -\frac{mp^2}{4q} \equiv -\frac{p}{2q} \pmod{1}.$$

Since q is relatively prime to p , $\frac{1}{2}p \in H_1(\Sigma(2, p, q))$ is also a generator of $H_1(\Sigma(2, p, q))$. Therefore taking $x = \frac{1}{2}p$ proves the lemma. $\square$

Remark 5.5 Feller and Golla [\[13\]](#) proved that the linking form of the Brieskorn sphere $\Sigma(2, 2p, 2pk \pm 1)$ represents $\mp p/(2pk + 1)$ as a square. This is coherent with our general computation in [Proposition 5.4](#).

5.2 Number theory interlude

Before we can prove [Theorem 1.13](#), we need some number theoretic results.

Lemma 5.6 Suppose that p is even and $\frac{1}{2}p$ is not a perfect square. There exists an integer r , relatively prime to p , such that for any prime s satisfying $s \equiv r \pmod{2p}$, the integers $\frac{1}{2}p$ and $-\frac{1}{2}p$ are not quadratic residues mod s .

To prove this, we need the following lemma.

Lemma 5.7 [8, Lemma 1.14] *Let n be an integer such that $n \equiv 0, 1 \pmod{4}$. Then there exists a unique group homomorphism*

$$\chi_n: (\mathbb{Z}/n\mathbb{Z})^\times \rightarrow \{\pm 1\}$$

such that $(n/s) = \chi_n(s)$ for all odd primes s not dividing n . Here, (n/s) denotes the Legendre symbol. Furthermore, we have

$$\chi_n(-1) = \begin{cases} 1 & \text{if } n > 0, \\ -1 & \text{if } n < 0. \end{cases}$$

Proof of Lemma 5.6 Since $\frac{1}{2}p$ is not a perfect square, $2p$ is a multiple of 4 which is not a perfect square. Using Lemma 5.7, there is a unique nontrivial group homomorphism $\chi_{2p}: (\mathbb{Z}/2p\mathbb{Z})^\times \rightarrow \{\pm 1\}$ such that $(2p/s) = \chi_{2p}(s)$ for all odd primes s not dividing p . It is a well-known fact (see, for example, [26, Theorem 3 of Section 5.2]) that for any nonsquare integer n , there exist infinitely many primes for which n is a quadratic nonresidue; since $2p$ is not a perfect square, we see that χ_{2p} cannot be a trivial map. Hence the size of $\ker(\chi_{2p})$ is exactly half of its domain $(\mathbb{Z}/2p\mathbb{Z})^\times$.

We claim that there exists an element r not contained in $\ker(\chi_{2p})$ such that $r \equiv 1 \pmod{4}$. Suppose not. Then for each $s \notin \ker(\chi_{2p})$, $s \equiv 3 \pmod{4}$. Since half of the elements in $(\mathbb{Z}/2p\mathbb{Z})^\times$ are of the form $4n+3$ and the other half are of the form $4n+1$, we have that $\ker \chi_{2p} = \{s \in (\mathbb{Z}/2p\mathbb{Z})^\times \mid s \equiv 1 \pmod{4}\}$. However, we have $\chi_{2p}(-1) = 1$ since $2p > 0$, while $-1 \equiv 3 \pmod{4}$. Hence, we have reached a contradiction and there must exist an element $r \notin \ker(\chi_{2p})$ such that $r \equiv 1 \pmod{4}$.

Now, for any prime s congruent to $r \pmod{2p}$, we have

$$-1 = \chi_{2p}(r) = \left(\frac{2p}{s}\right) = \left(\frac{2}{s}\right)^2 \left(\frac{p/2}{s}\right) = \left(\frac{p/2}{s}\right).$$

Moreover, since $s \equiv 1 \pmod{4}$, we have that $(-1/s) = 1$ and, consequently,

$$\left(\frac{-p/2}{s}\right) = \left(\frac{-1}{s}\right) \left(\frac{p/2}{s}\right) = -1.$$

Therefore both $\frac{1}{2}p$ and $-\frac{1}{2}p$ are quadratic nonresidues modulo s . □

5.3 Proof of Theorem 1.13

Recall the following theorem of Murakami and Yasuhara, which is a direct corollary of their lower bound on γ_4^{top} from Lemma 2.7.

Theorem 5.8 [39, Corollary 2.7] *If a knot K in S^3 bounds a locally flat Möbius band in B^4 , then the linking form λ on $H_1(\Sigma_2(S^3, K); \mathbb{Z})$ splits into a direct sum $(G_1, \lambda_1) \oplus (G_2, \lambda_2)$, where λ_1 is represented by the (1×1) -matrix $(\pm 1/|G_1|)$ and λ_2 is metabolic.*

To state the next lemma, we introduce the following definition.

Definition 5.9 Let n be a nonzero integer and let p be a prime factor of n . Let r be the maximal positive integer such that p^r divides n . When r is odd, we say that p is an *odd-power prime factor* of n .

Now using [Theorem 5.8](#), we can prove the following lemma.

Lemma 5.10 Let K be a knot in S^3 satisfying $H_1(\Sigma_2(S^3, K); \mathbb{Z}) \cong \mathbb{Z}/n\mathbb{Z}$ for some positive integer n . Write the linking form λ of $\Sigma_2(S^3, K)$ as $\lambda(x, x) = (m/n)x^2 \pmod{1}$ for some integer m . Suppose that there exists an odd-power prime factor s of n such that m and $-m$ are not quadratic residues mod s . Then K does not bound a locally flat Möbius band in B^4 .

Proof Suppose that K bounds a Möbius band. Then we have a splitting

$$(H_1(\Sigma_2(S^3, K); \mathbb{Z}), \lambda) = (G_1, \lambda_1) \oplus (G_2, \lambda_2),$$

where λ_1 is represented by a (1×1) -matrix $(\pm 1/|G_1|)$ and λ_2 is metabolic. The order of G_2 is a perfect square, so if we write $G_1 = \mathbb{Z}/k\mathbb{Z}$ and $G_2 = \mathbb{Z}/\ell\mathbb{Z}$, then ℓ is a perfect square and s divides k .

Now let 1 be a generator of $H_1(\Sigma_2(S^3, K); \mathbb{Z})$. Then G_1 is generated by $\ell = \ell \cdot 1$. By our assumption on λ_1 , there exists an integer r , relatively prime to k , such that $\lambda_1(r\ell, r\ell) = \pm 1/k$. But since

$$\lambda_1(r\ell, r\ell) = \lambda(r\ell, r\ell) = \frac{m\ell r^2}{k}$$

and ℓ is a perfect square, either m or $-m$ must be a quadratic residue mod k . Since s divides k , this implies that either m or $-m$ is a quadratic residue mod s , a contradiction. $\square$

Example 5.11 Consider the torus knots $T_{4,q}$, where q is odd. By [Proposition 5.4](#), the linking form on $\Sigma_2(S^3, T_{4,q})$ is given by $\lambda(x, x) = -2/q \pmod{1}$ for some generator x of $H_1(\Sigma_2(S^3, T_{4,q}); \mathbb{Z})$. It is well known that ± 2 are both quadratic nonresidues modulo a if and only if $a \equiv 5 \pmod{8}$. Thus, in this case, the only possible value of r in the statement of [Lemma 5.6](#) is $r = 5$. Now by [Lemma 5.10](#), if q has an odd-power prime factor $r \equiv 5 \pmod{8}$, then $T_{4,q}$ does not bound a locally flat Möbius band.

More generally, we can prove the following

Proposition 5.12 Let $p, q > 0$ be relatively prime integers, where p is even. Suppose that $\frac{1}{2}p$ is not a perfect square. Then there exists an integer r , relatively prime to p , such that $T_{p,q}$ does not bound a locally flat Möbius band in B^4 whenever q admits an odd-power prime divisor congruent to $r \pmod{2p}$.

Proof By [Lemma 5.6](#), there exists an integer r , relatively prime to p , such that for any prime s congruent to $r \pmod{2p}$, $\frac{1}{2}p$ and $-\frac{1}{2}p$ are quadratic nonresidues mod s . The lemma then follows directly from [Lemma 5.10](#). $\square$

For the proof of [Theorem 1.13](#), we will use the following estimate.

Lemma 5.13 *Let P be a set consisting of primes, such that the infinite sum $\sum_{s \in P} 1/s$ diverges to ∞ . Given a positive integer N , denote the set of positive integers n less than N whose odd-exponent prime factors are not contained in P by $S_{P,N}$, and the set of positive integers n less than N which are relatively prime to p as $T_{p,N}$. Then*

$$\lim_{N \rightarrow \infty} \frac{|S_{P,N} \cap T_{p,N}|}{|T_{p,N}|} = 0.$$

Proof Without loss of generality, we can assume that P does not contain any prime factors of p . Then, by the inclusion-exclusion principle, the given limit can be computed as the infinite product

$$\prod_{s \in P} \left(1 - \frac{1}{s} + \frac{1}{s^2} - \frac{1}{s^3} + \cdots\right) = \prod_{s \in P} \frac{s}{s+1}.$$

Taking logarithms gives

$$\log \left(\prod_{s \in P} \frac{s}{s+1} \right) = - \sum_{s \in P} (\log(s+1) - \log(s)) \leq - \sum_{s \in P} \frac{1}{s} = -\infty.$$

Here, we used the fact that $\log(s+1) - \log(s) \geq 1/s$ and $\sum_{s \in P} 1/s = \infty$. Therefore the limit given in the statement of the lemma converges to zero. $\square$

Recall the following estimate on the sum of reciprocals of primes in arithmetic progressions, which is a well-known corollary of the prime number theorem.

Lemma 5.14 *For any choice of positive integers N , k , and ℓ , such that ℓ is relatively prime to k , there exists a constant $C_{k,\ell}$ which depends only on k and ℓ such that*

$$\sum_{\substack{s \leq N, s \equiv \ell \pmod{k}, \\ s \text{ is prime}}} \frac{1}{s} = \frac{\log \log N}{\phi(k)} + C_{k,\ell} + O\left(\frac{1}{\log N}\right).$$

We are finally ready to prove the desired asymptotics.

Proof of Theorem 1.13 By Proposition 5.12, there exists some r , relatively prime to p , such that $T_{p,q}$ does not bound a locally flat Möbius band in B^4 whenever q admits an odd-power prime divisor congruent to $r \pmod{2p}$. Moreover, by Lemma 5.14, the sum of reciprocals of all primes congruent to $r \pmod{2p}$ diverges to ∞ . The theorem then follows from Lemma 5.13. $\square$

References

- [1] **S Allen**, *Nonorientable surfaces bounded by knots: a geography problem*, New York J. Math. 29 (2023) 1038–1059 [MR](#) [Zbl](#)
- [2] **W Ballinger**, *Concordance invariants from the $E(-1)$ spectral sequence on Khovanov homology*, preprint (2020) [arXiv 2004.10807](#)

- [3] **J Batson**, *Nonorientable slice genus can be arbitrarily large*, Math. Res. Lett. 21 (2014) 423–436 [MR](#) [Zbl](#)
- [4] **J Bodnár, D Celoria, M Golla**, *A note on cobordisms of algebraic knots*, Algebr. Geom. Topol. 17 (2017) 2543–2564 [MR](#) [Zbl](#)
- [5] **M Borodzik, J Hom**, *Involutive Heegaard Floer homology and rational cuspidal curves*, Proc. Lond. Math. Soc. 118 (2019) 441–472 [MR](#) [Zbl](#)
- [6] **M Borodzik, C Livingston**, *Heegaard Floer homology and rational cuspidal curves*, Forum Math. Sigma 2 (2014) art. id. e28 [MR](#) [Zbl](#)
- [7] **B E Clark**, *Crosscaps and knots*, Int. J. Math. Math. Sci. 1 (1978) 113–123 [MR](#) [Zbl](#)
- [8] **D A Cox**, *Primes of the form $x^2 + ny^2$: Fermat, class field theory and complex multiplication*, Wiley, New York (1989) [MR](#) [Zbl](#)
- [9] **A Daemi, C Scaduto**, *Chern–Simons functional, singular instantons, and the four-dimensional clasp number*, J. Eur. Math. Soc. 26 (2024) 2127–2190 [MR](#) [Zbl](#)
- [10] **I Dai, C Manolescu**, *Involutive Heegaard Floer homology and plumbed three-manifolds*, J. Inst. Math. Jussieu 18 (2019) 1115–1155 [MR](#) [Zbl](#)
- [11] **S K Donaldson**, *The orientation of Yang–Mills moduli spaces and 4-manifold topology*, J. Differential Geom. 26 (1987) 397–428 [MR](#) [Zbl](#)
- [12] **H Fan**, *Unoriented cobordism maps on link Floer homology*, PhD thesis, University of California, Los Angeles (2019) Available at <https://escholarship.org/uc/item/9823t959>
- [13] **P Feller, M Golla**, *Non-orientable slice surfaces and inscribed rectangles*, Ann. Sc. Norm. Super. Pisa Cl. Sci. 24 (2023) 1463–1485 [MR](#) [Zbl](#)
- [14] **P Feller, D Krcatovich**, *On cobordisms between knots, braid index, and the epsilon-invariant*, Math. Ann. 369 (2017) 301–329 [MR](#) [Zbl](#)
- [15] **P Feller, J Park, A Ray**, *On the epsilon invariant and satellite knots*, Math. Z. 292 (2019) 1431–1452 [MR](#) [Zbl](#)
- [16] **K A Frøyshov**, *Equivariant aspects of Yang–Mills Floer theory*, Topology 41 (2002) 525–552 [MR](#) [Zbl](#)
- [17] **N Ghanbarian**, *The non-orientable 4-genus for knots with 10 crossings*, J. Knot Theory Ramifications 31 (2022) art. id. 2250034 [MR](#) [Zbl](#)
- [18] **P M Gilmer, C Livingston**, *The nonorientable 4-genus of knots*, J. Lond. Math. Soc. 84 (2011) 559–577 [MR](#) [Zbl](#)
- [19] **M Golla, M Marengon**, *Correction terms and the nonorientable slice genus*, Michigan Math. J. 67 (2018) 59–82 [MR](#) [Zbl](#)
- [20] **S Gong, M Marengon**, *Nonorientable link cobordisms and torsion order in Floer homologies*, Algebr. Geom. Topol. 23 (2023) 2627–2672 [MR](#) [Zbl](#)
- [21] **C M Gordon, R A Litherland**, *On the signature of a link*, Invent. Math. 47 (1978) 53–69 [MR](#) [Zbl](#)
- [22] **C M Gordon, R A Litherland, K Murasugi**, *Signatures of covering links*, Canadian J. Math. 33 (1981) 381–394 [MR](#) [Zbl](#)
- [23] **M Hedden**, *On knot Floer homology and cabling, II*, Int. Math. Res. Not. 2009 (2009) 2248–2274 [MR](#) [Zbl](#)
- [24] **K Hendricks, C Manolescu**, *Involutive Heegaard Floer homology*, Duke Math. J. 166 (2017) 1211–1299 [MR](#) [Zbl](#)

- [25] **M Hogancamp, C Livingston**, *An involutive epsilon knot invariant*, preprint (2017) [arXiv 1710.08360](#)
- [26] **K Ireland, M Rosen**, *A classical introduction to modern number theory*, 2nd edition, Grad. Texts in Math. 84, Springer (1990) [MR](#) [Zbl](#)
- [27] **S Jabuka, T Kelly**, *The nonorientable 4-genus for knots with 8 or 9 crossings*, *Algebr. Geom. Topol.* 18 (2018) 1823–1856 [MR](#) [Zbl](#)
- [28] **S Jabuka, C A Van Cott**, *Comparing nonorientable three genus and nonorientable four genus of torus knots*, *J. Knot Theory Ramifications* 29 (2020) art. id. 2050013 [MR](#) [Zbl](#)
- [29] **S Jabuka, C A Van Cott**, *On a nonorientable analogue of the Milnor conjecture*, *Algebr. Geom. Topol.* 21 (2021) 2571–2625 [MR](#) [Zbl](#)
- [30] **A Juhász**, *Cobordisms of sutured manifolds and the functoriality of link Floer homology*, *Adv. Math.* 299 (2016) 940–1038 [MR](#) [Zbl](#)
- [31] **A Juhász, D Thurston, I Zemke**, *Naturality and mapping class groups in Heegaard Floer homology*, *Mem. Amer. Math. Soc.* 1338, Amer. Math. Soc., Providence, RI (2021) [MR](#) [Zbl](#)
- [32] **A Juhász, I Zemke**, *New Heegaard Floer slice genus and clasp number bounds*, preprint (2020) [arXiv 2007.07106](#)
- [33] **P B Kronheimer, T S Mrowka**, *Gauge theory for embedded surfaces, I*, *Topology* 32 (1993) 773–826 [MR](#) [Zbl](#)
- [34] **J Levine**, *Polynomial invariants of knots of codimension two*, *Ann. of Math.* 84 (1966) 537–554 [MR](#) [Zbl](#)
- [35] **A Lobb**, *A counterexample to Batson’s conjecture*, *Math. Res. Lett.* 26 (2019) 1789 [MR](#) [Zbl](#)
- [36] **V Longo**, *An infinite family of counterexamples to Batson’s conjecture*, preprint (2020) [arXiv 2011.00122](#)
- [37] **C Manolescu, B Owens**, *A concordance invariant from the Floer homology of double branched covers*, *Int. Math. Res. Not.* 2007 (2007) art. id. rnm077 [MR](#) [Zbl](#)
- [38] **L Moser**, *Elementary surgery along a torus knot*, *Pacific J. Math.* 38 (1971) 737–745 [MR](#) [Zbl](#)
- [39] **H Murakami, A Yasuhara**, *Four-genus and four-dimensional clasp number of a knot*, *Proc. Amer. Math. Soc.* 128 (2000) 3693–3699 [MR](#) [Zbl](#)
- [40] **Y Ni, Z Wu**, *Cosmetic surgeries on knots in S^3* , *J. Reine Angew. Math.* 706 (2015) 1–17 [MR](#) [Zbl](#)
- [41] **P S Ozsváth, A I Stipsicz, Z Szabó**, *Unoriented knot Floer homology and the unoriented four-ball genus*, *Int. Math. Res. Not.* 2017 (2017) 5137–5181 [MR](#) [Zbl](#)
- [42] **P Ozsváth, Z Szabó**, *Absolutely graded Floer homologies and intersection forms for four-manifolds with boundary*, *Adv. Math.* 173 (2003) 179–261 [MR](#) [Zbl](#)
- [43] **P Ozsváth, Z Szabó**, *On knot Floer homology and lens space surgeries*, *Topology* 44 (2005) 1281–1300 [MR](#) [Zbl](#)
- [44] **J Rasmussen**, *Lens space surgeries and a conjecture of Goda and Teragaito*, *Geom. Topol.* 8 (2004) 1013–1031 [MR](#) [Zbl](#)
- [45] **W H Stevens**, *On the homology of branched cyclic covers of knots*, PhD thesis, Louisiana State University and Agricultural & Mechanical College (1996) Available at <https://www.proquest.com/docview/304312454>
- [46] **N Tairi**, *The smooth nonorientable 4-genus of an infinite family of torus knots*, master’s thesis, Universität Bern (2020)

- [47] **M Teragaito**, *Crosscap numbers of torus knots*, Topology Appl. 138 (2004) 219–238 [MR](#) [Zbl](#)
- [48] **O Y Viro**, *Positioning in codimension 2, and the boundary*, Uspehi Mat. Nauk 30 (1975) 231–232 [MR](#) [Zbl](#)
In Russian.
- [49] **C T C Wall**, *Singular points of plane curves*, Lond. Math. Soc. Stud. Texts 63, Cambridge Univ. Press (2004) [MR](#) [Zbl](#)
- [50] **S Wang**, *Semigroups of L -space knots and nonalgebraic iterated torus knots*, Math. Res. Lett. 25 (2018) 335–346 [MR](#) [Zbl](#)
- [51] **A Yasuhara**, *Connecting lemmas and representing homology classes of simply connected 4-manifolds*, Tokyo J. Math. 19 (1996) 245–261 [MR](#) [Zbl](#)
- [52] **I Zemke**, *Connected sums and involutive knot Floer homology*, Proc. Lond. Math. Soc. 119 (2019) 214–265 [MR](#) [Zbl](#)
- [53] **I Zemke**, *Link cobordisms and functoriality in link Floer homology*, J. Topol. 12 (2019) 94–220 [MR](#) [Zbl](#)

FB: Department of Mathematics, Princeton University
Princeton, NJ, United States

SK: Mathematical Institute, University of Oxford
Oxford, United Kingdom

JS: School of Mathematics, Georgia Institute of Technology
Atlanta, GA, United States

PT: Max Planck Institute for Mathematics
Bonn, Germany

fb1673@princeton.edu, sungkyung3838@gmail.com, jsimone7@gatech.edu,
truoel@mpim-bonn.mpg.de

Received: 13 March 2023 Revised: 6 February 2024

Generalised Baumslag–Solitar groups and hierarchically hyperbolic groups

JACK O BUTTON

We look at isometric actions on arbitrary hyperbolic spaces of generalised Baumslag–Solitar groups of any rank (the rank of the free abelian vertex and edge subgroups). It is known that being a hierarchically hyperbolic group is not a quasiisometric invariant. We show that virtually being a hierarchically hyperbolic group is not invariant under quasiisometry either, and nor is property (QT).

20F65; 20F67

1 Introduction

Gromov’s notion of a word hyperbolic group encapsulates what it means for a finitely presented group to be negatively curved. In particular this property is not just a commensurability invariant (that is, if H is a finite index subgroup of G , for which we write $H \leq_f G$, then H has the property if and only if G has the property) but it is also invariant under quasiisometry.

However trying to come up with an equivalent notion of nonpositive curvature for finitely presented groups which is also invariant under quasiisometry seems less clear. One definition that is invariant is that of having (at most) quadratic Dehn function but this contains groups that would not naturally be thought of as being nonpositively curved. To give but one example, de Cornulier and Tessera [13] showed that for each $n \geq 2$ there exists a metabelian group with quadratic Dehn function but which contains the Baumslag–Solitar group $BS(1, n)$, whereas one might hope that any finitely generated soluble subgroup of a nonpositively curved group is virtually abelian.

Even if invariance under quasiisometry fails, one could at least hope for a definition of nonpositive curvature which is invariant under commensurability. However for the property of being $CAT(0)$ (a group which acts geometrically on a $CAT(0)$ space), which is known not to be invariant under quasiisometry, it is currently open whether H being a $CAT(0)$ group and $H \leq_f G$ implies that G is a $CAT(0)$ group. It is also still open whether all hyperbolic groups are $CAT(0)$.

A more recent property that also aims to encapsulate nonpositive curvature is that of being a hierarchically hyperbolic group (a HHG for short). This does include all hyperbolic groups but it was recently shown by Petyt and Spriano [34] that this is not a commensurability invariant as there are groups which are not HHGs but which have a finite index subgroup equal to $\mathbb{Z}^2$, which is. Note that neither $CAT(0)$ groups or HHGs contain each other.

However there is an easy way to turn a group theoretic property $\mathcal{P}$ into a commensurability invariant provided it is preserved by finite index subgroups (which many properties are, including being CAT(0) and an HHG). We merely alter it to being virtually $\mathcal{P}$, that is it has some finite index subgroup with $\mathcal{P}$. We can thus ask instead: for finitely generated groups, is being virtually $\mathcal{P}$ a quasiisometric invariant? For instance this question has been considered for the properties of being virtually free, virtually cyclic, virtually abelian, virtually nilpotent, virtually polycyclic and virtually soluble. None of these would be commensurability invariants if virtually were removed, but the first four are known to be quasiisometry invariants by various deep results. The fifth is unknown and the sixth is false for finitely generated groups but unknown for finitely presented groups (see de la Harpe [23, IV.B.50]).

Therefore it is natural to ask in this context whether virtually being an HHG is a quasiisometry invariant. Note that virtually being a CAT(0) group is not. This is demonstrated by the well known example of a group G which is a central extension of $\mathbb{Z}$ by a closed surface group S_g for $g \geq 2$ which does not virtually split and so is not a CAT(0) group (see Bridson and Haefliger [9, II.7.26]) but which is quasiisometric to $\mathbb{Z} \times S_g$. This group G has all of its finite index subgroups of the same form, so it is not virtually a CAT(0) group either. We also note here that $\mathbb{Z} \times S_g$ acts geometrically not just on a CAT(0) space but on a CAT(0) cube complex too, thus the property $\mathcal{P}$ of acting geometrically on a CAT(0) cube complex is not a quasiisometric invariant and nor is being virtually $\mathcal{P}$ (note that here virtually $\mathcal{P}$ is not equal to $\mathcal{P}$ as shown by Hagen [21]).

However the group G , and more generally central extensions of $\mathbb{Z}$ by any nonelementary hyperbolic group, was recently shown to be a HHG by Hagen, Russell, Sisto and Spriano [22, Corollary 4.3], so will not provide a counterexample to virtually HHGs being invariant under quasiisometry. Nor will the examples which are not HHGs in [34, Theorem 4.4] work because these are all virtually $\mathbb{Z}^k$ for some k .

In this paper we show in Corollary 5.4 that there exist finitely presented groups with no finite index subgroup being a HHG but which are quasiisometric to an HHG. Thus even adding virtually to the property of being an HHG will not make it a quasiisometric invariant. The particular group used in Corollary 5.4 is the Leary–Minasyan group first appearing in [28]. A Leary–Minasyan group denotes any group formed by starting with $A = \mathbb{Z}^n$ for $n \geq 1$ which is regarded as sitting naturally in $\mathbb{Q}^n$ and then taking a matrix $M: \mathbb{Q}^n \rightarrow \mathbb{Q}^n$ which is in $\mathrm{GL}(n, \mathbb{Q})$ and a finite index subgroup B of $A \cap M^{-1}(A)$. The group is then defined to be the HNN extension with base A and associated subgroups B and $M(B)$ using the matrix M . We refer to THE Leary–Minasyan group L as the case where $n = 2$,

$$M = \begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

and $B = A \cap M^{-1}(A)$ which is a CAT(0) group. It was shown in [28, Theorem 1.1] that a Leary–Minasyan group is biautomatic if and only if it is virtually biautomatic which occurs if and only if the matrix M has finite order. (It is unknown whether there exist virtually biautomatic groups which are not biautomatic.) In particular the group L is CAT(0) but not (virtually) biautomatic: this was the first known example. As for

whether every HHG is biautomatic, this was answered negatively by Hughes and Valiunas [25]. In their introduction, it is asked whether any nonbiautomatic Leary–Minasyan group is an HHG, with the answer expected to be no. In this paper we confirm this in [Corollary 5.2](#) by showing that a Leary–Minasyan group is virtually an HHG if and only if the matrix M has finite order.

It turns out that our arguments work in a somewhat wider class of groups than Leary–Minasyan groups, that of generalised Baumslag–Solitar groups of arbitrary rank. A generalised Baumslag–Solitar group of rank n , or GBS_n group for short, is a finite graph of groups where all vertex and edge groups are isomorphic to $\mathbb{Z}^n$ (sometimes the term generalised Baumslag–Solitar group refers only to the case $n = 1$ but we will use the term in its wider context). A GBS_n group G gives rise to a homomorphism from G to $\text{GL}(n, \mathbb{Q})$, defined up to conjugacy, which we call the modular homomorphism of G . (Strictly speaking this should be the modular homomorphism of the decomposition of G as a GBS_n group but in all but some basic cases this homomorphism is unique, even if the decomposition is not.) In the case of a Leary–Minasyan group the graph of groups has one vertex and one edge with the modular homomorphism sending the stable letter to the matrix M and all of the vertex subgroup $\mathbb{Z}^n$ to the identity.

We begin in [Section 2](#) with some introductory material on how groups act on hyperbolic spaces. In [Section 3](#) we introduce generalised Baumslag–Solitar groups G of arbitrary rank n and define the aforementioned modular homomorphism. We use this to examine the free abelianisation of G , that is the usual abelianisation G/G' but with the torsion of G/G' quotiented out. This gives us a dichotomy between groups with finite and infinite monodromy, where the monodromy is the image of G in $\text{GL}(n, \mathbb{Q})$ under the modular homomorphism. Indeed it is shown in [Theorem 3.4](#) that a GBS_n group with finite monodromy is virtually equal to $\mathbb{Z}^n \times F_r$ for some finite rank free group F_r (and hence in this case G is virtually a HHG), whereas a GBS_n group with infinite monodromy cannot be of this form. (As a consequence, whether a GBS_n group has finite or infinite monodromy depends in all cases only on the group and not the decomposition.)

In [Section 3.3](#) we use this information to examine the possible actions by isometries of a GBS_n group G on an arbitrary hyperbolic space. Whilst a group of the form $\mathbb{Z}^n \times F_r$ will have many different actions on hyperbolic spaces, we show in [Theorem 3.5](#) that if G has infinite monodromy then there is a nontrivial element of the base $\mathbb{Z}^n$ -subgroup which never acts loxodromically in any action of G on a hyperbolic space.

As for particular actions on hyperbolic spaces, the concept of an acylindrical action of a group G on an arbitrary hyperbolic space X was introduced by Bowditch [8]. However if X is bounded then any action by isometries is acylindrical, so we require acylindrical actions of G with unbounded orbits in order to obtain anything useful. This was developed by Osin [33] to obtain the much studied concept of an acylindrically hyperbolic group G , in which G has an acylindrical action on some hyperbolic space that is nonelementary, which is equivalent to saying that the action has unbounded orbits and G is not virtually cyclic. In the theory of HHGs, such a group comes equipped with various hyperbolic spaces, including one on which G acts by isometries called the maximal domain S . It was shown by Behrstock, Hagen and

Sisto [4, Theorem 14.3] that if G is an HHG then the resulting action of G on S is acylindrical. However it is perfectly possible that S is a bounded metric space even if G is an infinite HHG (for instance for groups acting geometrically on the product of two trees).

Now it can be seen using standard results that GBS_n groups are never acylindrically hyperbolic but this does not allow us to conclude that they are not HHGs as well, because it could be that in some potential HHG structure on such a group, the maximal domain S is bounded. However this can be avoided by appealing to recent results by Petyt and Spriano [34] which show that if G is any HHG then there is a finite collection of unbounded domains, invariant under the action of G and which are pairwise orthogonal, such that any unbounded domain in the HHG structure is nested in one of these. If G is an infinite group then this collection will be nonempty. Moreover the action of G on each of these domains can be combined to give an action of G on their product and this action is also acylindrical (as well as containing a loxodromic element). Of course a product of at least two unbounded hyperbolic spaces will not itself be hyperbolic but the definition of an acylindrical action makes sense for an arbitrary metric space. Whilst we do not know of any group theoretic consequences if there exists an acylindrical action with a loxodromic element on a product P of hyperbolic spaces, in Section 4 we look at such actions which also preserve the factors of P , which we call a product acylindrical action. We show in Theorem 4.3 that every element in such an action must either be elliptic or loxodromic, as was shown by Bowditch [8, Lemma 2.2] to be the case for a single hyperbolic space.

This is then used to obtain Theorem 4.4 which states that a GBS_n group with infinite monodromy has no product acylindrical action. In Section 5 we give some background on the properties of HHGs that we will need, as opposed to giving the full definition. We then conclude in Corollary 5.2 that a GBS_n group is virtually an HHG if and only if it has finite monodromy, by virtue of the (non)existence of some product acylindrical action. However having finite monodromy is not a quasiisometric invariant, as the Leary–Minasyan group shows.

Bestvina, Bromberg and Fujiwara [6] introduced the property (QT), which applies to finitely generated groups. Such a group has (QT) if it acts isometrically on a finite product of quasitrees such that the orbit map is a quasiisometric embedding. They show there that many groups do have (QT). We show in Corollary 6.4, once again using the Leary–Minasyan group L , that (QT) is not preserved under quasiisometry. For this we do not need knowledge of possible product acylindrical actions of generalised Baumslag–Solitar groups G . We only need to know which elements will be elliptic in all actions of G on a hyperbolic space.

Acknowledgements

The author would like to acknowledge helpful conversations with Mark Hagen and Sam Hughes, as well as the referee for useful suggestions that improved the exposition (including pointing out the reference [26] by Koubi).

2 Actions on hyperbolic spaces

In this paper we consider groups G acting by isometries on a hyperbolic space (or later a finite product of hyperbolic spaces). Here a hyperbolic space X will always mean that X is a geodesic metric space satisfying any of the equivalent definitions of δ -hyperbolicity. Note that no further conditions such as properness of the space will be assumed. As X is hyperbolic we can look at the action of G by homeomorphisms on the (Gromov) boundary ∂X of X to obtain the limit set $\partial_G X$ which is a subset of ∂X . This subset is G -invariant and we have $\partial_H X \subseteq \partial_G X$ if H is a subgroup of G .

We have the standard classification of the individual elements $g \in G$ as follows:

- Definition 2.1** (i) The element g is *elliptic* under the given action if the subgroup $\langle g \rangle$ has bounded orbits. This happens if and only if $\partial_{\langle g \rangle} X = \emptyset$.
- (ii) The element g is *loxodromic* under the given action if the subgroup $\langle g \rangle$ embeds quasiisometrically in X under the (or an) orbit map, namely there is $c > 0$ (which exists independently of $x \in X$) such that for all $n \in \mathbb{Z}$ we have $d(g^n(x), x) \geq |n|c$. This occurs if and only if $\partial_{\langle g \rangle} X$ consists of exactly 2 points $\{g^\pm\}$ for any $x \in X$ and this is the fixed point set of g on ∂X .
- (iii) The element g is *parabolic* exactly when it is not elliptic or loxodromic, which occurs if and only if $\partial_{\langle g \rangle} X$ consists of exactly 1 point and again this is the fixed point set of g on ∂X .

Note that an element $g \in G$ is elliptic/loxodromic/parabolic if and only if g^n is (for some/all $n \in \mathbb{Z} \setminus \{0\}$) and if and only if some conjugate of g is.

Moving back now to arbitrary groups, if G acts by isometries on an arbitrary hyperbolic space X then we have the Gromov classification dividing possible actions into five very different classes (for these facts and related references, see [1; 15]):

- (1) The action has *bounded orbits*.
- (2) The action is *parabolic* (or horocyclic), meaning that $\partial_G X$ has exactly one point p . In this case the action can never be cobounded.
- (3) The action is *lineal*, meaning that $\partial_G X = \{p, q\}$ has exactly 2 points (which in general can be swapped or fixed pointwise by the action of G). In this case there will exist some loxodromic element in G with limit set $\{p, q\}$.
- (4) The action is *quasiparabolic* (or focal). This says that the limit set has at least 3 points, so is infinite, but there is some point $p \in \partial_G X$ which is globally fixed by G . This implies that G contains a pair of loxodromic elements with limit sets $\{p, q\}$ and $\{p, r\}$ for p, q, r distinct points.
- (5) The action is *general*: the limit set is infinite and we have two loxodromic elements with disjoint limit sets.

We will be interested in the question: given a specific group G and a particular element $g \in G$, can we find an action of G on some hyperbolic space X where g acts loxodromically? If we first consider obstructions, the most obvious is if g has finite order or if G is finitely generated and $g \in G$ is distorted in G . As for sufficient conditions, if we have a homomorphism $\theta: G \rightarrow \mathbb{R}$ with $\theta(g) \neq 0$ then we can realise θ as an action of G on $\mathbb{R}$ by translations in which g acts loxodromically. Indeed this also works for a homomorphism $\theta: G \rightarrow \text{Isom}(\mathbb{R}) \cong \mathbb{R} \rtimes C_2$ if $\theta(g)$ has infinite order, as here the elements swapping the ends of $\mathbb{R}$ will have order 2. However, much more generally, this also works by [1, Proposition 4.9] if we have a homogeneous quasimorphism $q: G \rightarrow \mathbb{R}$ with $q(g) \neq 0$, whereupon X is a quasiline.

We also have a type of converse which we will use later: given an action of a locally compact group G on a hyperbolic space X where there is a point $p \in \partial X$ on the boundary fixed by all of G , there exist *Busemann functions* on G . Here all groups with such an action that we consider are countable and discrete, hence locally compact. In particular, see [10] where it is shown that there exists a homogeneous quasimorphism $q: G \rightarrow \mathbb{R}$, the Busemann quasicharacter, where the quasikernel $\{g \in G \mid q(g) = 0\}$ consists exactly of those elements of G which are not acting loxodromically on X . In particular, although this results in q being trivial in actions of the first or second type, in the third (assuming the two limit points are fixed pointwise) or the fourth types of actions we do obtain a nontrivial homogeneous quasimorphism on G . Moreover [10] also shows that this Busemann quasicharacter is a genuine homomorphism if either G is amenable or X is proper. Indeed the first point follows from the well known fact that the only homogeneous quasimorphisms on an amenable group are the homomorphisms.

3 Generalised rank n Baumslag–Solitar groups

Let G be the fundamental group of a finite graph of groups where all edge and vertex groups are isomorphic to $\mathbb{Z}$. The classical Baumslag–Solitar groups are obtained when the underlying graph has a single vertex and one loop. For an arbitrary finite graph, these groups go back at least to [27] and at some later point they became known as generalised Baumslag–Solitar groups. However one can replace $\mathbb{Z}$ with $\mathbb{Z}^n$ for a fixed integer $n \geq 1$, which are also sometimes known as generalised Baumslag–Solitar groups. Therefore for clarity we introduce the following definition.

Definition 3.1 A *generalised rank n Baumslag–Solitar group* (GBS_n group) G is the fundamental group of a finite graph Γ of groups where all vertex and edge groups are isomorphic to $\mathbb{Z}^n$.

This notation occurs in [3], where also the term $v\text{GBS}$ group is used (with v standing for variable) for when the rank of the various vertex and edge groups $\mathbb{Z}^k$ need not be constant. However $v\text{GBS}$ groups need not enjoy the finite valence property mentioned below and so we will not consider those groups further.

3.1 The modular homomorphism

We can obtain a finite presentation for a GBS_n group G by taking each vertex $v_i \in \Gamma$ and corresponding vertex group V_i and fixing a free basis $a_{i,1}, \dots, a_{i,n}$ for V_i . We take a maximal tree T_0 in the finite graph

Γ defining G and form the amalgamation of the V_i over the finite index edge subgroups. We then add a stable letter t_j to the generators for each of the r (say) edges e_j of $\Gamma \setminus \{T_0\}$ and the corresponding relations where t_j conjugates the inclusion of the edge group E_j at one end of e_j to the inclusion of E_j at the other end. Note this means that $G = N \rtimes F_r$, where N is given by the normal closure of the vertex groups V_i in G and $F_r = \langle t_1, \dots, t_r \rangle$.

For $n = 1$ a modular homomorphism from G to $\mathbb{Q}^\times$ was considered in [29]. This was generalised to arbitrary n in [14], where the homomorphism is now from G to $\mathbb{R}^n \rtimes \mathrm{GL}(n, \mathbb{R})$. We present our own more basic version here for arbitrary n which has the advantage that it works for the finite index subgroups of G as well. It is equivalent to the modular representation in [12, Section 3] but is a more concrete version, making it easier to calculate in specific examples.

As a GBS_n group is formed from a graph of groups, it acts (by automorphisms without inverting edges) coboundedly on a simplicial tree T . The crucial property here is that this tree has finite valence because a subgroup of $\mathbb{Z}^n$ which is itself isomorphic to $\mathbb{Z}^n$ must be a finite index subgroup. Now given any GBS_n group G , we first make arbitrary choices of a base vertex v_0 in the tree T on which G acts, a finite index subgroup $A \cong \mathbb{Z}^n$ of the stabiliser $\mathrm{Stab}_G(v_0) \cong \mathbb{Z}^n$ and an ordered basis $a_1, \dots, a_n$ of A . We refer to A as our *base $\mathbb{Z}^n$ -subgroup* for G . But given any element $g \in G$, we have that A and $g^{-1}Ag$ are commensurable subgroups of G , that is their intersection has finite index in both. This is because $g^{-1}Ag \leq_f \mathrm{Stab}_G(g^{-1}(v_0)) = g^{-1}\mathrm{Stab}(v_0)g$ and the finite valence of the tree T implies that any two vertex stabilisers are commensurable so this holds for $\mathrm{Stab}_G(v_0)$ and $g^{-1}\mathrm{Stab}(v_0)g$, hence also for A and $g^{-1}Ag$.

Our modular homomorphism $\mathcal{M}$ which we now define will be a homomorphism from G to $\mathrm{GL}(n, \mathbb{Q})$ and it is easily checked that changing any of these choices results in a homomorphism that is conjugate in $\mathrm{GL}(n, \mathbb{Q})$ to $\mathcal{M}$. We first note that for any finite index subgroup H of $\mathbb{Z}^n$, there is $m \in \mathbb{N}$ (depending on H) such that $m\mathbb{Z}^n \leq_f H \leq_f \mathbb{Z}^n$ (here using additive notation).

Definition 3.2 Let G be a GBS_n group for $n \geq 1$ with some base subgroup $A \cong \mathbb{Z}^n$. We define the *modular homomorphism* $\mathcal{M}: G \rightarrow \mathrm{GL}(n, \mathbb{Q})$ in the following way. For any $g \in G$, we have just remarked that (taking $\mathbb{Z}^n$ to be our base subgroup A and H to be $A \cap g^{-1}Ag$) there is $m > 0$ such that

$$mA = \langle a_1^m, \dots, a_n^m \rangle \leq_f A \cap g^{-1}Ag \leq_f A.$$

In particular for any $a = a_1^{l_1} \cdots a_n^{l_n} \in A$ with $l_1, \dots, l_n \in \mathbb{Z}$, we have that $a^m \in g^{-1}Ag$ and so the element $ga^m g^{-1}$ of G is actually in A . This means that on taking a to be each of our basis elements $a_1, \dots, a_n$ in turn, we have uniquely defined integer coefficients g_{ij} with

$$ga_j^m g^{-1} = a_1^{g_{1j}} \cdots a_n^{g_{nj}}$$

and our definition of the modular map $\mathcal{M}$ is that it sends g to the matrix whose ij^{th} entry is $g_{ij}/m \in \mathbb{Q}$.

The *monodromy* of a GBS_n group G is the image $\mathcal{M}(G)$.

Note that this definition of $\mathcal{M}(g)$ is independent of the value of m taken as if we replace m with m' for any appropriate $m' > 0$ then $(ga^m g^{-1})^{m'}$ is equal to $(ga^{m'} g^{-1})^m$. Moreover for any $g, h \in G$ we have $\mathcal{M}(g)\mathcal{M}(h) = \mathcal{M}(gh)$ so that $\mathcal{M}$ maps to $\text{GL}(n, \mathbb{Q})$ and is a homomorphism. Clearly the subgroup A is in the kernel of $\mathcal{M}$. We also note here that any element g of G acting elliptically on the tree T is sent to the identity by $\mathcal{M}$ because g lies in some vertex stabiliser $\text{Stab}(v) \cong \mathbb{Z}^n$ and by finite valence of T there will be a finite index subgroup $B \leq_f A$ with $B \leq \text{Stab}(v)$, whereupon $gbg^{-1} = b$ for all $b \in B$. Thus $\mathcal{M}$ factors through the decomposition of G into $N \rtimes F_r$ above and so can also be thought of as a homomorphism from F_r to $\text{GL}(n, \mathbb{Q})$. In particular if $r = 0$ so that the underlying graph Γ is actually a finite tree then $\mathcal{M}$ is the trivial homomorphism. This can also happen for $r > 0$, for instance the HNN extension $\langle a, b, t \mid [a, b], tat^{-1} = a, tbt^{-1} = b \rangle$ when $n = 2$.

Given our modular homomorphism $\mathcal{M}: G \rightarrow \text{GL}(n, \mathbb{Q})$, we can restrict $\mathcal{M}$ to a subgroup H of G . If $H \leq_f G$ then H is also a GBS_n group because the restriction of the action of G on the tree T to H is also cobounded, with edge and vertex subgroups which are finite index subgroups of $\mathbb{Z}^n$, hence are all isomorphic to $\mathbb{Z}^n$ too. This description of H gives rise to its own modular homomorphism $\mathcal{M}_H$ but we can regard this, up to conjugacy, as the restriction of the modular homomorphism $\mathcal{M}$ for G . This is because as G and H are acting on the same tree T , we can first take the same base vertex v_0 in T . Then since $H \leq_f G$, we have that $H \cap \text{Stab}_G(v_0) \leq_f \text{Stab}_G(v_0)$. Thus we can choose our subgroup A to be $H \cap \text{Stab}_G(v_0)$ both when defining $\mathcal{M}$ and $\mathcal{M}_H$ and we also choose the same ordered basis for A in both cases. We are now in the position that the definition of $\mathcal{M}_H$ is exactly the definition of $\mathcal{M}$ but just for elements $h \in H$.

3.2 The free abelianisation

If G is a finitely generated group and G' is its commutator subgroup then G/G' is the abelianisation of G . It is a finitely generated abelian group and so is of the form $\mathbb{Z}^k \oplus T$, where the torsion subgroup T is finite. Moreover every abelian quotient of G factors through G/G' .

Here we will consider the *free abelianisation* $\bar{G}$ where we further quotient out by the torsion in the abelianisation to obtain $\bar{G} = \mathbb{Z}^k$ for some k . This has the corresponding universal property that any homomorphism from G to a torsion free abelian group factors through $\bar{G}$.

If we are given a finite presentation for G with m generators then it is easy to calculate G/G' and $\bar{G}$ by abelianising these relations and considering them as defining a subgroup S of $\mathbb{Z}^m$ so that G/G' is the quotient abelian group $\mathbb{Z}^m/S$. In fact the process is even easier for $\bar{G}$ because of the lack of torsion: we can work over $\mathbb{Q}$ to get that the rank k of $\bar{G}$ is the dimension of the quotient space $\mathbb{Q}^m/R$, where $\mathbb{Q}^m$ is the vector space spanned by the given generators for G and R is the subspace spanned by the relators for G , once these relators have been abelianised and regarded as elements of $\mathbb{Q}^m$. This also says that an element $g \in G$ has infinite order (equivalently is nontrivial) in $\bar{G}$ if on expressing g as a word in the generators and abelianising this word, the corresponding $\mathbb{Q}^m$ -vector is not in the subspace R .

This process works out especially well for a GBS_n group G . As before, we take our base vertex v_0 , finite index subgroup A of $\text{Stab}_G(v_0)$ and basis $a_1, \dots, a_n$ for A . On considering the $\mathbb{Q}$ -vector space W of dimension n spanned by this basis, let us consider how this relates to forming the group G as the fundamental group of a finite graph of $\mathbb{Z}^n$ groups and how it also relates to $\bar{G}$. Each time we introduce a new vertex group V_i and form its amalgamation with the previous vertex groups over the appropriate edge group, we are giving an identification of the $\mathbb{Q}$ -vector space spanned by V_i with our original vector space W . Thus if there are no stable letters then $\bar{G} = \mathbb{Z}^n$. However on taking one of the stable letters t_j with its edge running from the vertex v_j to the vertex $v_{j'}$, this introduces n new relations in G of the form

$$t_j x_1^{l_1} \cdots x_n^{l_n} t_j^{-1} = y_1^{m_1} \cdots y_n^{m_n},$$

where $x_1, \dots, x_n$ is a basis for the vertex group V_j and $y_1, \dots, y_n$ a basis for $V_{j'}$. Thus in $\bar{G}$ we obtain the abelianised relation $l_1 x_1 + \cdots + l_n x_n = m_1 y_1 + \cdots + m_n y_n$, so that the corresponding relator can be expressed using our identifications above as an element of W which is trivial in $\bar{G}$. Thus the span of these rn relators forms a subspace R of $W \cong \mathbb{Q}^n$ and our free abelianisation $\bar{G}$ can be described over $\mathbb{Q}$ as the quotient vector space $\mathbb{Q}^r \oplus (W/R)$, where the first summand comes from the stable letters. Note that we can again work out easily whether an element g of G is nontrivial in $\bar{G}$ by writing g in terms of the generators obtained from the graph of groups decomposition and abelianising. Indeed g will be trivial in $\bar{G}$ if and only if each stable letter appears in g with exponent sum 0 and such that the resulting abelianisation of the word representing g , which will now lie in W , also lies in R . Thus in particular we have from this discussion:

Theorem 3.3 *Suppose that G is a GBS_n group with base $\mathbb{Z}^n$ -subgroup A . Then A embeds in the free abelianisation $\bar{G}$ if and only if $R = \{0\}$. Moreover this happens if and only if the monodromy $\mathcal{M}(G)$ is trivial because the modular homomorphism is defined by what it does on the stable letters.*

Note: from this, we see that the set of elements in A which are trivial (equivalently have finite order) in the free abelianisation of G form a subgroup of A . This is because these are the elements of A which, when considered as elements of $\mathbb{Q}^n$, lie in the subspace R of W .

We can now present a dichotomy in the behaviour of GBS_n groups (which for $n = 1$ is shown in [29, Proposition 2.6] and for general n was stated in [12, Section 3]).

Theorem 3.4 *If G is any GBS_n group with finite monodromy then G is virtually $\mathbb{Z}^n \times F_r$ for some $r \geq 0$.*

Proof First drop down to a finite index subgroup H of G where H has trivial monodromy, for instance we could take H to be the kernel of the modular homomorphism $\mathcal{M}$. As H is the fundamental group of a finite graph of groups where all of the finitely many edge and vertex groups are commensurable, we can intersect them to get a subgroup B of H which can be used as a base $\mathbb{Z}^n$ -subgroup H . Now let us consider the presentation we obtain for H from this graph of groups decomposition. Our finite generating

set consists of generators g_i of the vertex groups along with the stable letters t_j . Now any element $b \in B$ will also lie in any vertex group and so will commute with every element g_i . On taking a stable letter t which is obtained from the edge joining the vertices v_1 and v_2 (possibly the same vertex) with vertex groups V_1, V_2 say and respective edge inclusions $E_1 \leq V_1$ and $E_2 \leq V_2$, we have that $tE_1t^{-1} = E_2$.

But as the monodromy is trivial, there is $M > 0$ (depending on $b \in B$) such that $b^M \in B \cap t^{-1}Bt$ and $tb^Mt^{-1} = b^M$. However B lies in every edge and vertex group, so that $b \in E_1$ and hence $tb^Mt^{-1} \in E_2 \cong \mathbb{Z}^n$. Thus tb^Mt^{-1} must be an element in E_2 such that $(tb^Mt^{-1})^M = b^M \in E_2$. Clearly the element b has this property as $B \leq E_2$ as well. Moreover M^{th} roots are unique in $\mathbb{Z}^n$, thus $tb^Mt^{-1} = b$.

Hence we conclude that B is normal and indeed central in H . As it lies in (and is normal in) every vertex and edge group, we can consider H/B . This group itself admits a graph of groups decomposition with the same underlying finite graph, but with vertex groups V_i/B and edge groups E_j/B . These are all finite groups so H/B is virtually free. Hence we can pull back a finite index free subgroup of H/B to obtain a subgroup $L \leq_f H$ with $L/B \cong F_r$. But as this quotient is free the extension splits, so there is a copy of F_r in L with $L = B \rtimes F_r$. As B is central this is simply a direct product, so that G has the finite index subgroup $L \cong \mathbb{Z}^n \times F_r$. $\square$

3.3 Actions of GBS_n groups on arbitrary hyperbolic spaces

We can now use the above to see how a given GBS_n group G and its finite index subgroups H act on hyperbolic spaces. Certainly we have the action of G on its Bass–Serre tree where all elements of the base $\mathbb{Z}^n$ -subgroup A act elliptically. However the point is that if some element a of A is loxodromic when G acts on a hyperbolic space, the fact that the base subgroup A is commensurated in all of G means that the action must be very restricted.

Theorem 3.5 *Suppose the monodromy of a GBS_n group G with given base $\mathbb{Z}^n$ -subgroup A has infinite order. Then there is a nontrivial element $z \in A$ such that for any isometric action of G on a hyperbolic space, the element z does not act loxodromically.*

Proof Consider any isometric action of G on some hyperbolic space X and suppose that there is some element $a \in A$ which is acting loxodromically. We have the two limit points $p^\pm \in \partial X$ for the action of $\langle a \rangle$ and these are the only points in ∂X fixed by any nontrivial power of a . Now for any $g \in G$, we have that A and $g^{-1}Ag$ are commensurable subgroups. As $a \in A$, we can find $j > 0$ such that a^j is in both of these subgroups. In particular $ga^jg^{-1} \in A$ and it acts loxodromically on X because it is a conjugate in G of the element a^j . But as A is abelian, this means ga^jg^{-1} sends the fixed point set of a^j to itself and so swaps or fixes the two points p^+ and p^- . However the case where p^+ and p^- are swapped cannot occur: if so then the loxodromic element ga^jg^{-1} would have two other fixed points q^+, q^- say on ∂X . But then the square of ga^jg^{-1} , which is also loxodromic, would fix at least four distinct points $p^\pm, q^\pm$ on ∂X , which loxodromic elements cannot do. Thus for every $g \in G$ we have that a^j and ga^jg^{-1} both

have the two fixed points p^+ and p^- . But the latter element actually has fixed points $g(p^+)$ and $g(p^-)$ in the action of G on ∂X , so $\{p^+, p^-\}$ is preserved by every element of G and hence the action must be lineal of type (3). Here we can have elements of G which swap the two points, but if so then we can avoid this by dropping down to a subgroup of index 2.

Therefore let G_0 be the index-1 or index-2 subgroup of G where the action preserves p^+ and p^- pointwise. We can now take the Busemann quasicharacter $q: G_0 \rightarrow \mathbb{R}$ (at p^+ say) which is a homogeneous quasimorphism on G_0 that restricts to a genuine homomorphism $\theta = q|_{G_0 \cap A}$ on $G_0 \cap A$ (which is amenable). Note that θ is nontrivial because $a \in G_0 \cap A$ is acting loxodromically.

We will now show that we can “lift” θ to G_0 , in that there is some homomorphism $\Theta: G_0 \rightarrow \mathbb{R}$ which restricts to θ on $G_0 \cap A$. As G is a GBS $_n$ group with a graph of groups decomposition giving us a presentation for G , we can do the same for G_0 by restricting the action of G on the Bass–Serre tree T to G_0 and then taking a quotient. We already have Θ defined on $G_0 \cap A$, which is a finite index subgroup of the vertex group at the base vertex in the finite graph $\Gamma = G_0 \backslash T$ and we can extend Θ to the group generated by all vertex groups but without the relations from the stable letters. This is because all other defining relations are each given by an isomorphism between finite index subgroups of vertex groups, so we proceed inductively by extending Θ to the image of the edge group in the new vertex group, then we can extend over this vertex group.

However we must also consider the defining relations for G_0 coming from the stable letters. These are all of the form

$$tx_1^{l_1} \cdots x_n^{l_n} t^{-1} = y_1^{m_1} \cdots y_n^{m_n},$$

where t is one of these stable letters. We allow Θ to send the stable letters anywhere but $h_1 := x_1^{l_1} \cdots x_n^{l_n}$ is an element in the vertex group at one end of the edge defining t and $h_2 := y_1^{m_1} \cdots y_n^{m_n}$ is in the vertex group at the other end. We must now show that Θ is still well defined when these relations of the form $th_1 t^{-1} = h_2$ are added. Now h_1 and h_2 both lie in vertex stabilisers so by commensurability there will be $M > 0$ such that h_1^M and h_2^M are both in our base $\mathbb{Z}^n$ -subgroup $G_0 \cap A$. But t, h_1, h_2 all lie in G_0 which is the domain of the homogeneous quasimorphism q , thus we have $q(th_1^M t^{-1}) = q(h_2^M)$. Now homogeneous quasimorphisms are invariant under conjugation (because $|q(yxy^{-1}) - q(x)|$ is bounded independently of x , so we can replace x with x^m and let m tend to infinity), thus this becomes $q(h_1^M) = q(h_2^M)$. But $q(h_i^M) = \theta(h_i^M)$ for $i = 1, 2$ as h_1^M, h_2^M are in the domain of θ and we have $\Theta(h_i) = (1/M)\Theta(h_i^M)$ by definition of Θ above. Also $\Theta(th_1 t^{-1}) = \Theta(h_1)$ as we are mapping to $\mathbb{R}$, so this is also equal to

$$(1/M)\Theta(h_1^M) = (1/M)\theta(h_1^M) = (1/M)\theta(h_2^M) = (1/M)\Theta(h_2^M) = \Theta(h_2)$$

so Θ is indeed well defined.

We can now finish the proof. The subgroup G_0 obtained above has index 1 or 2 in G and there are only finitely many index-2 subgroups in G as it is finitely generated. Let G_2 be the intersection of all of

these index-2 subgroups, which will have finite index in G . As G has infinite monodromy, the same is true for G_2 where we can take our base $\mathbb{Z}^n$ -subgroup to be $G_2 \cap A$. Thus by [Theorem 3.3](#) applied to G_2 with base $\mathbb{Z}^n$ -subgroup $G_2 \cap A$, there is some nonidentity $z \in G_2 \cap A$ which is trivial in the free abelianisation of G_2 . Let us now suppose that G does have an action on some hyperbolic space X in which z is loxodromic, so we can run through the above argument with z equal to a . But then we will obtain a homomorphism Θ from some index-2 subgroup of G to $\mathbb{R}$ with $\Theta(z) \neq 0$ and this subgroup will contain G_2 . Thus we can restrict Θ to G_2 which is a contradiction because z is trivial in the free abelianisation of G_2 . $\square$

Note that this theorem can be used to complement [Theorem 3.4](#) in that if a GBS_n group G has infinite monodromy then it can have no finite index subgroup H which is isomorphic to a direct product of a free group and copies of $\mathbb{Z}$. This is because H would have infinite monodromy as well, thus by [Theorem 3.5](#) there would be a nontrivial element of H which cannot be loxodromic in any action of H on a hyperbolic space. But if H has the above form then we can use actions on trees to make any given element loxodromic.

This at least deals with a potential ambiguity which we have previously glossed over: that G could have different decompositions as a generalised Baumslag–Solitar group. For instance $\mathbb{Z}^{n+1}$ is both a GBS_{n+1} group in a trivial way where the underlying graph is a single vertex and also a GBS_n group $\mathbb{Z}^n \times \mathbb{Z}$ with graph a vertex and an edge, where the stable letter acts trivially by conjugation and where we have the freedom to take any primitive element as the stable letter. Here the monodromy is trivial in all cases, but if we take the fundamental group of the Nil torus bundle

$$\langle t, a, b \mid [a, b] = e, tat^{-1} = ba, tbt^{-1} = b \rangle,$$

which is of the form $\mathbb{Z}^2 \rtimes \mathbb{Z}$, we see that we have a decomposition with base group $\langle a, b \rangle$ and stable letter t , or base group $\langle t, b \rangle$ and stable letter a . The respective modular homomorphisms of these two decompositions are different but both have infinite monodromy.

The above tells us that whether the monodromy is finite or infinite (which is our main concern in this paper) is an invariant only of the group G , not of how it decomposes as a generalised Baumslag–Solitar group. In fact in nearly all cases the modular homomorphism itself and thus the monodromy (up to conjugation in $\text{GL}(n, \mathbb{Q})$) is well defined. We state this as:

Proposition 3.6 *Take any two decompositions of the same group G as a generalised Baumslag–Solitar group. If neither of the two actions on the corresponding Bass–Serre trees have an invariant line or point, the modular homomorphisms corresponding to the two decompositions are the same.*

Proof Here we adapt the result in [\[20, Corollary 6.10\]](#) which achieves this for $n = 1$. Suppose that a group G has two decompositions as a generalised Baumslag–Solitar group with the first of rank n , say. By considering cohomological dimension, the second is also of rank n unless one decomposition is trivial (in which case G is abelian and so the modular homomorphism obtained from any possible decomposition

will be trivial too). Hence we have two actions of G on trees T_1 and T_2 , say, and hence two partitions $\{E_1, H_1\}$ and $\{E_2, H_2\}$ of G into elliptic and hyperbolic elements according to these actions. But if these partitions are the same then the resulting modular homomorphisms are the same. This is because elliptic elements are sent to the identity and the image of a hyperbolic element g is determined by how it conjugates elements in any base $\mathbb{Z}^n$ -subgroup. But under any action of G on a tree (by automorphisms without inverting edges), a finitely generated purely elliptic subgroup will fix some vertex, so we can use the same base $\mathbb{Z}^n$ -subgroup in either decomposition.

Thus we are done if we have a characterisation of the hyperbolic and elliptic elements of G which does not depend on the particular action. In [20] this is achieved (in all but these exceptional cases) for $n = 1$ by showing that the elliptics are the elements of G which are commensurable with all their conjugates. Here we can instead argue: suppose that $g \in G$ lies in $E_2 \setminus E_1$ so that it fixes a vertex v_2 when G acts on T_2 but which is hyperbolic when G acts on T_1 . The action of G on T_2 gives rise to a generalised Baumslag–Solitar decomposition of G where we can take the base group to be $\text{Stab}_G(v_2)$. Thus on taking g equal to a in the proof of Theorem 3.5 where we use the action of G on T_1 with g loxodromic, we conclude by the same argument that G fixes setwise the axis of g . This means that this GBS_n decomposition of G gives rise to an invariant line when G acts on the Bass–Serre tree which is ruled out by hypothesis. (In fact a brief check reveals that in the exceptional case, G can only be equal to $\mathbb{Z}^n$, $\mathbb{Z}^n \rtimes_{\alpha} \mathbb{Z}$ for α some automorphism of $\mathbb{Z}^n$, or $\mathbb{Z}^n *_C \mathbb{Z}^n$, where C has index 2 in both copies of $\mathbb{Z}^n$.) Thus we see that g was not loxodromic and so $E_2 \subseteq E_1$. We can now swap the actions and argue again to conclude that $E_1 = E_2$ and $H_1 = H_2$. $\square$

We finish this section by noting though that even though the modular homomorphism is well defined (away from these exceptional cases), a group might have many decompositions as a generalised Baumslag–Solitar group which are not obviously related. For instance the isomorphism problem is open just amongst GBS_1 groups (see [11] for some recent progress on this question). Moreover [7] exhibits GBS_4 groups with unsolvable conjugacy problem.

4 Acylindrical actions on products of hyperbolic spaces

4.1 Product acylindrical actions

Given a group G acting by isometries on a metric space X , a well known definition is that of G acting *acylindrically*: that is, given any $\epsilon \geq 0$ we have N, R such that if $x, y \in X$ are two points which are at least distance R apart then the set of group elements moving both x and y by at most ϵ has cardinality at most N . This definition is generally used when X is a hyperbolic metric space whereupon it gives rise to the concept of a group being *acylindrically hyperbolic*. This is where there exists an acylindrical action which is of type (5) in the list on page 2257 (in fact actions of types (2) and (4) can never be acylindrical on hyperbolic spaces) and which implies a number of consequences, for instance such a group must be SQ-universal.

Observe that the definition of an acylindrical action makes sense for any group action by isometries on any metric space. However such a concept is not useful in this generality. First of all if X is bounded then any action is trivially acylindrical, so any suitable notion needs to avoid this case. (Indeed even if X is unbounded then an acylindrical action can still have bounded orbits, but not all actions with bounded orbits need be acylindrical.) But even this is problematic because any geometric action on any metric space is uniformly metrically proper, which in turn implies that the action is acylindrical. Therefore any finitely generated group acts acylindrically on its own Cayley graph and consequently there is no chance that an acylindrical action automatically implies any group theoretic consequences, even if we could agree on what a suitable acylindrical action meant in this context.

But one option is to restrict the metric space X to being hyperbolic-like or to have nonpositive curvature in some sense, whilst making it more general than just a hyperbolic space. In this section we look at what happens when we allow X to be a finite product of hyperbolic spaces rather than just one. This will allow us to create obstructions to GBS_n groups having such an action, which we will then compare to the class of hierarchically hyperbolic groups in the next section. Suppose we have a product of r metric spaces $P = X_1 \times \cdots \times X_r$ (where P is equipped with the ℓ_1 product metric) and an isometric action of a group G on P . Note that $\text{Isom}(X_1) \times \cdots \times \text{Isom}(X_r)$ is naturally a subgroup of $\text{Isom}(P)$ using the diagonal action. We say that G acts on P *preserving factors* if the image of this action lies inside $\text{Isom}(X_1) \times \cdots \times \text{Isom}(X_r)$, whereupon we can think of any element $g \in G$ as having an expression $(g_1, \dots, g_r)$ with g_i an isometry of $\text{Isom}(X_i)$. We now introduce our main definition of this section.

Definition 4.1 If a group G acts isometrically on $P = X_1 \times \cdots \times X_r$, where each X_i is a hyperbolic space, then we say that the action is *product acylindrical* if the action on P is acylindrical, preserves the factors of P and such that there is an element $g = (g_1, \dots, g_r)$ in G where some g_i acts as a loxodromic element on the space X_i .

Note that if $r = 1$ then, as opposed to the standard definition of an acylindrically hyperbolic group, we allow actions of type (3) (whereupon our group will be virtually cyclic) with an infinite order element acting loxodromically as we will not need to treat this as a special case. However in common with the standard definition, we rule out any action with bounded orbits.

We make some points about the above definition. First it need not be the case that if G acts acylindrically on a space X and arbitrarily on another space Y then the product action on $X \times Y$ is acylindrical. (One could take any acylindrical action of some group G on X where a point has an infinite stabiliser and then set $Y = \mathbb{R}$ with G acting as the identity on Y .) However if G acts uniformly metrically properly on X and arbitrarily on Y then the product action on $X \times Y$ will also be uniformly metrically proper and hence will be an acylindrical action. Thus if G does have a product acylindrical action then we can say nothing in general about the action on any individual factor. Moreover we do obtain groups which have a product acylindrical action but which are not themselves acylindrically hyperbolic, for instance $F_2 \times F_2$

or Burger–Mozes–Wise groups acting geometrically on a product of two trees. As these last groups can be virtually simple, we also note that we do not have any result for product acylindrical actions of the form: a group with such an action is SQ-universal or virtually cyclic which we do have for unbounded acylindrical actions on a single hyperbolic space.

In Definition 2.1 we saw a division of isometries into a trichotomy of elliptic/loxodromic/parabolic elements when the space is hyperbolic. Note that we still have this trichotomy for a group acting on an arbitrary metric space X : elliptic elements have bounded orbits, loxodromic elements have orbits which quasiisometrically embed in X and everything else is a parabolic element. In general we might not have a nice description of how these isometries act on the boundary of X (indeed there might not even be a suitable boundary of X). However we do have a simple way of classifying an element of a group acting on a finite product of metric spaces preserving factors if we know how this element is behaving on each of the factors.

Lemma 4.2 *Let G act on the product of metric spaces $P = X_1 \times \cdots \times X_r$ by isometries preserving factors and take $g = (g_1, \dots, g_r) \in G$, where g_i acts isometrically on X_i . Then:*

- (i) *g acts elliptically on P if and only if g_i acts elliptically on X_i for each $1 \leq i \leq r$.*
- (ii) *g acts loxodromically on P if and only if there is some i where g_i acts loxodromically on X_i .*
- (iii) *g acts parabolically on P if and only if no g_i acts loxodromically on X_i but some g_i acts parabolically on X_i .*

Proof The first case is straightforward to establish in both directions, just by using the fact that if in the product space $P = X_1 \times \cdots \times X_r$ we have points $x = (x_1, \dots, x_r)$ and $y = (y_1, \dots, y_r)$ then $d_{X_i}(x_i, y_i) \leq d_P(x, y)$. The same is true for the reverse implication of the second case.

Thus let us now suppose without loss of generality that g_1 acts parabolically on X_1 and g_i does not act loxodromically on X_i for any $1 \leq i \leq r$. Then the orbit of any point $x_1 \in X_1$ under $\langle g_1 \rangle$ will be unbounded so the orbit of any $x \in P$ under $\langle g \rangle$ will be too. Thus in order to establish the forward direction of (ii) and the reverse direction of (iii), we just need to show that g is not loxodromic, for which we can use stable translation length. In particular for any point $x = (x_1, \dots, x_r) \in P$, we have $d_i(g_i^m(x_i), x_i)/m$ tending to zero as m tends to infinity because no g_i acts loxodromically. But then by adding and using inequalities we have that $d_P(g^m(x), x)/m$ tends to zero too, thus we cannot have $c > 0$ and $\epsilon \geq 0$ with $d_P(g^m(x), x) \geq cm - \epsilon$ for all $m \in \mathbb{N}$ and so g does not act loxodromically on P . Finally if g acts parabolically on P then we cannot be in cases (i) or (ii) by what we have already shown. $\square$

It was shown in [8, Lemma 2.2] that if G acts acylindrically on a hyperbolic space X then no element can act parabolically. For a general metric space X , we can certainly have acylindrical actions with elements acting parabolically. For instance take any finitely generated group G with a distorted infinite cyclic subgroup $\langle g \rangle$. Then G acts geometrically and hence acylindrically on its own Cayley graph, but the action of g here will be parabolic.

We now show however that we cannot have parabolic elements in a product acylindrical action.

Theorem 4.3 *Suppose that a group G acts acylindrically and preserving factors on the product P of hyperbolic spaces $X_1 \times \cdots \times X_r$. Then no element of G acts as a parabolic element on P .*

Proof Suppose otherwise, so that by Lemma 4.2 we have $g = (g_1, \dots, g_r)$ acting on P where without loss of generality the action of g_1 on X_1 is parabolic and none of the actions of g_i on X_i are loxodromic for $1 \leq i \leq r$. We now invoke [26, Proposition 3.2]. This states that if X is a geodesic δ -hyperbolic space and S is any finite set of isometries of X such that neither S nor S^2 contains a loxodromic element, then the joint minimum displacement $L(S)$ is at most 100δ . Here $L(S)$ is defined to be the infimum over all points $x \in X$ of the joint displacement $\max_{s \in S} d(x, s(x))$.

For each i we apply this result to the δ_i -hyperbolic space X_i . We have no loxodromic elements in $\langle g_i \rangle$, thus for any finite subset S of $\langle g_i \rangle$ there is some point $x_i \in X_i$ which is moved by at most distance $101\delta_i$ by any $s \in S$. In particular, for any $N > 0$ this applies to the set $S_N = \{g_i, \dots, g_i^N\}$. We will label the point obtained above when this result is applied with $S = S_N$ as $x_i^{(N)} \in X_i$.

We now show that this action of $\langle g \rangle$ on P is not acylindrical and hence nor is the action of G on P . Set ϵ to be $101\delta r$, where $\delta = \max(\delta_1, \dots, \delta_r)$ and suppose we are given R and N . Now as g_1 does act parabolically on X_1 , the orbit of $x_1^{(N)}$ under $\langle g_1 \rangle$ is not bounded and so we can find g_1^K where, on setting $y_1^{(N)} := g_1^K(x_1^{(N)})$, we have $d_1(x_1^{(N)}, y_1^{(N)}) \geq R$, where d_i is the distance in X_i . Let $y_i^{(N)}$ be $g_i^K(x_i^{(N)})$.

Now for all $1 \leq i \leq r$ and $1 \leq j \leq N$ we have that $d_i(g_i^j(x_i^{(N)}), x_i^{(N)}) \leq 101\delta$ and also

$$d_i(g_i^j(y_i^{(N)}), y_i^{(N)}) = d_i(g_i^{K+j}(x_i^{(N)}), g_i^K(x_i^{(N)})) = d_i(g_i^j(x_i^{(N)}), x_i^{(N)})$$

which therefore is at most 101δ as well. Thus if we take the two points $x = (x_1^{(N)}, \dots, x_r^{(N)})$ and $y = (y_1^{(N)}, \dots, y_r^{(N)})$ of P then we have $d(x, y) \geq d_1(x_1^{(N)}, y_1^{(N)}) \geq R$ but the N distinct elements g^j for $1 \leq j \leq N$ satisfy

$$d(g^j(x), x) = d_1(g_1^j(x_1^{(N)}), x_1^{(N)}) + \cdots + d_r(g_r^j(x_r^{(N)}), x_r^{(N)}) = d(g^j(y), y) \leq 101\delta r$$

so that each of these elements moves both x and y by at most ϵ . □

4.2 Acylindrical actions of generalised Baumslag–Solitar groups

A generalised Baumslag–Solitar group can never be acylindrically hyperbolic. This can be seen by using [32, Theorem 3.7], which states that if G is acylindrically hyperbolic and has a subgroup H where $H \cap gHg^{-1}$ is infinite for all $g \in G$ then H must itself be acylindrically hyperbolic. For a GBS_n group G we can of course take H to be a base $\mathbb{Z}^n$ -subgroup for a contradiction.

However certainly there are GBS_n groups which possess a product acylindrical action, for instance $\mathbb{Z}^n \times F_r$ has an obvious geometric (and hence acylindrical) action on $\mathbb{R} \times \cdots \times \mathbb{R} \times T_{2r}$, where T_d is

the regular tree of degree d . We are interested in when a GBS_n group G has a finite index subgroup H possessing a product acylindrical action. (Note that even for acylindrical hyperbolicity, it is not known whether H having this property and $H \leq_f G$ implies that G has this property too. See the correction to [32] which applies to Lemma 3.9.) However if we consider the property of virtually having a proper acylindrical action then we can now give a complete answer in the case of GBS_n groups.

Theorem 4.4 *For any $n \geq 1$, a GBS_n group G has a finite index subgroup H possessing a product acylindrical action if and only if the monodromy of G is finite.*

Proof First if the monodromy of G is finite then by Theorem 3.4 we have $H \leq_f G$ with H of the form $\mathbb{Z}^n \times F_r$ and this has a geometric, hence acylindrical, action on the product of $n + 1$ hyperbolic spaces which preserves factors.

Now say that the monodromy of G is infinite. Then so will the monodromy of H for any finite index subgroup H , hence we just need to rule out that G has a product acylindrical action.

We thus suppose that G acts on the product P of hyperbolic spaces $X_1 \times \cdots \times X_r$ by isometries, and that this action preserves factors and is acylindrical. By Theorem 3.5 we have an infinite order element z of G which lies in our base $\mathbb{Z}^n$ -subgroup A and which cannot be loxodromic in any action of G on a hyperbolic space. Hence on splitting z into its component parts $(z_1, \dots, z_r)$ with each z_i acting as an isometry of the hyperbolic space X_i , we have that z_i must act parabolically or elliptically on X_i . By Lemma 4.2, z must be a parabolic or elliptic element in the action of G on P . But by Theorem 4.3 there are no parabolic elements if we have a product acylindrical action of G .

Thus z must be an elliptic element. Whilst we can certainly have acylindrical actions with elliptic infinite order elements, or indeed with every element acting elliptically, we are not in the latter case here because in the definition of a product acylindrical action we must have some other element of G which is acting loxodromically.

Hence let $g \in G$ be this element, with $g = (g_1, \dots, g_r)$ and without loss of generality g_1 acts loxodromically on X_1 . Let us set G_2 to be the intersection of the index-2 subgroups of G as before and note that by the proof of Theorem 3.5, we can take for z any nonidentity element of $A \cap G_2$ which is trivial in the free abelianisation of G_2 . Consider the subset Z of $A \cap G_2$ consisting of elements with this property along with the identity and recall that this forms a subgroup of $A \cap G_2$ and so it is a finitely generated free abelian group. It is also infinite (for instance it contains all powers of a given $z \in Z$). Now for the element g as above, we can replace it by a power (which we will continue to call g) which lies in G_2 .

We will show that this action of G on P cannot in fact be acylindrical. Pick any point $x_0 \in P$ and let D be an upper bound for the set $\{d_P(z(x_0), x_0) \mid z \in Z\}$. This exists because every element z is acting elliptically and so the orbit of x under Z is bounded, because Z is finitely generated and abelian.

Set $\epsilon = D$ and suppose we are given any $R > 0$. As g acts loxodromically, there will exist $K > 0$ such that $d_P(g^K(x_0), x_0) \geq R$ and set $y_R := g^K(x_0)$. Now by definition of the modular homomorphism, there will be an integer $m > 0$ (depending on R) such that $g^K z^m g^{-K}$ is in our base subgroup A and hence in $G_2 \cap A$ as $g^K \in G_2$. But note that if z^m is trivial in the free abelianisation of G_2 then so is $g^K z^m g^{-K}$ as $g^K \in G_2$, so z^m and $g^K z^m g^{-K}$ will have the same image in the free abelianisation of G_2 . Notice this also works for powers z^{im} and $g^K z^{im} g^{-K}$ for any $i > 0$.

Thus we have infinitely many elements $\{g^K z^{im} g^{-K} \mid i \in \mathbb{N}\}$ (which are distinct because z^{im} are) lying in Z and so they each move x_0 by at most a distance D in P . But clearly we also have

$$d_P(g^K z^{im} g^{-K}(y_R), y_R) = d_P(g^K z^{im}(x_0), g^K(x_0)) = d_P(z^{im}(x_0), x_0),$$

which is at most D , and so this infinite set of elements moves both x_0 and y_R a distance at most ϵ , but $d_P(x_0, y_R) \geq R$. Thus G is not acting acylindrically on P . $\square$

5 Hierarchically hyperbolic groups and generalised Baumslag–Solitar groups

5.1 Hierarchically hyperbolic groups

The notion of a hierarchically hyperbolic space (HHS) was introduced in [4] as a way of generalising hyperbolic spaces to include mapping class groups and many CAT(0) cube complexes. We do not give a definition here but roughly speaking an HHS is a quasigeodesic metric space X together with a structure given in terms of projections $\pi_i: X \rightarrow U_i$ to a family (infinite in general) of hyperbolic spaces $\{U_i \mid i \in I\}$ called the domains (these spaces need not be proper in general and they can also be bounded). A hierarchically hyperbolic group (HHG) G is not merely a finitely generated group quasiisometric to a HHS but one where there is a HHS X where the group G acts (or rather quasiacts) geometrically on X and permutes the family of hyperbolic spaces by isometries. This family also has a nesting and an orthogonality relation. If G is a HHG then we can take X to be the Cayley graph of G with respect to a finite generating set along with the usual action of G on itself by left multiplication (thus without loss of generality we do have a genuine action of G on X rather than a quasiaction). We can also assume without loss of generality that the image $\pi_i(G)$ is coarsely dense in U_i and that this is uniform over $i \in I$.

At this point we might wonder what properties are possessed by HHGs. We have:

- (i) Hyperbolic groups are HHGs (this is seen by taking the family of hyperbolic spaces to be a single hyperbolic space).
- (ii) If G is an HHG and $H \leq_f G$ then H is an HHG [2, Lemma 2.25].
- (iii) If G_1 and G_2 are HHGs then so is $G_1 \times G_2$ [5, Corollary 8.28].

We might also wonder about obstructions for a given group G to be an HHG. Here we have:

- (1) Any HHG G is finitely presented and has quadratic isoperimetric inequality [5, Corollary 7.5].
- (2) Given any finitely generated subgroup H of an HHG G , either H is virtually abelian or $F_2 \leq H$ ([17, Theorem 9.15] where the condition of being finitely generated is not used, but see also the correction in [18]).
- (3) Every infinite order element $g \in G$ is undistorted in G with respect to word length of a finite generating set (see [17, Theorem 7.1] but again see also the correction in [18]).

Note that for each ordered pair of these three statements, there exists a group satisfying the first but not the second.

More recently another obstruction was found in [34]. We take the following result from Remark 4.9 of that paper.

Theorem 5.1 *If G is an infinite HHG then there exist finitely many unbounded hyperbolic spaces $X_1, \dots, X_r$ for $r \geq 1$ and an isometric action of some finite index subgroup H of G on the product space $P = X_1 \times \dots \times X_r$ (with the l_1 metric) such that H preserves each factor X_i and H acts acylindrically on the product P . Moreover there exists a loxodromic element in the action of H on X_i , thus this action of H is product acylindrical.*

Proof Theorem 3.2 of [34] shows that for any HHG G there is a finite G -invariant set $\mathcal{W} = \{W_1, \dots, W_r\}$ of unbounded domains which are pairwise orthogonal and such that any unbounded domain U_i is nested in one of these W_j . In the case when G is finite $\mathcal{W}$ is necessarily empty as all domains will be bounded. Conversely an HHS with all domains bounded is itself bounded, so whenever G is infinite there will be some unbounded domain and therefore $\mathcal{W}$ is nonempty.

Consequently there is a finite index subgroup H of G with

$$H(W_j) = W_j \quad \text{for all } 1 \leq j \leq r$$

and H acts by isometries on each domain W_j . Moreover the action of H on each W_j is cobounded. This is because we can assume for any domain U_i and any $g \in G$, $x \in X (= G)$ that $g(\pi_i(x)) = \pi_i(gx)$ by Remark 2.1 of [18]. But we mentioned above that $\pi_i(G) = G(\pi_i(\text{id}))$ is coarsely dense and therefore so is $\pi_i(H) = H(\pi_i(\text{id}))$ as H has finite index in G . However any cobounded action of a group on an unbounded hyperbolic space must contain a loxodromic element, as mentioned above when we listed the five types of action.

We can now get H to act on the product $W_1 \times \dots \times W_r$ with the ℓ_1 metric of these unbounded domains using the diagonal action. This action clearly preserves factors and we have said that it will contain a loxodromic element. But as pointed out in [34, Remark 4.9], the proof in [4, Theorem 14.3] that G acts acylindrically on S in the case where $\mathcal{W}$ consists of the single unbounded domain S applies equally to the above action of H on $W_1 \times \dots \times W_r$ which preserves factors and contains a loxodromic element in the action, thus this action of H is product acylindrical. $\square$

5.2 Generalised Baumslag–Solitar groups and HHGs

We can now use the results above for our main application. Our initial question might be which generalised Baumslag–Solitar groups are HHGs. However it is possible for a group G to have a finite index subgroup which is an HHG but for G not to be. This was shown in [34, Corollary 4.5] by using Theorem 3.2 in that paper which we have already quoted above. Specifically they show that the (orientation preserving) $(3, 3, 3)$ triangle group is not an HHG but of course it has the finite index subgroup $\mathbb{Z}^2$ which is. A group G which is virtually an HHG will still have good group theoretic and geometric properties as these will be inherited from the finite index subgroup, so we now give a complete answer to which generalised Baumslag–Solitar groups are virtually HHGs.

Corollary 5.2 *If G is a GBS_n group then G is virtually a HHG if and only if G has finite monodromy.*

Proof If G has finite monodromy then G is virtually $\mathbb{Z}^n \times F_r$ by Theorem 3.4 and the latter group is an HHG as it is a direct product of HHGs.

Now suppose the GBS_n group G has a finite index subgroup L which is a HHG. By Theorem 5.1 there is some finite index subgroup H of L which has a product acylindrical action (unless L is finite, in which case G was not a generalised Baumslag–Solitar group). But H is then a finite index subgroup of G too and so by Theorem 4.4 we have that the monodromy of G must be finite. $\square$

The result above in [34] that the $(3, 3, 3)$ triangle group is not a HHG even though it is virtually $\mathbb{Z}^2$ established that being a HHG is not a quasiisometry invariant and indeed not even a commensurability invariant. However, whenever we have a property $\mathcal{P}$ of abstract groups which is invariant under taking finite index subgroups, as is the case for being a HHG, we can recover a commensurability invariant by using the property of being virtually $\mathcal{P}$. Thus although [34] tells us that being a HHG is not a commensurability invariant and thus not a quasiisometric invariant, being virtually a HHG is certainly a commensurability invariant so the obvious question now is whether it is preserved under quasiisometry. By our results here along with famous recent examples of Leary and Minasyan, we can show the answer is no. The next theorem is an adaptation of [28, Theorem 7.2] to GBS_n groups.

Theorem 5.3 *Suppose that G is a GBS_n group with monodromy $\mathcal{M} \leq \text{GL}(n, \mathbb{R})$ which is conjugate in $\text{GL}(n, \mathbb{R})$ to a subgroup of the orthogonal group $O(n)$. Then G is a $\text{CAT}(0)$ group acting properly, freely and cocompactly on $T \times \mathbb{R}^n$, where T is a locally finite tree. Conversely if G is virtually a $\text{CAT}(0)$ group then its monodromy can be conjugated into $O(n)$.*

Proof We take for T the Bass–Serre tree of the decomposition of G as a GBS_n group, so that G acts on T by tree automorphisms. We also pick some base vertex v_0 of T and we set our base $\mathbb{Z}^n$ -subgroup A to be the intersection of all vertex and edge subgroups obtained from this decomposition, which will have finite index in each of these subgroups, along with a free generating set $\langle a_1, \dots, a_n \rangle$ for A . We will now define a homomorphism θ from G to $\text{Isom}(\mathbb{R}^n)$ and then let G act on $T \times \mathbb{R}^n$ via the diagonal action, which will

be also be by isometries. (This homomorphism is similar to the version of the modular homomorphism used in [14], which maps to $\mathbb{R}^n \rtimes \mathrm{GL}_n(\mathbb{R})$ but which does not specifically consider orthogonal matrices.)

Suppose that the matrix $X \in \mathrm{GL}(n, \mathbb{R})$ is such that $X\mathcal{M}(G)X^{-1}$ lies in $O(n)$. If $e_1, \dots, e_n$ is the standard orthogonal basis for $\mathbb{R}^n$ then set α_i to be the vector $X(e_i)$ in $\mathbb{R}^n$. We now let the base subgroup A act via θ on $\mathbb{R}^n$ as a group of translations, where we define $\theta(a_i)$ to be the map $x \mapsto x + \alpha_i$. Recalling how we obtain a finite presentation for G , we next extend θ to all vertex subgroups by using the appropriate translations, where we map consistently across edge subgroups for the edges in some maximal tree, by working out from v_0 . This works because any edge subgroup E is a copy of $\mathbb{Z}^n$ with finite index in its neighbouring vertex subgroup V , so a homomorphism from E to $\mathbb{R}^n$ has a unique extension from V to $\mathbb{R}^n$.

We now need to set $\theta(t)$ for all stable letters t . Once we have chosen $\theta(t)$ to be some isometry of $\mathbb{R}^n$, θ will be well defined if t acts by conjugation on A in the correct way because A has finite index in any vertex group.

Therefore we send t to the orthogonal matrix $X\mathcal{M}(t)X^{-1}$. Let us write the matrix $\mathcal{M}(t)$ as $\mathcal{T}$ with ij^{th} entry equal to τ_{ij} . We then have that $\theta(t)\theta(a_j)\theta(t^{-1})$ is the translation by

$$X\mathcal{T}X^{-1}(\alpha_j) = X\mathcal{T}(e_j) = X(\tau_{1j}e_1 + \dots + \tau_{nj}e_n) = \tau_{1j}\alpha_1 + \dots + \tau_{nj}\alpha_n.$$

As for the translation given by $\theta(ta_jt^{-1})$, we can assume ta_jt^{-1} is in A by replacing a_j with a power if necessary, whereupon we have

$$ta_jt^{-1} = a_1^{\tau_{1j}} \dots a_n^{\tau_{nj}}$$

by definition of the modular homomorphism and so $\theta(ta_jt^{-1})$ will be the translation of $\mathbb{R}^n$ by

$$\tau_{1j}\alpha_1 + \dots + \tau_{nj}\alpha_n$$

too. Thus θ is well defined and so G acts on $T \times \mathbb{R}^n$ by isometries.

This action of G on $T \times \mathbb{R}^n$ is (metrically) proper: for instance see [14, Lemma 4.5]. This implies that the action is free (and in particular faithful) because G is torsion free. It is also cocompact, as we can take the compact set $C \times D \subseteq T \times \mathbb{R}^n$, where C is a closed ball centred at v_0 and containing a fundamental domain for the action of G on T , whereas D is a closed parallelepiped containing a fundamental domain for the action of A on $\mathbb{R}^n$. Given any point $(t, x) \in T \times \mathbb{R}^n$, we can find $g \in G$ which moves t into C , and then take some element $a \in A$ which moves $g(x)$ into D . As $a \in \mathrm{Stab}(v_0)$ we have $a(C) = C$ so that $ag(t, x) \in C \times D$.

For the converse, first suppose that G is a CAT(0) group, acting geometrically (or even properly and semisimply) on the CAT(0) space X . Then by [28, Theorem 6.4(4')] (itself based on the flat torus theorem of [9, Theorem II.7.1], but replacing elements that normalise a copy of $\mathbb{Z}^n$ with elements that commensurate it), there is an inner product on $A \otimes \mathbb{R} \cong \mathbb{R}^n$ such that the monodromy of any isometry of X that commensurates A preserves this inner product, so that we can conjugate the monodromy of G to lie in $O(n)$.

Finally if G has a finite index subgroup H that is a CAT(0) group then by the above we can conjugate the monodromy $\mathcal{M}(G)$ so that the finite index subgroup $\mathcal{M}(H)$ lies in $O(n)$. We can now further conjugate so that all of $\mathcal{M}(G)$ is in $O(n)$ too (either by mimicking the proof when $\mathcal{M}(H)$ is trivial, or by $O(n)$ being a maximal compact subgroup of $\mathrm{GL}(n, \mathbb{R})$). $\square$

We can now bring in examples from [28].

Corollary 5.4 *The Leary–Minasyan group given by the finite presentation*

$$L = \langle t, a, b \mid [a, b], ta^2b^{-1}t^{-1} = a^2b, tab^2t^{-1} = a^{-1}b^2 \rangle$$

is a GBS_2 group which is not virtually a HHG but which is quasiisometric to a HHG.

Proof We see that L is a GBS_2 group with base $\mathbb{Z}^2$ -subgroup $\langle a, b \rangle$ by using the graph of groups decomposition of one vertex and one edge associated to the HNN extension, with the edge subgroup having index 5 in the vertex group at both ends. Thus the Bass–Serre tree is the regular tree T_{10} . The monodromy is determined by the conjugation action of the single stable letter and we can work this out explicitly by noting that the above relations imply that

$$ta^5t^{-1} = a^3b^4 \quad \text{and} \quad tb^5t^{-1} = a^{-4}b^3$$

both hold in G . Thus the monodromy of L is generated by the matrix

$$\begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$$

which has infinite order. Hence L has infinite monodromy and is not virtually a HHG by Corollary 5.2.

However this matrix is orthogonal so Theorem 5.3 says that L acts properly and cocompactly by isometries on $T_{10} \times \mathbb{R}^2$, as does $M = F_5 \times \mathbb{Z}^2$. Thus, by the Švarc–Milnor lemma, L and M are both quasiisometric to this space and hence to each other, but M is certainly a HHG. $\square$

6 Property (QT) is not invariant under quasiisometry

Another property of groups which, like HHGs, considers how a group can act on different hyperbolic spaces, is Bestvina, Bromberg and Fujiwara’s property (QT) from [6]. Here a quasitree will always be a graph equipped with its path metric which is quasiisometric to a simplicial tree but which need not be locally finite. A finitely generated group G (equipped with the word metric with respect to some finite generating set) is said to have (QT) if it acts by isometries on a finite product P of quasitrees equipped with the ℓ_1 product metric such that the orbit map (using an arbitrary basepoint of P) is a quasiisometric embedding from G to P . This is a strong definition: for instance it implies that this action is metrically proper. Nevertheless it is shown in [6] that mapping class groups and all residually finite hyperbolic groups have (QT). It is also a consequence of [16] that every Coxeter group has (QT).

Moreover property (QT) has good closure properties, in fact these are better than for HHGs. It is certainly the case that if G has (QT) and H has finite index in G then H also has (QT) but the definition ensures that it also holds more generally when H is an undistorted finitely generated subgroup of G . Moreover if G_1 has (QT) via an action on the space P_1 and G_2 on the space P_2 then it can be checked directly that $G_1 \times G_2$ has (QT) by letting it act on the direct product $P_1 \times P_2$ (with the ℓ_1 product metric) using the action on each factor and summing the word metrics on G_1 and G_2 , which is the word metric on $G_1 \times G_2$ with the obvious generating set.

However property (QT) is also a commensurability invariant because if H has index i in G and H acts isometrically on the product P of quasitrees then we can induce an isometric action of G on the product P^i of copies of P . This will also turn the orbit map under G into a quasiisometric embedding. In particular a group which virtually has (QT) does itself have (QT). Therefore the questions of whether possessing (QT) and virtually possessing (QT) are quasiisometry invariants are in fact the same question. Here we will answer this by first considering how an isometry on a product of graphs $\Gamma_1 \times \cdots \times \Gamma_m$ (each equipped with the path metric and then using the ℓ_1 product metric) breaks up, or at least virtually breaks up, into individual isometries on each Γ_i . That this can be done using the ℓ_∞ metric is a result of W Malone in [30] and we will now mimic his proof for the (easier) ℓ_1 case.

Theorem 6.1 *Suppose that $X = \Gamma_1 \times \cdots \times \Gamma_m$ is a finite product of connected graphs, where each Γ_i has the induced path metric and X has the ℓ_1 or the ℓ_∞ product metric. Suppose that G is any group acting by isometries on X . Then G has a finite index subgroup H which preserves factors and acts as an isometry on each factor.*

Proof On giving a graph the induced path metric, it becomes not just a metric space which is geodesic but a geodesic metric space which is locally uniquely geodesic. Thus if X is given the ℓ_∞ metric then this result is a direct consequence of [30], which is established by using the fact that for certain directions in the product space, geodesics are unique if they are unique in the factors. This is certainly true in an ℓ_1 product space if we travel “horizontally/vertically”, so in this case we proceed as follows:

Lemma 6.2 *Let $X = X_1 \times \cdots \times X_m$ have the ℓ_1 product metric d_X for geodesic metric spaces $(X_1, d_1), \dots, (X_m, d_m)$. Suppose we have two points $\mathbf{x}, \mathbf{y} \in X$ of the form*

$$\mathbf{x} = (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m) \quad \text{and} \quad \mathbf{y} = (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_m)$$

for $1 \leq i \leq m$. Then any geodesic between $\mathbf{x}$ and $\mathbf{y}$ only varies in the i^{th} coordinate. Conversely if we have two points $\mathbf{x}, \mathbf{y} \in X$ which differ in at least two coordinates then there is more than one geodesic from $\mathbf{x}$ to $\mathbf{y}$ in X .

Proof Suppose we have a (unit speed) geodesic $\gamma: [0, d] \rightarrow X$ from $\mathbf{x}$ to $\mathbf{y}$, where $d = d_X(\mathbf{x}, \mathbf{y})$. If there is $k \neq i$ and $t \in (0, d)$ such that the k^{th} coordinate u_k of $\gamma(t) = (u_1, \dots, u_m)$ is not equal to x_k

then

$$\begin{aligned}
 d_i(x_i, y_i) &= d_X(\mathbf{x}, \mathbf{y}) = d(\mathbf{x}, \boldsymbol{\gamma}(t)) + d(\boldsymbol{\gamma}(t), \mathbf{y}) \\
 &= \sum_{j=1, j \neq i}^m (d_j(x_j, u_j) + d_j(u_j, x_j)) + d_i(x_i, u_i) + d_i(u_i, y_i) \\
 &\geq d_k(x_k, u_k) + d_k(u_k, x_k) + d_i(x_i, u_i) + d_i(u_i, y_i) \\
 &> d_i(x_i, y_i) \quad \text{since } x_k \neq u_k,
 \end{aligned}$$

which is a contradiction.

Now suppose we have two points $\mathbf{x} = (x_1, \dots, x_m)$ and $\mathbf{y} = (y_1, \dots, y_m)$ in X which differ in (without loss of generality) at least the first two coordinates. Take constant speed geodesics γ_i in each X_i running from x_i to y_i and set

$$\delta_1(t) = (x_1, \gamma_2(t), \dots, \gamma_m(t)), \quad \delta_2(t) = (\gamma_1(t), y_2, \dots, y_m))$$

and

$$\delta_3(t) = (\gamma_1(t), x_2, \gamma_3(t), \dots, \gamma_m(t)), \quad \delta_4(t) = (y_1, \gamma_2(t), y_3, \dots, y_m)),$$

where in δ_1 the geodesics $\gamma_2, \dots, \gamma_m$ are reparametrised to have domain $[0, d_X(\mathbf{x}, \mathbf{y}) - d_1(x_1, y_1)]$ but γ_1 and hence δ_2 remain unit speed. We also do the same for δ_3 and δ_4 . Then it is easily checked that following δ_1 then δ_2 and also following δ_3 then δ_4 are both unit speed geodesics from $\mathbf{x}$ to $\mathbf{y}$. Moreover they are distinct as the first geodesic passes through a point that projects to (x_1, y_2) in the first two coordinates, whereas the second geodesic never does since $x_1 \neq y_1$ and $x_2 \neq y_2$. $\square$

Now we return to the setting in [Theorem 6.1](#). Each Γ_i is a geodesic metric space (which we assume without loss of generality is not a single point). Thus given any isometry g of X , take any point $\mathbf{x} = (x_1, \dots, x_i, \dots, x_m)$ in X and for a given $1 \leq i \leq m$, let z_i be another point in Γ_i near x_i such that there is only one geodesic γ_i in Γ_i from x_i to z_i . Then by [Lemma 6.2](#) we have that $\boldsymbol{\gamma}(t)$, which is equal to $(x_1, \dots, \gamma_i(t), x_{i+1}, \dots, x_m)$, is the unique geodesic in X between $\mathbf{x}$ and $\mathbf{z} = (x_1, \dots, x_{i-1}, z_i, x_{i+1}, \dots, x_m)$ and so under g it must map to a unique geodesic between $g(\mathbf{x})$ and $g(\mathbf{z})$, by considering g^{-1} . Thus by [Lemma 6.2](#) again these two points (which are not the same because g is a bijection) differ in only one coordinate, say the j^{th} .

Now take y_i to be an arbitrary point in Γ_i . Given any geodesic in Γ_i from x_i to y_i , we can split it into a finite number of subgeodesics, with some overlap that is more than a point, where each subgeodesic is the unique geodesic between its own endpoints. Thus the image of each of these subgeodesics under g is a subset of X where only one coordinate varies, and as these subgeodesics overlap this will always be the j^{th} coordinate. In other words given $1 \leq i \leq m$ we have $1 \leq j \leq m$ such that for fixed

$$x_1 \in \Gamma_1, \quad \dots, \quad x_{i-1} \in \Gamma_{i-1}, \quad x_{i+1} \in \Gamma_{i+1}, \quad \dots, \quad x_m \in \Gamma_m$$

there is

$$y_1 \in \Gamma_1, \quad \dots, \quad y_{j-1} \in \Gamma_{j-1}, \quad y_{j+1} \in \Gamma_{j+1}, \quad \dots, \quad y_m \in \Gamma_m$$

and a function $f: \Gamma_i \rightarrow \Gamma_j$ with

$$g(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_m) = (y_1, \dots, y_{j-1}, f(x), y_{j+1}, \dots, y_m)$$

for all $x \in \Gamma_i$. Moreover f is a bijection and an isometry because g is. By varying over all m coordinates we see that g permutes $\Gamma_1, \dots, \Gamma_m$ and for each $1 \leq k \leq m$ it acts as an isometry from Γ_k to $\Gamma_{g(k)}$ (but clearly only isometric factors can be permuted). We can therefore set H to be the kernel of this finite permutation action, so that H has finite index in G , preserves factors and acts as an isometry on each factor. $\square$

We can now use the same example as for HHGs to show that property (QT) is not a quasiisometric invariant.

Theorem 6.3 *A GBS_n G group with infinite monodromy does not possess any metrically proper action by isometries on any finite product of graphs (using the path metric) which are quasitrees with the l_1 (or l_∞) product metric. In particular G does not have property (QT).*

Proof If G did have such an action then by Theorem 6.1 G would have a finite index subgroup H (also acting metrically properly) which preserves factors and acts by isometries on each factor. But H is also a GBS_n group with infinite monodromy. Thus by Theorem 3.5 applied to H , there is a nontrivial element z of H which does not act loxodromically in any action of H on a hyperbolic space. But as isometries of quasitrees can only be loxodromic or elliptic by [31] the action of the infinite order element z is elliptic on each quasitree. Hence the action of $\langle z \rangle$ on X is also bounded, so we have an infinite subgroup of G acting on X with bounded orbits, meaning that the action of G is not metrically proper. Therefore the orbit map from G to X cannot be a quasiisometric embedding because there are only finitely many elements in G with word length at most a given value. $\square$

Corollary 6.4 *The property of having (QT) is not preserved under quasiisometries.*

Proof We take the Leary–Minasyan group L as given in Corollary 5.4 and again note that it has infinite monodromy, so it does not have (QT) by Theorem 6.3. However it is quasiisometric to $M = \mathbb{Z}^2 \times F_5$ which acts geometrically on the product $P = \mathbb{R} \times \mathbb{R} \times T_{10}$ of simplicial trees and hence the orbit map from M to P is a quasiisometry by Švarc–Milnor. $\square$

Note that as P is a product not just of arbitrary quasitrees but of bounded valence simplicial trees, changing the property (QT) by replacing quasitrees with trees or inserting the bounded valence condition, in any combination, will still not result in a quasiisometry invariant.

Note also that if we equip our product spaces with the ℓ_2 product metric then the Leary–Minasyan group does indeed act geometrically on a product of quasitrees with the orbit map a quasiisometric embedding. Here the equivalent of Theorem 6.1 does not hold as we can have a de Rham factor (see [19] for this case). However the definition of (QT) uses the ℓ_1 product metric.

Finally we also note that the Leary–Minasyan group acts geometrically on the CAT(0) cube complex P by *isometries* using the CAT(0) metric but it has no geometric action on any CAT(0) cube complex by *cubical automorphisms* and nor does any finite index subgroup. This is noted in [28] as a consequence of [24].

References

- [1] C Abbott, S H Balasubramanya, D Osin, *Hyperbolic structures on groups*, Algebr. Geom. Topol. 19 (2019) 1747–1835 [MR](#) [Zbl](#)
- [2] C R Abbott, T Ng, D Spriano, R Gupta, H Petyt, *Hierarchically hyperbolic groups and uniform exponential growth*, Math. Z. 306 (2024) art. id. 18 [MR](#) [Zbl](#)
- [3] B Beeker, *Abelian JSJ decomposition of graphs of free abelian groups*, J. Group Theory 17 (2014) 337–359 [MR](#) [Zbl](#)
- [4] J Behrstock, M F Hagen, A Sisto, *Hierarchically hyperbolic spaces, I: Curve complexes for cubical groups*, Geom. Topol. 21 (2017) 1731–1804 [MR](#) [Zbl](#)
- [5] J Behrstock, M Hagen, A Sisto, *Hierarchically hyperbolic spaces, II: Combination theorems and the distance formula*, Pacific J. Math. 299 (2019) 257–338 [MR](#) [Zbl](#)
- [6] M Bestvina, K Bromberg, K Fujiwara, *Proper actions on finite products of quasi-trees*, Ann. H. Lebesgue 4 (2021) 685–709 [MR](#) [Zbl](#)
- [7] O Bogopolski, A Martino, E Ventura, *Orbit decidability and the conjugacy problem for some extensions of groups*, Trans. Amer. Math. Soc. 362 (2010) 2003–2036 [MR](#) [Zbl](#)
- [8] B H Bowditch, *Tight geodesics in the curve complex*, Invent. Math. 171 (2008) 281–300 [MR](#) [Zbl](#)
- [9] M R Bridson, A Haefliger, *Metric spaces of non-positive curvature*, Grundle. Math. Wissen. 319, Springer (1999) [MR](#) [Zbl](#)
- [10] P-E Caprace, Y Cornuier, N Monod, R Tessera, *Amenable hyperbolic groups*, J. Eur. Math. Soc. 17 (2015) 2903–2947 [MR](#) [Zbl](#)
- [11] M Casals-Ruiz, I Kazachkov, A Zakharov, *Commensurability of Baumslag–Solitar groups*, Indiana Univ. Math. J. 70 (2021) 2527–2555 [MR](#)
- [12] C H Cashen, G Levitt, *Mapping tori of free group automorphisms, and the Bieri–Neumann–Strebel invariant of graphs of groups*, J. Group Theory 19 (2016) 191–216 [MR](#) [Zbl](#)
- [13] Y de Cornulier, R Tessera, *Metabelian groups with quadratic Dehn function and Baumslag–Solitar groups*, Confluentes Math. 2 (2010) 431–443 [MR](#) [Zbl](#)
- [14] Y Cornulier, A Valette, *On equivariant embeddings of generalized Baumslag–Solitar groups*, Geom. Dedicata 175 (2015) 385–401 [MR](#) [Zbl](#)
- [15] R B Coulon, *Partial periodic quotients of groups acting on a hyperbolic space*, Ann. Inst. Fourier (Grenoble) 66 (2016) 1773–1857 [MR](#) [Zbl](#)
- [16] A Dranishnikov, T Januszkiewicz, *Every Coxeter group acts amenably on a compact space*, Topology Proc. 24 (1999) 135–141 [MR](#) [Zbl](#)
- [17] M G Durham, M F Hagen, A Sisto, *Boundaries and automorphisms of hierarchically hyperbolic spaces*, Geom. Topol. 21 (2017) 3659–3758 [MR](#) [Zbl](#)

- [18] **M G Durham, M F Hagen, A Sisto**, *Correction to the article Boundaries and automorphisms of hierarchically hyperbolic spaces*, Geom. Topol. 24 (2020) 1051–1073 [MR](#)
- [19] **T Foertsch, A Lytchak**, *The de Rham decomposition theorem for metric spaces*, Geom. Funct. Anal. 18 (2008) 120–143 [MR](#)
- [20] **M Forester**, *Deformation and rigidity of simplicial group actions on trees*, Geom. Topol. 6 (2002) 219–267 [MR](#) [Zbl](#)
- [21] **M F Hagen**, *Cocompactly cubulated crystallographic groups*, J. Lond. Math. Soc. 90 (2014) 140–166 [MR](#) [Zbl](#)
- [22] **M Hagen, J Russell, A Sisto, D Spriano**, *Equivariant hierarchically hyperbolic structures for 3-manifold groups via quasimorphisms*, Ann. Inst. Fourier (Grenoble) (online publication July 2024)
- [23] **P de la Harpe**, *Topics in geometric group theory*, Univ. Chicago Press (2000) [MR](#) [Zbl](#)
- [24] **J Huang, T Prytuła**, *Commensurators of abelian subgroups in CAT(0) groups*, Math. Z. 296 (2020) 79–98 [MR](#) [Zbl](#)
- [25] **S Hughes, M Valiunas**, *Commensurating HNN-extensions: hierarchical hyperbolicity and biautomaticity*, Comment. Math. Helv. 99 (2024) 397–436 [MR](#) [Zbl](#)
- [26] **M Koubi**, *Croissance uniforme dans les groupes hyperboliques*, Ann. Inst. Fourier (Grenoble) 48 (1998) 1441–1453 [MR](#) [Zbl](#)
- [27] **P H Kropholler**, *Baumslag–Solitar groups and some other groups of cohomological dimension two*, Comment. Math. Helv. 65 (1990) 547–558 [MR](#) [Zbl](#)
- [28] **I J Leary, A Minasyan**, *Commensurating HNN extensions: nonpositive curvature and biautomaticity*, Geom. Topol. 25 (2021) 1819–1860 [MR](#) [Zbl](#)
- [29] **G Levitt**, *On the automorphism group of generalized Baumslag–Solitar groups*, Geom. Topol. 11 (2007) 473–515 [MR](#) [Zbl](#)
- [30] **W Malone**, *Isometries of products of path-connected locally uniquely geodesic metric spaces with the sup metric are reducible*, preprint (2009) [arXiv 0912.3471](#)
- [31] **J F Manning**, *Quasi-actions on trees and property (QFA)*, J. Lond. Math. Soc. 73 (2006) 84–108 [MR](#) [Zbl](#)
- [32] **A Minasyan, D Osin**, *Acylindrical hyperbolicity of groups acting on trees*, Math. Ann. 362 (2015) 1055–1105 [MR](#) [Zbl](#) Correction in 373 (2019) 895–900
- [33] **D Osin**, *Acylindrically hyperbolic groups*, Trans. Amer. Math. Soc. 368 (2016) 851–888 [MR](#) [Zbl](#)
- [34] **H Petyt, D Spriano**, *Unbounded domains in hierarchically hyperbolic groups*, Groups Geom. Dyn. 17 (2023) 479–500 [MR](#) [Zbl](#)

Selwyn College, University of Cambridge
Cambridge, United Kingdom

jb128@cam.ac.uk

Received: 30 May 2023 Revised: 5 January 2024

Geometric and arithmetic properties of Löbell polyhedra

NIKOLAY BOGACHEV

SAMI DOUBA

The Löbell polyhedra form an infinite family of compact right-angled hyperbolic polyhedra in dimension 3. We observe, through both elementary and more conceptual means, that the “systoles” of the Löbell polyhedra approach 0, so that these polyhedra give rise to particularly straightforward examples of closed hyperbolic 3-manifolds with arbitrarily small systole, and constitute an infinite family even up to commensurability. By computing number-theoretic invariants of these polyhedra, we refine the latter result, and also determine precisely which of the Löbell polyhedra are quasi-arithmetic.

[57K32](#); [22E40](#), [57M50](#)

1 Introduction

For each $n \geq 5$, consider a combinatorial 3-polyhedron whose “top” and “bottom” faces are n -gons and whose “lateral” surface consists of $2n$ pentagons (see [Figure 1](#) and [Figure 3](#), left, for illustrations of the case $n = 6$). By Andreiev’s theorem (see Andreiev [\[3\]](#) and Roeder, Hubbard, and Dunbar [\[27\]](#)), this abstract polyhedron can be realized as a compact right-angled polyhedron L_n in $\mathbb{H}^3$ for any $n \geq 5$. For instance, the polyhedron L_5 is the right-angled hyperbolic dodecahedron.

The polyhedra L_n are called *Löbell polyhedra*, after FR Löbell, who constructed the first example [\[18\]](#) of a closed oriented hyperbolic 3-manifold by gluing eight copies of the polyhedron L_6 . The subgroup $\Gamma_n < \text{Isom}(\mathbb{H}^3)$ generated by the reflections in the faces of L_n is a cocompact right-angled reflection group; we refer to the quotients $\mathbb{H}^3 / \Gamma_n$ as *Löbell orbifolds*. Such an orbifold $\mathbb{H}^3 / \Gamma_n$ can be visualized as the polyhedron L_n itself with reflective singularities in its faces. The example of Löbell is a particular instance of a general construction of degree-8 manifold covers of right-angled reflection 3-orbifolds; see Vesnin’s [\[30\]](#) or [\[33, Section 3.1\]](#).

Much is understood about Löbell polyhedra. In particular, Vesnin [\[32\]](#) computed their volumes $\text{vol}(L_n)$, and showed that $\text{vol}(L_n)$ is an increasing function of n . Moreover, Mednykh and Vesnin [\[21\]](#) showed that the distance δ_n between the top and bottom faces of the Löbell polyhedron L_n satisfies

$$\cosh \delta_n = \frac{\cos(\pi/n)}{\cos(2\pi/n)}.$$

The following is an elementary consequence of the above identity:

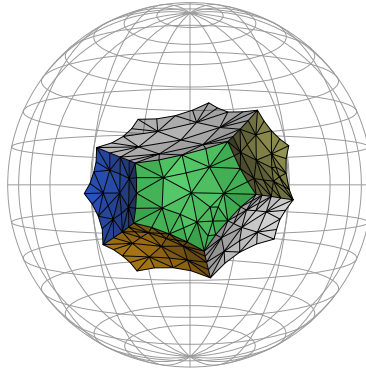


Figure 1: The Löbell polyhedron L_6 , displayed in the Poincaré ball model of $\mathbb{H}^3$, based on an image created with Geomview [25].

Observation 1.1 *The distance δ_n between the top and bottom faces of the Löbell polyhedron L_n approaches 0 as $n \rightarrow \infty$.*

Since the systole of a compact right-angled reflection orbifold is twice the minimal distance between two nonadjacent walls (see Section 3.4 for the definition of the systole of a hyperbolic orbifold), and by an application of a collar lemma of Basmajian [6] to the lateral faces of L_n , whose areas are fixed, the following is immediate from Observation 1.1:

Corollary 1.1.1 *For sufficiently large n , the systole of the Löbell orbifold L_n is precisely $2\delta_n$.*

Observation 1.1 also yields the following:

Corollary 1.1.2 *Let M_n be manifold covers of the Löbell orbifolds L_n of uniformly bounded degree (for instance, the degree-8 manifold covers discussed above). Then the systole of M_n approaches 0 as $n \rightarrow \infty$.*

Indeed, for n a multiple of 3, we provide an explicit degree-8 oriented manifold cover of L_n whose systole is precisely $2\delta_n$ for sufficiently large n ; see Figure 2. We thus obtain straightforward examples of closed hyperbolic 3-manifolds with arbitrarily small systole. Such manifolds were known to exist in dimension 3 by Thurston's hyperbolic Dehn filling theorem; see for instance Benedetti and Petronio [8, Sections E.5 and E.6]. In fact, the Löbell polyhedron L_n decomposes into $2n$ copies of a polyhedron T_n (see Section 3.1) which may be viewed as a Dehn filling of a cusped polyhedron T_∞ , so that Observation 1.1 ultimately also follows from a Dehn filling argument; see Section 4.

Agol [1] provided another construction that, given an input arithmetic lattice $\Gamma < \text{Isom}(\mathbb{H}^3)$ and $\varepsilon > 0$, outputs a closed hyperbolic 3-manifold with systole at most ε . Agol originally suggested this construction in dimension 4 (where the problem had theretofore been open), but it evidently also applies in lower dimensions, and in fact applies in all dimensions by work of Belolipetsky and Thomson [7] (alternatively, by a result of Bergeron, Haglund, and Wise [9, Theorem 1.4]). While the output lattices are nonarithmetic for

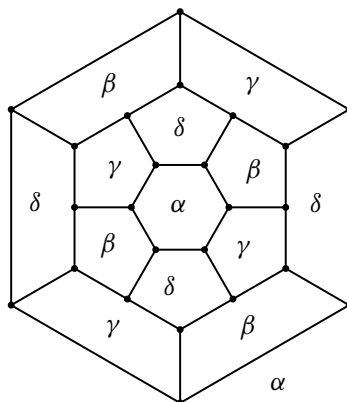


Figure 2: The above 4-coloring of the Löbell polyhedron L_6 and its analogues for the polyhedra L_{3k} for $k \geq 2$ determine (orientable) degree-8 manifold covers M_{3k} as in [33, Section 3.1]. Since the top and bottom faces of L_{3k} receive the same color, the manifold M_{3k} contains a closed geodesic of length $2\delta_{3k}$. For sufficiently large k , the systole of M_{3k} is in fact precisely $2\delta_{3k}$ by Corollary 1.1.1.

fixed¹ Γ and sufficiently small ε , they are nevertheless all quasi-arithmetic, as observed by Thomson [29]. On the other hand, the reflection lattices Γ_n are eventually not quasi-arithmetic.

Theorem 1.2 *The Löbell polyhedron L_n is quasi-arithmetic if and only if $n = 5, 6, 8, 12$, and is properly quasi-arithmetic only when $n = 12$.*

Recall that a lattice is said to be *properly quasi-arithmetic* if it is quasi-arithmetic but not arithmetic. See Section 2.2 for definitions. We remark that it is unclear to us whether Theorem 1.2 exemplifies a more general phenomenon. More precisely, it appears that the following question is open:

Question 1.3 *Are only finitely many Dehn fillings of a complete finite-volume noncompact hyperbolic 3-orbifold quasi-arithmetic?*

The answer to Question 1.3 is known to be affirmative if one replaces “quasi-arithmetic” with “arithmetic”; see Maclachlan and Reid [19, Corollary 11.2.2].

Returning to our discussion on Löbell polyhedra, Vesnin [31] observed that the Löbell polyhedron L_n is nonarithmetic for $n \neq 5, 6, 7, 8, 10, 12, 18$. Antolín-Camarena, Maloney, and Roeder [5] later showed that L_n is arithmetic if and only if $n = 5, 6, 8$.

Our proof of Theorem 1.2 is straightforward and uses only classical tools from Vinberg’s theory of hyperbolic reflection groups. Along the way, we compute the adjoint trace fields of the lattices Γ_n .

Theorem 1.4 *The adjoint trace field k_n of Γ_n is $\mathbb{Q}(\cos(2\pi/n))$. In particular, if $p, q \geq 5$ are distinct primes, then the Löbell polyhedra L_p and L_q are incommensurable.*

¹Conjecturally, there is no dependence on Γ ; see the discussion following [1, Theorem 4.3], and [7, Section 5.1].

It is shown in [19, Section 4.7.3] that there are infinitely many pairwise-incommensurable compact Coxeter polyhedra in $\mathbb{H}^3$. However, we could not find in the literature a justification of the existence of infinitely many pairwise-incommensurable *right-angled* such polyhedra. Indeed, this was our initial motivation for considering the Löbell polyhedra. Since $\deg(\cos(2\pi/n)) = \frac{1}{2}\phi(n) \rightarrow \infty$ as $n \rightarrow \infty$, where ϕ is Euler's totient function, one concludes that there is no infinite subsequence of Γ_n consisting entirely of pairwise-commensurable lattices. This fact is also implied by [Observation 1.1](#); see [Remark 3.3](#).

We remark that identical degree considerations also demonstrate that the ideal right-angled *antiprisms* A_n form an infinite family of finite-volume *noncompact* right-angled polyhedra in $\mathbb{H}^3$ even up to commensurability. Meyer, Millichap, and Trapp [22] and Kellerhals [16] have shown moreover that A_n and A_m are incommensurable for $n \neq m$, and that the adjoint trace field of A_n is again $\mathbb{Q}(\cos(2\pi/n))$. There is in fact more to be said about the relationship between the families A_n and L_n : indeed, as observed by Kolpakov [17, Section 5.1], the Löbell polyhedra L_n can be viewed as having been obtained from the antiprisms A_n via Dehn filling; see [Section 4](#). Another way of phrasing this is that the antiprisms are obtained from Löbell polyhedra by “contracting” certain edges to ideal vertices. We explain in [Section 4](#) how this trick of contracting edges of a finite-volume right-angled polyhedron is a general method for constructing infinitely many commensurability classes of such polyhedra in dimension 3, and also why an analogous trick is not available in dimension 4 ([Theorem 4.2](#)), where the existence of such an infinite family appears to be open.

Acknowledgements

We are grateful to Misha Belolipetsky, Sasha Kolpakov, Greg Kuperberg, Nicolas Tholozan, and Andrei Vesnin for helpful discussions and comments. We thank the Institut des Hautes Études Scientifiques for hosting Bogachev in the fall of 2022, during which most of this work was completed. Bogachev was partially supported by the Russian Science Foundation, grant 21-41-09011. Douba was supported by the Huawei Young Talents program.

2 Preliminaries

2.1 Hyperbolic lattices

Let $\mathbb{R}^{d,1}$ be the real vector space $\mathbb{R}^{d+1}$ equipped with the standard quadratic form f of signature $(d, 1)$, namely,

$$f(x) = -x_0^2 + x_1^2 + \cdots + x_d^2.$$

The hyperboloid $\mathcal{H} = \{x \in \mathbb{R}^{d,1} \mid f(x) = -1\}$ has two connected components

$$\mathcal{H}^+ = \{x \in \mathcal{H} \mid x_0 > 0\} \quad \text{and} \quad \mathcal{H}^- = \{x \in \mathcal{H} \mid x_0 < 0\}.$$

The d -dimensional hyperbolic space $\mathbb{H}^d$ is the manifold $\mathcal{H}^+$ with the Riemannian metric ρ induced by restricting f to each tangent space $T_p(\mathcal{H}^+)$ for $p \in \mathcal{H}^+$. This hyperbolic metric ρ satisfies $\cosh \rho(x, y) = -(x, y)$, where (x, y) is the scalar product in $\mathbb{R}^{d,1}$ associated to f . The hyperbolic d -space $\mathbb{H}^d$ is known to be the unique simply connected complete Riemannian d -manifold with constant sectional curvature -1 . *Hyperplanes* of $\mathbb{H}^d$ are intersections of linear hyperplanes of $\mathbb{R}^{d,1}$ with $\mathcal{H}^+$, and are totally geodesic submanifolds of codimension 1 in $\mathbb{H}^d$.

Let $O_{d,1} = O(f, \mathbb{R})$ be the orthogonal group of the form f , and $O'_{d,1} < O_{d,1}$ be the subgroup (of index 2) preserving $\mathcal{H}^+$. The group $O'_{d,1}$ preserves the metric ρ on $\mathbb{H}^d$, and is in fact the full group $\text{Isom}(\mathbb{H}^d)$ of isometries of the latter.

If $\Gamma < O'_{d,1}$ is a lattice, ie if Γ is a discrete subgroup of $O'_{d,1}$ with a finite-volume fundamental domain in $\mathbb{H}^d$, then the quotient $M = \mathbb{H}^d / \Gamma$ is a complete finite-volume *hyperbolic orbifold*. If Γ is torsion-free, then M is a complete finite-volume Riemannian manifold, and is called a *hyperbolic manifold*.

Now set $G = O'_{d,1}$, and suppose $\mathbf{G}$ is an admissible (for G) algebraic k -group, ie $\mathbf{G}(\mathbb{R})^o$ is isomorphic to G^o and $\mathbf{G}^\sigma(\mathbb{R})$ is a compact group for any nonidentity embedding $\sigma: k \hookrightarrow \mathbb{R}$. Then any subgroup $\Gamma < G$ commensurable with the image in G of $\mathbf{G}(\mathcal{O}_k)$ is an *arithmetic lattice* (in G) with *ground field* k .

Since G also admits nonarithmetic lattices, we discuss some weaker notions of arithmeticity for lattices in G . Following Vinberg [34], a lattice $\Gamma < G$ is called *quasi-arithmetic* with *ground field* k if some finite-index subgroup of Γ is contained in the image in G of $\mathbf{G}(k)$, where $\mathbf{G}$ is some admissible algebraic k -group, and is called *properly quasi-arithmetic* if Γ is quasi-arithmetic, but not arithmetic on the nose.

It is worth stressing that the notion of quasi-arithmeticity is indeed broader than that of arithmeticity; as was mentioned in the introduction, the nonarithmetic closed hyperbolic manifolds constructed by Agol [1] and Belolipetsky and Thomson [7] exist in all dimensions and, as observed by Thomson [29], are quasi-arithmetic. The first examples of properly quasi-arithmetic lattices in dimensions 3, 4, and 5 were constructed by Vinberg [34] via reflection groups.

2.2 Convex polyhedra and arithmetic properties of hyperbolic reflection groups

A (*hyperbolic*) *reflection group* is a discrete subgroup of $O'_{d,1}$ generated by reflections in hyperplanes. The fixed hyperplanes of the reflections in a finite-covolume reflection group $\Gamma < O'_{d,1}$ divide $\mathbb{H}^d$ into isometric copies of a single finite-volume convex polyhedron $P \subset \mathbb{H}^d$. The polyhedron P is a *Coxeter polyhedron*, that is, a finite-sided convex polyhedron in which the dihedral angle between any two adjacent facets is an integral submultiple of π . We say P is a *fundamental chamber* for Γ . Conversely, given a finite-volume Coxeter polyhedron $P \subset \mathbb{H}^d$, the group generated by the reflections in all the supporting hyperplanes, or *walls*, of P is a finite-covolume reflection group $\Gamma < O'_{d,1}$ with fundamental chamber P . We thus frequently conflate finite-volume Coxeter polyhedra in $\mathbb{H}^d$ with their corresponding lattices in $O'_{d,1}$ (or their corresponding hyperbolic orbifolds).

Let $H_e = \{x \in \mathbb{H}^d \mid (x, e) = 0\}$ be a hyperplane in $\mathbb{H}^d \subset \mathbb{R}^{d,1}$ whose linear span in $\mathbb{R}^{d,1}$ has normal vector $e \in \mathbb{R}^{d,1}$ with $(e, e) = 1$, and $H_e^- = \{x \in \mathbb{H}^d \mid (x, e) \leq 0\}$ be the half-space associated with it. If

$$P = \bigcap_{j=1}^N H_{e_j}^-$$

is a Coxeter polyhedron in $\mathbb{H}^d$, then the matrix $G(P) = \{g_{ij}\}_{i,j=1}^N = \{(e_i, e_j)\}_{i,j=1}^N$ is its *Gram matrix*. We write $K(P) = \mathbb{Q}(\{g_{ij}\}_{i,j=1}^N)$ and denote by $k(P)$ the field generated by all possible cyclic products of the entries of $G(P)$; we call the field $k(P)$ the *ground field* of P . For convenience, the set of all cyclic products of entries of a given matrix $A = (a_{ij})_{i,j=1}^N$, ie the set of all possible products of the form $a_{i_1 i_2} a_{i_2 i_3} \cdots a_{i_k i_1}$, will be denoted by $\text{Cyc}(A)$. Thus $k(P) = \mathbb{Q}(\text{Cyc}(G(P))) \subset K(P)$.

The following criterion allows us to determine if a given finite-covolume hyperbolic reflection group Γ with fundamental chamber P is arithmetic, quasi-arithmetic, or neither:

Theorem 2.1 (Vinberg’s arithmeticity criterion [34]) *Let $\Gamma < O'_{d,1}$ be a reflection group with finite-volume fundamental chamber $P \subset \mathbb{H}^d$. Then Γ is arithmetic if and only if each of the following conditions holds:*

- (V1) $K(P)$ is a totally real algebraic number field.
- (V2) For any embedding $\sigma: K(P) \rightarrow \mathbb{R}$ such that $\sigma|_{k(P)} \neq \text{Id}$, the matrix $G^\sigma(P)$ is positive semidefinite.
- (V3) $\text{Cyc}(2G(P)) \subset \mathcal{O}_{k(P)}$.

In this case, the ground field of Γ is $k(P)$. The group Γ is quasi-arithmetic if and only if it satisfies conditions (V1)–(V2), but not necessarily (V3), and, in this case, the ground field of Γ is again $k(P)$.

Remark 2.2 Note that $2 \cos(\pi/n)$ is always an algebraic integer. Thus, if there are no dashed edges in the Coxeter–Vinberg diagram of a finite-volume Coxeter polyhedron P , then condition (V3) above automatically holds, and there is no distinction between arithmeticity and quasi-arithmeticity for the associated reflection group Γ . In particular, a triangle group acting on $\mathbb{H}^2$ is quasi-arithmetic precisely when it is arithmetic.

Remark 2.3 Work of Vinberg [35, Section 4] implies that the ground field of a finite-covolume hyperbolic reflection group coincides with its adjoint trace field, and is thus a commensurability invariant.

3 Geometry and arithmetic of Löbell orbifolds

3.1 A decomposition of L_n

For any $n \geq 5$ the Löbell polyhedron L_n admits a decomposition into $2n$ isometric “slices” T_n , each of which may be regarded as a twice-truncated tetrahedron; this decomposition is illustrated in Figure 3 for $n = 6$, where the hyperbolic triangles ABC and $A'B'C'$ are exactly the results of these truncations.

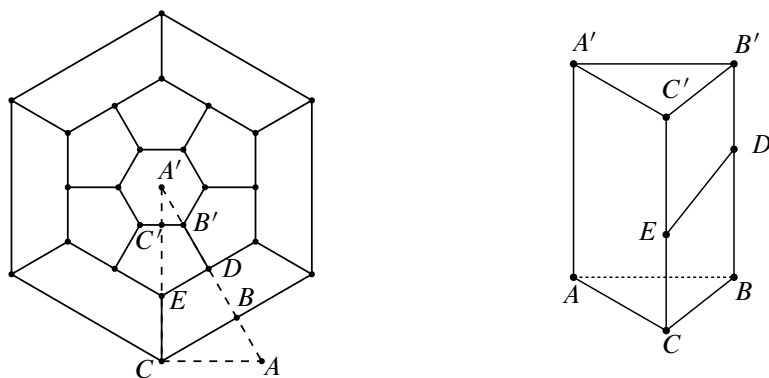


Figure 3: The Löbell polyhedron L_6 and its slice, the truncated tetrahedron T_6 . We use the analogous edge labeling for all the truncated tetrahedra T_n .

The polyhedron T_n is a Coxeter polyhedron whose edges are labeled in Figure 3, right. The Coxeter–Vinberg diagram for T_n is given in Figure 4. The weight d_n in this diagram is equal to $\cosh \delta_n$ since the distance between the top and bottom faces of L_n is the same as that for T_n ; note that δ_n is also equal to the length of the edge AA' .

The Gram matrix of T_n is

$$G_n := G(T_n) = \begin{pmatrix} 1 & -\cos(\pi/n) & 0 & 0 & 0 & -\frac{1}{2}\sqrt{2} \\ -\cos(\pi/n) & 1 & 0 & 0 & -\frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 & -d_n & 0 & -a_n \\ 0 & 0 & -d_n & 1 & -a_n & 0 \\ 0 & -\frac{1}{2}\sqrt{2} & 0 & -a_n & 1 & 0 \\ -\frac{1}{2}\sqrt{2} & 0 & -a_n & 0 & 0 & 1 \end{pmatrix}.$$

We know that the signature² of G_n is $(3, 1, 2)$, since $T_n \subset \mathbb{H}^3$ is compact. Therefore, using the fact $\det G_n$ and all 5×5 principal minors of G_n vanish, we obtain that $d_n^2 = (2a_n^2 - 1)(a_n^2 - 1)$. This allows us to compute d_n and a_n :

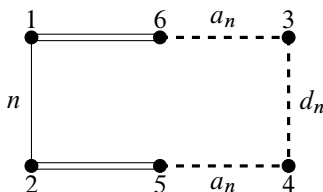
$$d_n = \cosh \delta_n = \frac{\cos(\pi/n)}{\cos(2\pi/n)}, \quad a_n = \sqrt{1 + \frac{1}{2\cos(2\pi/n)}}.$$

3.2 Proof of Theorem 1.4

By Remark 2.3, the adjoint trace field k_n of the lattice Γ_n coincides with the ground field of the polyhedron T_n for $n \geq 5$. Our computations in Section 3.1 allow us to determine the cyclic products $C_n := \text{Cyc}(G_n)$ using the Coxeter–Vinberg diagram shown in Figure 4. We have

$$\mathbb{Q}(C_n) = \mathbb{Q}\left(\left\{\cos(2\pi/n), 1 + \frac{1}{2\cos(2\pi/n)}, \frac{\cos^2(\pi/n)}{\cos^2(2\pi/n)}, \frac{\cos^2(\pi/n)}{\cos(2\pi/n)}\left(1 + \frac{1}{2\cos(2\pi/n)}\right)\right\}\right).$$

²The signature of a real symmetric matrix A is the triple (p, q, r) of numbers of positive, negative, and zero eigenvalues of A , respectively.

Figure 4: The Coxeter–Vinberg diagram of the polytope T_n .

Thus

$$k_n = \mathbb{Q}(C_n) = \mathbb{Q}(\cos(2\pi/n)).$$

Since $\deg(k_n) = \frac{1}{2}\phi(n)$, where ϕ is Euler's totient function, we have that for distinct primes p and q the fields k_p and k_q have different degrees, so that the polyhedra L_p and L_q are not commensurable. $\square$

3.3 Proof of Theorem 1.2

The reflection group Γ_n is commensurable with the group Λ_n generated by reflections in the walls of T_n . It thus remains to check quasi-arithmeticity of Λ_n using Vinberg's arithmeticity criterion (see Theorem 2.1).

The main result in [10] implies in particular that any face of a quasi-arithmetic hyperbolic Coxeter 3-polyhedron that is itself a Coxeter polygon is also quasi-arithmetic with the same ground field. Note that if some face F of a Coxeter 3-polyhedron meets its adjacent faces at even angles, ie angles of the form $\pi/(2m)$ for some $m \geq 1$, then F is a Coxeter polygon. It is shown in [10] that, in the latter case, if P is moreover arithmetic, then F is arithmetic as well.

The polyhedron T_n has a $(2, 4, n)$ -triangular face orthogonal to all adjacent faces. For the $(2, 4, n)$ -triangle group, arithmeticity is equivalent to quasi-arithmeticity (see Remark 2.2). Takeuchi [28] showed that these triangle groups are arithmetic only for $n = 5, 6, 7, 8, 10, 12, 18$. Thus, by the previous paragraph, we have that Λ_n is not quasi-arithmetic for n outside these values (and hence neither is Γ_n). It then suffices to check the conditions of Vinberg's criterion for the Gram matrix of the Coxeter polyhedron T_n for n within these values.

Denote by σ_l the embeddings of the totally real number field $k_n = \mathbb{Q}(\cos(2\pi/n))$, enumerated as follows:

$$\sigma_l(\cos(2\pi/n)) = \cos(2\pi l/n) \quad \text{where } (l, n) = 1 \text{ and } 1 \leq l < n.$$

Note that $\sigma_l(k_n) = \mathbb{Q}(\cos(2\pi l/n))$.

We see from Table 1, which is a result of computations made in Sage, that Λ_n is quasi-arithmetic only for $n = 5, 6, 8, 12$. In order to show that Λ_n is properly quasi-arithmetic if and only if $n = 12$, one needs to check the condition Theorem 2.1(V3), that is, that $\text{Cyc}(2G_n) \subset \mathcal{O}_{k_n}$ if and only if $n = 5, 6, 8$. This can easily be done even without a computer. To avoid being redundant, we record here only the (most interesting) case $n = 12$. Indeed, the other cases were already verified in [5].

n	k_n	d_n	G_n
5	$\mathbb{Q}(\cos(2\pi/5)) = \mathbb{Q}(\sqrt{5})$	2	$G_5^{\sigma_l} \geq 0$ for all $l \neq 1$
6	$\mathbb{Q}$	1	no nonidentity embeddings
7	$\mathbb{Q}(\cos(2\pi/7))$	3	$G_7^{\sigma_2}$ has signature (3, 1)
8	$\mathbb{Q}(\sqrt{2})$	2	$G_8^{\sigma_l} \geq 0$ for all $l \neq 1$
10	$\mathbb{Q}(\cos(\pi/5)) = \mathbb{Q}(\sqrt{5})$	2	$G_{10}^{\sigma_3}$ has signature (3, 1)
12	$\mathbb{Q}(\cos(\pi/6)) = \mathbb{Q}(\sqrt{3})$	2	$G_{12}^{\sigma_l} \geq 0$ for all $l \neq 1$
18	$\mathbb{Q}(\cos(2\pi/18))$	3	$G_{18}^{\sigma_5}$ has signature (3, 1)

Table 1: Properties of the Gram matrices G_n under the nonidentity embeddings σ_l of the fields k_n for $n = 5, 6, 7, 8, 10, 12, 18$.

Lemma 3.1 *The reflection group Λ_{12} is properly quasi-arithmetic.*

Proof Notice that $\cos(\pi/12) = \frac{1}{4}(\sqrt{6} + \sqrt{2})$. The Gram matrix G_{12} of T_{12} is

$$\begin{pmatrix} 1 & -\frac{1}{4}(\sqrt{2}(1+\sqrt{3})) & 0 & 0 & 0 & -\frac{1}{2}\sqrt{2} \\ -\frac{1}{4}(\sqrt{2}(1+\sqrt{3})) & 1 & 0 & 0 & -\frac{1}{2}\sqrt{2} & 0 \\ 0 & 0 & 1 & -\frac{1}{6}(\sqrt{2}(3+\sqrt{3})) & 0 & -\sqrt{\sqrt{3}+3}/\sqrt{3} \\ 0 & 0 & -\frac{1}{6}(\sqrt{2}(3+\sqrt{3})) & 1 & -\sqrt{\sqrt{3}+3}/\sqrt{3} & 0 \\ 0 & -\frac{1}{2}\sqrt{2} & 0 & -\sqrt{\sqrt{3}+3}/\sqrt{3} & 1 & 0 \\ -\frac{1}{2}\sqrt{2} & 0 & -\sqrt{\sqrt{3}+3}/\sqrt{3} & 0 & 0 & 1 \end{pmatrix}.$$

We have that $k_{12} = \mathbb{Q}(\sqrt{3})$ and

$$K_{12} := K(T_{12}) = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{3+\sqrt{3}}) = \mathbb{Q}(\sqrt{2}, \sqrt{3+\sqrt{3}}).$$

One can diagonalize the quadratic form Q_{12} associated with the matrix G_{12} over the field k_{12} as $\text{diag}(1, -\frac{1}{4}\sqrt{3} + \frac{1}{2}, 1, -\frac{1}{3}\sqrt{3} + \frac{1}{3}, 0, 0)$. It is easy to verify that for each embedding $\sigma: K_{12} \rightarrow \mathbb{R}$ such that $\sigma|_{k_{12}} \neq \text{Id}$, the quadratic form Q_{12}^σ has signature (4, 0, 2).

Finally, we observe that not all cyclic products of $2G_{12}$ are algebraic integers. For instance, we have that $(2\sqrt{\sqrt{3}+3}/\sqrt{3})^2 = \frac{4}{3}(3+\sqrt{3}) \notin \mathbb{Z}[\sqrt{3}] = \mathcal{O}_{k_{12}}$. $\square$

3.4 Systoles of Löbell orbifolds

For a lattice $\Gamma < \text{Isom}(\mathbb{H}^d)$, the *systole* $\text{sys}(\Gamma)$ of Γ is the minimal translation length of a loxodromic element of Γ . The *systole* $\text{sys}(M)$ of the complete finite-volume hyperbolic orbifold $M = \mathbb{H}^d/\Gamma$ is simply the systole of the lattice Γ . We record in this section a couple of remarks about systoles of hyperbolic reflection orbifolds.

Remark 3.2 Suppose we have a sequence of reflection groups $\Gamma_n < \text{Isom}(\mathbb{H}^d)$ for $d \geq 2$ with the property that $\text{sys}(\Gamma_n) \rightarrow 0$. Then Γ_n is arithmetic for at most finitely many n . Indeed, suppose otherwise,

so that we may assume the Γ_n are all arithmetic. Since there are only finitely many maximal arithmetic reflection groups (see [2; 12; 24]), up to further extraction, we may also assume the Γ_n are all contained in a single lattice $\Gamma < \text{Isom}(\mathbb{H}^d)$, so that $\text{sys}(\Gamma_n) \geq \text{sys}(\Gamma)$, a contradiction.

Remark 3.3 It follows from Remark 3.2 that if $\Gamma_n < \text{Isom}(\mathbb{H}^d)$ for $d \geq 2$ is a sequence of finite-covolume reflection groups satisfying $\text{sys}(\Gamma_n) \rightarrow 0$, then for each $m \in \mathbb{N}$, we have that Γ_m is commensurable to Γ_n for at most finitely many n . Indeed, suppose otherwise. Then, up to passing to a subsequence, we may assume that the Γ_n are all commensurable. Since the Γ_n are nonarithmetic by Remark 3.2, it follows from a result of Margulis [20, Theorem 1, page 2] that their commensurator $\Lambda < \text{Isom}(\mathbb{H}^d)$ contains each Γ_n as a finite-index subgroup, so that $\text{sys}(\Gamma_n) \geq \text{sys}(\Lambda)$, a contradiction.

As observed in [7, Sections 5.2 and 5.3], the above conclusion in fact holds for any sequence of lattices $\Gamma_n < \text{Isom}(\mathbb{H}^d)$ satisfying $\text{sys}(\Gamma_n) \rightarrow 0$. Indeed, suppose one has such a sequence Γ_n where the Γ_n are all commensurable. If the Γ_n are nonarithmetic, then one obtains a contradiction as in the previous paragraph. If the Γ_n are arithmetic, then since they are commensurable, a uniform lower bound for $\text{sys}(\Gamma_n)$ is provided by the fact that, given $d \in \mathbb{N}$ and $\mu > 1$, there are only finitely many monic integer polynomials of degree d and Mahler measure at most μ . We remark that the strongest form of Lehmer's conjecture would imply a uniform lower bound on the systole of any arithmetic locally symmetric orbifold; see the discussion immediately following Conjecture 10.2 by Gelfander [13].

4 Löbell orbifolds and hyperbolic Dehn fillings of ideal right-angled antiprisms

Let $P \subset \mathbb{H}^3$ be a finite-volume Coxeter polyhedron with a compact edge e , and say the dihedral angle at e is π/m . Following Kolpakov [17], if $Q \subset \mathbb{H}^3$ is a finite-volume Coxeter polyhedron of the same combinatorial type and with the same dihedral angles as P except that the dihedral angle at e in Q is diminished to π/n for some $n \geq m$, we say that Q is obtained from P via a (π/n) -contraction at e . For example, for $n \geq 5$ and $k \geq 2$, the polyhedron $P_{n,k} \subset \mathbb{H}^3$, whose Coxeter–Vinberg diagram is shown in Figure 5, is obtained from the truncated tetrahedron T_n via $(\pi/(2k))$ -contractions at the (analogues of the) edges $B'D$ and EC , while T_n is in turn obtained from T_5 via a (π/n) -contraction at the edge AA' ; see Figures 3 and 4.

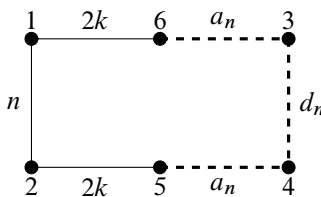
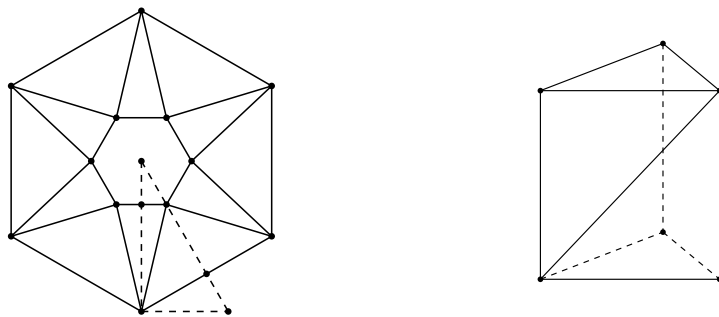


Figure 5: The Coxeter–Vinberg diagram of the polyhedron $P_{n,k}$.

Figure 6: The antiprism A_6 and its slice R_6 .

If instead Q and P differ (as labeled polyhedra) only in that, in Q , the edge e is replaced by an ideal vertex $v \in \partial\mathbb{H}^3$, then we say that Q is obtained from P by *contracting e to an ideal vertex*. Note that, if such a contraction exists, the dihedral angle at each edge adjacent to e in P (and each edge incident to v in Q) is $\frac{1}{2}\pi$, and the faces sharing e in P are at least 4-sided.

In [16], Kellerhals studies a family of ideal right-angled polyhedra known as the antiprisms $A_n \subset \mathbb{H}^3$ for $n \geq 3$, where she exploits a decomposition of each such polyhedron A_n into $2n$ copies of a polyhedron R_n , analogous to the decomposition of the L\"obell polyhedron L_n into $2n$ copies of T_n ; see Figure 6. In the above language, for $n \geq 5$, the polyhedron R_n is in fact obtained from T_n by contracting the edges $B'D$ and EC of T_n to ideal vertices. In particular, for such n , the antiprism A_n may be obtained from the L\"obell polyhedron L_n by a sequence of edge contractions to ideal vertices. This was already observed by Kolpakov [17, Section 5.1]. Indeed, the existence of the polyhedra $P_{n,k}$ and R_n for $n \geq 5$ and $k \geq 2$ is predicted by the following result of Kolpakov, whose proof rests on Andreev's theorem [3; 27]:

Theorem 4.1 (see the proof of [17, Proposition 1]) *Let $P \subset \mathbb{H}^3$ be a finite-volume Coxeter polyhedron with at least five faces and a compact edge e such that the dihedral angle at each edge adjacent to e is $\frac{1}{2}\pi$, and the faces sharing e in P are at least 4-sided. If the dihedral angle at e in P is π/m for $m \geq 2$, then P admits a (π/n) -contraction P_n at e for each $n \geq m$, as well as a contraction P_∞ of e to an ideal vertex.*

Since we may instead view the P_n as having been obtained from P_∞ via Dehn filling,³ it follows from work of Dunbar and Meyerhoff [11] that the length of the (analogue of the) edge $e \subset P_n$ approaches 0 as $n \rightarrow \infty$. Taking P_n to be the truncated tetrahedron T_n and e the edge AA' , we obtain another justification for Observation 1.1. Applying a similar argument to the R_n instead of the T_n , one concludes that the distance between the “top” and “bottom” faces of the antiprism A_n also approaches 0 as $n \rightarrow \infty$; for a justification of the latter that uses only elementary Euclidean geometry, see Figure 7.

We remark that the conditions of Theorem 4.1 are satisfied for any compact edge e of a finite-volume right-angled polyhedron $P \subset \mathbb{H}^3$. This suggests a method of constructing a family of new finite-volume

³More precisely, Dunbar and Meyerhoff adapt Thurston's theory of hyperbolic Dehn fillings to the setting of *oriented* hyperbolic 3-orbifolds. Applying this theory to the orientation covers of the P_n , where the P_n are viewed as reflection orbifolds, one concludes that the length of the edge e in P_n approaches 0 as $n \rightarrow \infty$.

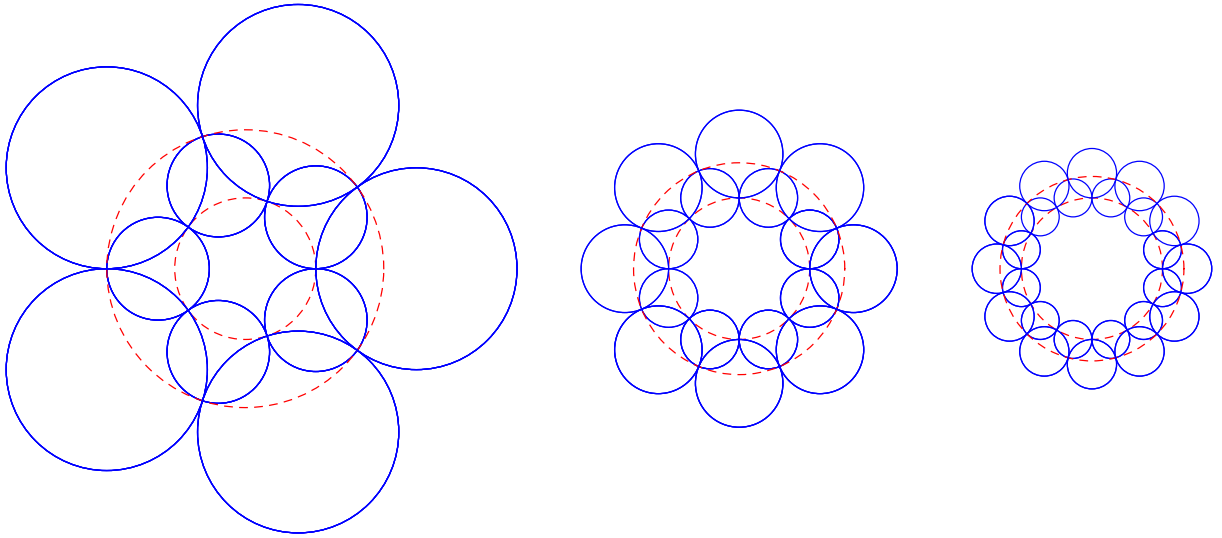


Figure 7: A visual proof that the distance between the “top” and “bottom” faces of the antiprism A_n approaches 0. Drawn above are the ideal boundaries of the walls of A_n for $n = 5, 8$, and 12 (from left to right), visualized via stereographic projection onto the page from the ideal boundary of $\mathbb{H}^3$ such that the circles corresponding to the top and bottom faces (the dashed circles) are concentric. Keeping fixed the Euclidean diameter of the inner dashed circle, as n approaches ∞ the diameter of an inner solid circle approaches 0, so that the ratio between the diameters of the inner and outer dashed circles approaches 1. In fact, an elementary exercise in Euclidean geometry shows that this ratio is precisely $(1 + \sin(\pi/n))/\cos(\pi/n)$, corroborating a computation of Kellerhals [16, Examples 1 and 2].

right-angled polyhedra starting from the polyhedron P . Namely, for $n \geq 2$, let $P_n \subset \mathbb{H}^3$ be the polyhedron obtained from P by a (π/n) -contraction at e , let $\Lambda_n < \text{Isom}(\mathbb{H}^3)$ be the associated reflection group, and let Δ_n be the stabilizer of e in Λ_n . Then the finite-volume polyhedron $Q_n = \bigcup_{\gamma \in \Delta_n} \gamma P_n$ is right-angled, and is compact if and only if P was. Moreover, for each $m \geq 2$, we have that Λ_m is commensurable to Λ_n for at most finitely many $n \geq 2$, for instance, because $\text{sys}(\Lambda_n) \rightarrow 0$ as $n \rightarrow \infty$ (see Remark 3.3). We conclude our discussion by observing that this strategy for producing an infinite family of pairwise incommensurable finite-volume right-angled hyperbolic polyhedra fails in dimension 4, where the existence of such a family appears to be open. The observation is essentially that codimension-2 contractions in dimensions higher than 3 yield orbifolds à la Gromov and Thurston [15].

Theorem 4.2 *Let $n, m \geq 3$ be of the same parity and $d \geq 4$. Suppose $P_n \subset \mathbb{H}^d$ is a finite-volume Coxeter polyhedron all of whose dihedral angles are right angles except for a single dihedral angle of π/n . Suppose there is also a finite-volume Coxeter polyhedron $P_m \subset \mathbb{H}^d$ with the same combinatorics and dihedral angles as P_n except that the exceptional dihedral angle of P_m is π/m . Then $n = m$.*

We will make use of the following lemma:

Lemma 4.3 Let P be a finite-volume Coxeter polyhedron in $\mathbb{H}^d$ for $d \geq 4$, and suppose all dihedral angles of P are right angles except possibly for one dihedral angle formed by walls H_1 and H_2 of P . Then the group $\Gamma < \text{Isom}(\mathbb{H}^d)$ generated by the reflections in all walls of P except H_1 and H_2 is Zariski-dense in $\text{Isom}(\mathbb{H}^d)$.

Proof of Lemma 4.3 Let $P' \subset \mathbb{H}^d$ be the (infinite-volume) polyhedron obtained from P by forgetting the walls H_1 and H_2 , and let P_i be the intersection of P' with the hyperplane H_i of $\mathbb{H}^d$ for $i = 1, 2$. Then the $(d-1)$ -dimensional right-angled hyperbolic polyhedron P_i is obtained from a finite-volume such polyhedron — namely, the polyhedron $P \cap H_i$ — by forgetting a single wall — namely, the intersection $H_1 \cap H_2$. Since $d-1 \geq 3$, it follows that the subgroup of $\text{Isom}(H_i)$ generated by the reflections in the walls of P_i is Zariski-dense in $\text{Isom}(H_i)$; see, for instance, [14, Section 1.7]. $\square$

Proof of Theorem 4.2 For $k = n, m$, let H_k and H'_k be the walls of P_k forming the exceptional dihedral angle of π/k , and let R_k be the union of the images of P_k under the reflection group D_k generated by the reflections in H_k and H'_k . Since n and m have the same parity, we may choose reflections $r_k \in D_k$ such that the $(d-1)$ -dimensional polyhedra $\text{Fix}(r_n) \cap R_n$ and $\text{Fix}(r_m) \cap R_m$ have the same combinatorics and dihedral angles and are thus isometric by Mostow–Prasad rigidity [23; 26]; see also Andreev [3; 4].

Now interbreed the R_k along $\text{Fix}(r_k) \cap R_k$ and let R be the resulting finite-volume polyhedron in $\mathbb{H}^d$. Then there is an obvious dihedral group $D_{(n+m)/2}$ of combinatorial symmetries of R preserving dihedral angles. By Mostow–Prasad rigidity, each of these combinatorial symmetries is a hyperbolic isometry. However, by Lemma 4.3, there is (up to inverses) a unique hyperbolic isometry γ_k rotating the surface of any given slice of R_k two slices over, namely, the composition of the reflections in H_k and H'_k , so that γ_k has order k . Since γ_n and γ_m each generate the commutator subgroup of $D_{(n+m)/2}$, it follows that $n = m$. $\square$

References

- [1] I Agol, *Systoles of hyperbolic 4-manifolds*, preprint (2006) [arXiv math/0612290](#)
- [2] I Agol, M Belolipetsky, P Storm, K Whyte, *Finiteness of arithmetic hyperbolic reflection groups*, Groups Geom. Dyn. 2 (2008) 481–498 [MR](#) [Zbl](#)
- [3] E M Andreev, *On convex polyhedra in Lobachevsky spaces*, Mat. Sb. 81 (1970) 445–478 [MR](#) [Zbl](#) In Russian; translated in Math. USSR-Sb. 10 (1970) 413–440
- [4] E M Andreev, *On convex polyhedra of finite volume in Lobachevsky space*, Mat. Sb. 83 (1970) 256–260 [MR](#) [Zbl](#) In Russian; translated in Math. USSR-Sb. 12 (1970) 255–259
- [5] O Antolín-Camarena, G R Maloney, R K W Roeder, *Computing arithmetic invariants for hyperbolic reflection groups*, from “Complex dynamics” (D Schleicher, editor), A K Peters, Wellesley, MA (2009) 597–631 [MR](#) [Zbl](#)
- [6] A Basmajian, *Tubular neighborhoods of totally geodesic hypersurfaces in hyperbolic manifolds*, Invent. Math. 117 (1994) 207–225 [MR](#) [Zbl](#)

- [7] **M V Belolipetsky, S A Thomson**, *Systoles of hyperbolic manifolds*, Algebr. Geom. Topol. 11 (2011) 1455–1469 [MR](#) [Zbl](#)
- [8] **R Benedetti, C Petronio**, *Lectures on hyperbolic geometry*, Springer (1992) [MR](#) [Zbl](#)
- [9] **N Bergeron, F Haglund, D T Wise**, *Hyperplane sections in arithmetic hyperbolic manifolds*, J. Lond. Math. Soc. 83 (2011) 431–448 [MR](#) [Zbl](#)
- [10] **N Bogachev, A Kolpakov**, *On faces of quasi-arithmetic Coxeter polytopes*, Int. Math. Res. Not. 2021 (2021) 3078–3096 [MR](#) [Zbl](#)
- [11] **W D Dunbar, G R Meyerhoff**, *Volumes of hyperbolic 3-orbifolds*, Indiana Univ. Math. J. 43 (1994) 611–637 [MR](#) [Zbl](#)
- [12] **D Fisher, S Hurtado**, *A new proof of finiteness of maximal arithmetic reflection groups*, Ann. H. Lebesgue 6 (2023) 151–159 [MR](#) [Zbl](#)
- [13] **T Gelander**, *Homotopy type and volume of locally symmetric manifolds*, Duke Math. J. 124 (2004) 459–515 [MR](#) [Zbl](#)
- [14] **M Gromov, I Piatetski-Shapiro**, *Nonarithmetic groups in Lobachevsky spaces*, Inst. Hautes Études Sci. Publ. Math. 66 (1988) 93–103 [MR](#) [Zbl](#)
- [15] **M Gromov, W Thurston**, *Pinching constants for hyperbolic manifolds*, Invent. Math. 89 (1987) 1–12 [MR](#) [Zbl](#)
- [16] **R Kellerhals**, *A polyhedral approach to the arithmetic and geometry of hyperbolic link complements*, J. Knot Theory Ramifications 32 (2023) art.id. 2350052 [MR](#) [Zbl](#)
- [17] **A Kolpakov**, *Deformation of finite-volume hyperbolic Coxeter polyhedra, limiting growth rates and Pisot numbers*, European J. Combin. 33 (2012) 1709–1724 [MR](#) [Zbl](#)
- [18] **F Löbell**, *Beispiele geschlossener dreidimensionaler Clifford–Kleinscher Räume negativer Krümmung*, Ber. Verh. Sächs Akad. Wiss. Leipzig Math.-Phys. Kl. 83 (1931) 167–174 [Zbl](#)
- [19] **C Maclachlan, A W Reid**, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Math. 219, Springer (2003) [MR](#) [Zbl](#)
- [20] **G A Margulis**, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Math. 17, Springer (1991) [MR](#) [Zbl](#)
- [21] **A D Mednykh, A Y Vesnin**, *Löbell manifolds revised*, Sib. Elektron. Mat. Izv. 4 (2007) 605–609 [MR](#) [Zbl](#)
- [22] **J S Meyer, C Millichap, R Trapp**, *Arithmeticity and hidden symmetries of fully augmented pretzel link complements*, New York J. Math. 26 (2020) 149–183 [MR](#) [Zbl](#)
- [23] **G D Mostow**, *Quasi-conformal mappings in n -space and the rigidity of hyperbolic space forms*, Inst. Hautes Études Sci. Publ. Math. 34 (1968) 53–104 [MR](#) [Zbl](#)
- [24] **V V Nikulin**, *Finiteness of the number of arithmetic groups generated by reflections in Lobachevsky spaces*, Izv. Ross. Akad. Nauk Ser. Mat. 71 (2007) 55–60 [MR](#) [Zbl](#) In Russian; translated in *Izv. Math.* 71 (2007) 53–56
- [25] **M Phillips, S Levy, T Munzner**, *Geomview* Available at <http://www.geomview.org>
- [26] **G Prasad**, *Strong rigidity of $\mathbf{Q}$ -rank 1 lattices*, Invent. Math. 21 (1973) 255–286 [MR](#) [Zbl](#)
- [27] **R K W Roeder, J H Hubbard, W D Dunbar**, *Andreev’s theorem on hyperbolic polyhedra*, Ann. Inst. Fourier (Grenoble) 57 (2007) 825–882 [MR](#) [Zbl](#)

- [28] **K Takeuchi**, *Arithmetic triangle groups*, J. Math. Soc. Japan 29 (1977) 91–106 [MR](#) [Zbl](#)
- [29] **S Thomson**, *Quasi-arithmeticity of lattices in $PO(n, 1)$* , Geom. Dedicata 180 (2016) 85–94 [MR](#) [Zbl](#)
- [30] **A Y Vesnin**, *Three-dimensional hyperbolic manifolds of Löbell type*, Sibirsk. Mat. Zh. 28 (1987) 50–53 [MR](#) [Zbl](#)
- [31] **A Y Vesnin**, *Three-dimensional hyperbolic manifolds with a general fundamental polyhedron*, Mat. Zametki 49 (1991) 29–32, 157 [MR](#) [Zbl](#) In Russian; translated in *Math. Notes* 49 (1991) 575–577
- [32] **A Y Vesnin**, *Volumes of Löbell 3-manifolds*, Mat. Zametki 64 (1998) 17–23 [MR](#) [Zbl](#) In Russian; translated in *Math. Notes* 64 (1998) 15–19
- [33] **A Y Vesnin**, *Right-angled polytopes and three-dimensional hyperbolic manifolds*, Uspekhi Mat. Nauk 72 (2017) 147–190 [MR](#) [Zbl](#) In Russian; translated in *Russian Math. Surv.* 72 (2017) 335–374
- [34] **E B Vinberg**, *Discrete groups generated by reflections in Lobachevsky spaces*, Mat. Sb. 72 (1967) 471–488 [MR](#) [Zbl](#) Correction in 73 (1967) 303–303 In Russian; translated in *Math. USSR-Sb.* 1 (1967) 429–444
- [35] **E B Vinberg**, *Rings of definition of dense subgroups of semisimple linear groups*, Izv. Akad. Nauk SSSR Ser. Mat. 35 (1971) 45–55 [MR](#) [Zbl](#) In Russian; translated in *Math. USSR-Izv.* 5 (1971) 45–55

NB: Department of Computer and Mathematical Sciences, University of Toronto Scarborough
Toronto, ON, Canada

Institute for Information Transmission Problems
Moscow, Russia

SD: Institut des Hautes Études Scientifiques
Bures-sur-Yvette, France

n.bogachev@utoronto.ca, douba@ihes.fr

Received: 19 July 2023

Formality of sphere bundles

JIAWEI ZHOU

We study the formality of orientable sphere bundles over connected compact manifolds. When the base manifold is formal, we prove that the formality of the bundle is equivalent to the vanishing of the Bianchi–Massey tensor introduced by Crowley and Nordström. As an example, this implies that the unit tangent bundle over a formal manifold can only be formal when the base manifold has vanishing Euler characteristic or a rational cohomology ring generated by one element. When the base manifold is not formal, we give an obstruction to the formality of sphere bundles whose Euler class is reducible.

55P62, 55R25

1 Introduction

In rational homotopy theory, the equivalence relation between simply connected topological spaces is defined by a continuous map $f: X \rightarrow Y$ inducing an isomorphism $\pi_*(X) \otimes \mathbb{Q} \rightarrow \pi_*(Y) \otimes \mathbb{Q}$ between the homotopy groups of their rationalizations. This condition on f is equivalent to $f^*: H^*(Y, \mathbb{Q}) \rightarrow H^*(X, \mathbb{Q})$ being an isomorphism. Based on the latter, Sullivan [10] used a commutative differential graded algebra (CDGA) as a model of these spaces. Such CDGAs are composed of polynomial differential forms, or just differential forms if the space is a smooth manifold. The equivalence between CDGAs are also defined by quasiisomorphisms, whose induced maps on cohomologies are isomorphisms. This equivalence relation on CDGA models also turn out to provide a classification of nonsimply connected spaces.

A CDGA is called formal if it is equivalent to its cohomology, and a topological space is called formal if its CDGA model is formal. In short, the rational homotopy type of a formal space can be described just by its cohomology ring, a much simpler structure.

Several examples of formal spaces stand out; for example, H-spaces, symmetric spaces, products of formal spaces, and k -connected compact manifolds whose total dimension does not exceed $4k + 2$ (Miller [9]). Additionally, compact Kähler manifolds are formal, as was first shown by Deligne, Griffiths, Morgan and Sullivan [5]. In general, a compact complex manifold is formal if it satisfies the dd^c -lemma. On the other hand, the dd^A -lemma, an analogous condition for symplectic manifolds which is equivalent to satisfy the hard Lefschetz property, does not imply formality, even if the manifold is simply connected (Cavalcanti [3]).

For general spaces, Sullivan [10] (see also [5]) used a special class of algebras to classify them. It consists of free graded algebras, each with a well-ordered set of generators. The differential of each generator belongs

to the subalgebra generated by elements with smaller indices. Such algebras are referred to as the Sullivan algebras. Every connected CDGA is equivalent to a Sullivan algebra uniquely up to quasiisomorphism, and a special type of Sullivan algebra uniquely up to isomorphism, which is called the minimal Sullivan algebra. If the CDGA is a model of some space, its equivalent Sullivan algebra and minimal Sullivan algebra are called the Sullivan model and the minimal Sullivan model of this space, respectively.

Adding a generator of odd degree k to the Sullivan model of a manifold is equivalent to constructing an orientable S^k -bundle over it. So we are interested in investigating how this process affects formality, and potentially applying this knowledge to study the formality for general fiber bundles. In this paper, manifolds are generally assumed to be smooth, connected and compact. Sphere bundles are assumed to be orientable, and moreover, the cohomologies are assumed to be over a field $\mathbb{K}$ of characteristic 0. Most of our proofs should also be applicable for general topological manifolds and spherical Serre fibrations since they do not rely on the smooth structure or the bundle structure.

The relationship between the formalities of the base and the total space was studied by Lupton [8] and Amann and Kapovitch [1]. For a fibration of simply connected spaces of finite type, if the fiber F is formal, rationally elliptic (ie both $\pi_*(F) \otimes \mathbb{Q}$ and $H^*(F, \mathbb{Q})$ are finite dimensional), and also satisfies the Halperin conjecture (ie that for any fibration $F \rightarrow E \rightarrow B$ with the base B being simply connected, the induced morphism $H^*(E, \mathbb{Q}) \rightarrow H^*(F, \mathbb{Q})$ is surjective), then the base is formal if and only if the total space is formal. This condition holds for bundles with even-dimensional sphere fibers.

Both Lupton's and the Amann–Kapovitch proof rely on the fact that when the fiber satisfies the Halperin conjecture, the cohomology ring of the total space can be expressed as a tensor product of the cohomologies of the base and that of the fiber. Since the fiber being an odd dimensional sphere does not satisfy this condition, the cohomology of the total space becomes more intricate. Actually, there are simple examples of a nonformal bundle over a formal manifold with the fiber being an odd-dimensional sphere. Such is the case of an orientable circle bundle over a torus with a nontrivial Euler class.

For circle bundles which adds an additional generator from the fiber circle, the breakdown of the formality of the total space is not drastic. In the case of a Boothby–Wang fibration, which is a circle bundle over a symplectic manifold with Euler class being the symplectic structure, Biswas, Fernández, Muñoz and Tralle [2, Proposition 4.5] proved that the higher order (greater than 3) Massey products of the total space vanish, if the base symplectic manifold is formal and satisfying the hard Lefschetz property. In an earlier work [12], we also showed that a sphere bundle over a formal manifold has an A_∞ -minimal model where the only nontrivial operations are m_2 and m_3 .

Regarding A_∞ -algebra, it is said to be formal if it has an A_∞ -minimal model with only m_2 nontrivial. The information of m_3 is encoded in the Bianchi–Massey tensor introduced by Crowley and Nordström [4], which is a linear map from a subspace of $(H^*)^{\otimes 4}$ to H^* . More precisely, a compact manifold whose Bianchi–Massey tensor vanishes has an A_∞ -minimal model with $m_3 = 0$. Therefore, it is natural to conjecture that a sphere bundle over a formal manifold is formal if its Bianchi–Massey tensor vanishes.

Unlike the possible different representatives of A_∞ -minimal models and Massey products, the Bianchi–Massey tensor is uniquely defined without ambiguity. So once we prove the above conjecture in [Section 3](#), it follows that the formality of a sphere bundle over a formal manifold can be determined by finite calculation. It turns out that we can construct a CDGA equivalence between the differential forms of the sphere bundle and its cohomology, and our proof will not use the language of A_∞ -algebra.

Theorem 1.1 *Suppose M is a compact formal manifold, and $\pi: X \rightarrow M$ is an orientable S^k -bundle. Then X is formal if and only if the Bianchi–Massey tensor of $\Omega^*(X)$ vanishes. Moreover, when k is even, X is always formal.*

It turns out that a trivial Euler class is a sufficient condition for the formality of a sphere bundle over a formal manifold. For nontrivial Euler classes, we can consider the special case that the Euler class is of top degree. This requires the manifold to be even-dimensional. We prove that such bundles are formal only when the rational cohomology rings of the base manifolds are generated by one element. This is a consequence of [Theorem 3.10](#).

Theorem 1.2 *Suppose that M is a $2k$ -dimensional compact orientable formal manifold. Let X be an S^{2k-1} -bundle with nontrivial Euler class. If X is formal, then $H^*(M) = \mathbb{K}[x]/(x^p)$ is a quotient of the polynomial ring with a single variable.*

It immediately follows that if the unit tangent bundle of a formal manifold M is formal, then either the Euler characteristic $\chi(M) = 0$, or $H^*(M)$ is generated by one element.

It is also interesting to consider the more general case when the base M is not necessarily formal. As $\Omega^*(M)$ may no longer be represented by the cohomology, we shall consider its minimal Sullivan model, labeled by $\mathcal{M}$, instead. Depending on the reducibility of the representative of the Euler class in $\mathcal{M}$, the minimal Sullivan model of the sphere bundles can have two types. In this paper, we mainly consider the reducible case, and give an obstruction to formality under certain conditions.

Theorem 1.3 *Let (M, ω) be a connected symplectic manifold satisfying the hard Lefschetz property. Suppose $[\omega]$ is an integral and reducible cohomology class, ie there exists some $x_i, y_i \in H^1(M)$ such that $[\omega] = \sum [x_i] \wedge [y_i]$. Then the Boothby–Wang fibration of M , ie the circle bundle with Euler class $[\omega]$, is nonformal.*

This obstruction can be generalized to S^{4k+1} -bundles, as long as the Euler class $[\omega]$ can be written as a sum of products of $(2k+1)$ -cohomology classes, and the hard Lefschetz property can be weakened in the following way: there exists some $s \geq 0$ such that taking the product with $[\omega]$ is an isomorphism from a nontrivial space $H^s(M)$ to $H^{s+4k+2}(M)$, and also, is injective from $H^{s-2k-1}(M)$ to $H^{s+2k+1}(M)$. The condition of the sphere dimension being $4k+1$ is difficult to relax. We will give a simple example of a formal S^{4k+3} -bundle satisfying other requirements.

This paper is organized as follows. In [Section 2](#), we review the algebraic tools that play a central role in this paper, including the Bianchi–Massey tensor and the minimal Sullivan model. We give the proof of [Theorems 1.1 and 1.2](#) in [Section 3](#). In [Section 4](#), we discuss the formality of general sphere bundles, and give the proof of [Theorem 1.3](#).

Acknowledgements The author thanks Ruizhi Huang, Si Li, Jianfeng Lin, Li-Sheng Tseng and Jie Wu for helpful discussions and valuable suggestions. The author is also grateful to the referee for carefully reading the paper and giving many helpful suggestions for improvement, especially for pointing out and resolving an issue with the proof of [Lemma 3.3](#), and suggesting a generalization of [Lemma 3.4](#). The author would like to acknowledge the support of the National Key Research and Development Program of China (number 2020YFA0713000).

2 Preliminary

2.1 Bianchi–Massey tensor

Let V be a graded vector space over a field $\mathbb{K}$ of characteristic 0. We let $\mathcal{G}^k V$ denote the k^{th} graded symmetric power of V , ie the quotient space of $V^{\otimes k}$ by the relations of graded commutativity. We will use $(x_1 \cdot x_2)$ denote the graded symmetric product of x_1 and x_2 . Hence, $(x_1 \cdot x_2) = (-1)^{|x_1||x_2|}(x_2 \cdot x_1)$, where $|x_1|, |x_2|$ are the degree of x_1 and x_2 , respectively. We can define $(x_1 \cdot x_2 \cdots x_k)$ similarly, and such elements generate $\mathcal{G}^k V$ when all $x_i \in V$.

Remark 2.1 In our setting, $\mathcal{G}^k V$ is isomorphic to the space of graded commutative k -tensors of V , although this is not true if the characteristic of the ground field is nonzero or V is replaced by an abelian group.

Let $K[\bullet]$ denote the kernel of a tensor space under full graded symmetrization. For example, if V is a graded vector space, $K[\mathcal{G}^2 \mathcal{G}^2 V]$ is the kernel of the symmetrization

$$\mathcal{G}^2 \mathcal{G}^2 V \rightarrow \mathcal{G}^4 V, \quad ((x \cdot y) \cdot (z \cdot w)) \mapsto (x \cdot y \cdot z \cdot w).$$

Thus, $(x \cdot y)(z \cdot w) - (-1)^{|y||z|}(x \cdot z)(y \cdot w) \in K[\mathcal{G}^2 \mathcal{G}^2 V]$.

Now suppose $\mathcal{A}$ is a CDGA (commutative differential graded algebra). Let

$$c: \mathcal{G}^2 H^*(\mathcal{A}) \rightarrow H^*(\mathcal{A}), \quad (x \cdot y) \mapsto xy$$

denote the product map, and $E^*(\mathcal{A}) = \ker c$, which we will simply write as E^* . Then we set

$$\mathcal{B}^*(H^*(\mathcal{A})) = \mathcal{G}^2 E^* \cap K[\mathcal{G}^2 \mathcal{G}^2 H^*(\mathcal{A})].$$

For simplicity, we will use $\mathcal{B}^*(\mathcal{A})$ to denote $\mathcal{B}^*(H^*(\mathcal{A}))$.

Let $\mathcal{Z}^*(\mathcal{A}) \subset \mathcal{A}^*$ be the subspace of d -closed elements. Pick a right inverse $\alpha: H^*(\mathcal{A}) \rightarrow \mathcal{Z}^*(\mathcal{A})$ for the projection to cohomology. Then the map

$$\alpha^2: \mathcal{G}^2 H^*(\mathcal{A}) \rightarrow \mathcal{A}^*, \quad (x \cdot y) \mapsto \alpha(x)\alpha(y)$$

takes exact values on E^* . So there exists a linear map $\gamma: E^* \rightarrow \mathcal{A}^{*-1}$ satisfying $d\gamma = \alpha^2$.

Definition 2.2 The map

$$\mathcal{G}^2 E^* \rightarrow \mathcal{A}^{*-1}, \quad (e \cdot e') \mapsto \gamma(e)\alpha^2(e') + (-1)^{|e||e'|}\gamma(e')\alpha^2(e)$$

takes closed values on $\mathcal{B}^*(\mathcal{A})$. So it induces a map

$$\mathcal{F}: \mathcal{B}^*(\mathcal{A}) \rightarrow H^{*-1}(\mathcal{A}),$$

which is called the *Bianchi–Massey tensor*. This map is independent of the choices of α and γ .

Another obstruction to formality is the uniform Massey triple product. It is defined as follows.

Definition 2.3 Given choices of α and γ as before, the map

$$\gamma\alpha: E^* \otimes H^*(\mathcal{A}) \rightarrow \mathcal{A}^{*-1}, \quad e \otimes x \mapsto \gamma(e)\alpha(x)$$

takes closed values on $K[E^* \otimes H^*(\mathcal{A})]$, which is the kernel of the full graded symmetrization

$$E^* \otimes H^*(\mathcal{A}) \rightarrow \mathcal{G}^3 H^*(\mathcal{A}).$$

So it induces a map

$$\mathcal{T}: K[E^* \otimes H^*(\mathcal{A})] \rightarrow H^{*-1}(\mathcal{A}),$$

which is called the *uniform Massey triple product*.

Unlike the Bianchi–Massey tensor, the uniform Massey triple product does depend on the choices of α and γ .

2.2 Minimal Sullivan model

Definition 2.4 A *Sullivan algebra* is a CDGA $\mathcal{M} = \Lambda V^*$ which is free as a graded algebra. V^* has a homogeneous basis $\{v_\alpha\}$ indexed by a well-ordered set, such that the degree $|v_\alpha| \geq 1$ and $dv_\alpha \in \Lambda V_{<\alpha}^*$. Here $V_{<\alpha}^*$ is spanned by v_β with $\beta < \alpha$.

If a Sullivan algebra also satisfies $\beta < \alpha$ whenever $|v_\beta| < |v_\alpha|$, it is called *minimal*.

In a minimal Sullivan algebra $\mathcal{M} = \Lambda V^*$, all dv_α are *reducible*, ie there exist $x_i, y_i \in \mathcal{M}^+$ such that $dv_\alpha = \sum x_i y_i$. Here $\mathcal{M}^+$ is the subspace of $\mathcal{M}$ spanned by elements with degree at least 1. Equivalently, we can say that all dv_α are in $\Lambda^{\geq 2} V$, the subspace generated by elements of wordlength at least 2.

Definition 2.5 If a CDGA $\mathcal{A}$ is quasiisomorphic to a (minimal) Sullivan algebra $\mathcal{M}$, we call $\mathcal{M}$ a (minimal) Sullivan model of $\mathcal{A}$. For a manifold M , say $\mathcal{M}$ is a (minimal) Sullivan model of M if it is a (minimal) Sullivan model of $\Omega^*(M)$.

Every simply connected manifold has a minimal Sullivan model generated by a graded vector space V^* of finite type, ie the subspace V^k of degree k is finite dimensional for all k . Moreover, the degrees of all elements in V^* are at least 2 [5]. More generally, every connected CDGA has a minimal Sullivan model generated by V^* , and the degrees of all elements in V^* are at least 1 [7]. The minimal Sullivan model is unique up to isomorphism.

Theorem 2.6 (Deligne, Griffiths, Morgan and Sullivan [5]) *A minimal Sullivan algebra $\mathcal{M}$ is formal if and only if the following statement holds. $\mathcal{M}$ is generated by $V^* = C^* \oplus N^*$, where C^* is the subspace of closed elements in V^* and N^* is a direct complement of C^* , and all the closed elements in the ideal $I(N^*)$ generated by N^* are exact.*

3 Formality of sphere bundles over formal spaces

3.1 The Bianchi–Massey tensor determines formality

Let $\omega \in \mathcal{A}^k$ be a closed element in a CDGA $\mathcal{A}$. Suppose θ has degree $k-1$, and satisfies $\theta^2 = \theta\theta = 0$ (this trivially holds when $k-1$ is odd). We can extend $\mathcal{A}$ by θ :

$$\mathcal{A}_\omega = \mathcal{A} \otimes \Lambda\theta = \{x + \theta y \mid x, y \in \mathcal{A}\},$$

with $d\theta = \omega$. Here $\Lambda\theta = \langle 1, \theta \rangle$ is an exterior algebra generated by θ .

$\mathcal{A}_\omega$ is also a CDGA. As the following lemma shows, its isomorphism class depends only on the cohomology class $[\omega] \in H^*(\mathcal{A})$ in numerous instances.

Lemma 3.1 *Suppose $\mathcal{A}$ is a connected CDGA, ie $H^0(\mathcal{A}) = \mathbb{K}$. Take an even-degree, closed but nonexact element $\omega \in \mathcal{A}$, then construct $\mathcal{A}_\omega$ as above. Let $\bar{\omega} \in \mathcal{A}$ be an arbitrary representative of $[\omega] \in H^*(\mathcal{A})$, and $\bar{\theta}$ be any element in $\mathcal{A}_\omega$ satisfying $d\bar{\theta} = \bar{\omega}$. Then $\mathcal{A}_\omega$ can also be written as $\mathcal{A}_{\bar{\omega}} = \mathcal{A} \otimes \Lambda\bar{\theta}$.*

Proof Write $\bar{\theta} = \xi + \theta\eta$ for some $\xi, \eta \in \mathcal{A}$, then

$$\bar{\omega} - \omega - d\xi = d(\theta(\eta - 1)) = (\eta - 1)\omega - \theta d\eta$$

is exact in $\mathcal{A}$. Thus, $d\eta = 0$ and $[(\eta - 1)\omega] = 0$ in $H^*(\mathcal{A})$. Since $H^0(\mathcal{A}) = \mathbb{K}$, the closed element $\eta \in \mathcal{A}^0$ has to be a constant. Moreover, $[\omega] \in H^*(\mathcal{A})$ is a nontrivial class by hypothesis, so η can only be 1. Hence, $\bar{\theta}$ is in the affine space $\theta + \mathcal{A}$. Then we can write $\mathcal{A}_\omega = \mathcal{A} \oplus \bar{\theta}\mathcal{A}$ as a graded vector space.

By hypothesis ω is of even degree, so the degree of $\bar{\theta}$ is odd and $\bar{\theta}^2 = 0$. This fact together with the algebraic structure on $\mathcal{A}_\omega$ identify this CDGA and $\mathcal{A} \otimes \Lambda\bar{\theta}$. $\square$

When $\mathcal{A}$ is formal, there exists a zigzag of quasiisomorphisms connecting $\mathcal{A}$ and $H^*(\mathcal{A})$. These quasiisomorphisms are naturally extended to quasiisomorphisms connecting $\mathcal{A}_\omega$ and $H^*(\mathcal{A}) \otimes \Lambda\theta$. For the later CDGA we set $d\theta = [\omega]$. Thus, the formality of $\mathcal{A}_\omega$ is determined by $H^*(\mathcal{A}) \otimes \Lambda\theta$ as long as $\mathcal{A}$ is formal.

Therefore, it is sufficient to consider the CDGA $\mathcal{A}$ with trivial differential $d = 0$. In this case, the space of exact elements in $\mathcal{A}_\omega$ is $\text{im } \omega$, the image of the map by left multiplying ω ,

$$\omega: \mathcal{A}^i \rightarrow \mathcal{A}^{i+k}, \quad x \mapsto \omega x.$$

The space of closed elements in $\mathcal{A}_\omega$ is $\mathcal{A} \oplus \theta \ker \omega$. Thus,

$$H^*(\mathcal{A}_\omega) \simeq \text{coker } \omega \oplus \theta \ker \omega.$$

Definition 3.2 Let H^* be a finite dimensional graded commutative algebra over $\mathbb{K}$. If there exists some $\alpha_H \in (H^n)^\vee$, the dual space of H^n , such that the linear map

$$\alpha_H \frown: H^i \rightarrow (H^{m-i})^\vee, \quad x \mapsto (y \mapsto \alpha_H(xy))$$

is an isomorphism for all i , then we say that H is *n-dimensional Poincaré* and α_H is a *Poincaré class*.

A CDGA is called *n-dimensional Poincaré* if its cohomology is.

For a Poincaré CDGA, the Bianchi–Massey tensor $\mathcal{F}$ and the uniform Massey product $\mathcal{T}$ are equivalent [4, Lemma 2.8]. In particular, if $\mathcal{F}$ vanishes, we can choose α and γ such that $\mathcal{T} = 0$. Based on their proof we will show that, when $\mathcal{A}_\omega$ is a Poincaré CDGA, there exists such a choice also satisfying $\text{im } \gamma \subset \mathbf{I}(\theta)$. Here $\mathbf{I}(\theta)$ is the ideal generated by θ .

Lemma 3.3 Let $\mathcal{A}$ be a CDGA with trivial differential. Suppose $\mathcal{A}_\omega$ is *n-dimensional Poincaré* and the Bianchi Massey tensor $\mathcal{F}: \mathcal{B}^{n+1}(\mathcal{A}_\omega) \rightarrow H^n(\mathcal{A}_\omega)$ vanishes. There are choices of α and γ such that $\text{im } \gamma \subset \mathbf{I}(\theta)$ and the morphism

$$\gamma\alpha: E^*(\mathcal{A}_\omega) \otimes H^*(\mathcal{A}_\omega) \rightarrow \mathcal{A}_\omega^{*-1}, \quad e \otimes a \mapsto \gamma(e)\alpha(a)$$

is trivial on $K[E^*(\mathcal{A}_\omega) \otimes H^*(\mathcal{A}_\omega)]$.

Proof We will use E^* to denote $E^*(\mathcal{A}_\omega)$ in this proof. It is sufficient to construct α and γ such that $\text{im } \gamma \subset \mathbf{I}(\theta)$ and the uniform Massey product $\mathcal{T} = 0$ on $K[E^* \otimes H^*(\mathcal{A}_\omega)]$. Then the image of $\gamma\alpha$ are exact elements in $\mathbf{I}(\theta)$. But as discussed earlier in this section, the exact elements in $\mathcal{A}_\omega$ are in $\text{im } \omega$. So $\gamma\alpha$ has to be 0 on $K[E^* \otimes H^*(\mathcal{A}_\omega)]$.

Pick a right inverse of the left multiplication map $\omega: \mathcal{A} \rightarrow \text{im } \omega$ and extend it to an endomorphism on $\mathcal{A}$. We use ω^{-1} to denote this map.

Choose an $\alpha: H^*(\mathcal{A}_\omega) \rightarrow \mathcal{A}_\omega^*$ such that $\alpha(\theta \ker \omega) \subset I(\theta)$. As $\alpha^2(E^*)$ consists of exact forms, it is in $\text{im } \omega \subset \mathcal{A}$. So we can set

$$\gamma = \theta \circ \omega^{-1} \circ \alpha^2: E^* \rightarrow \mathcal{A}_\omega^{*-1}.$$

Then $\text{im } \gamma \subset I(\theta)$. It follows that $\text{im } \mathcal{T} \subset \theta \ker \omega$. For an extreme case that $\ker \omega = 0$, we already have $\mathcal{T} = 0$. Note that in this case $\mathcal{A}$ is infinite dimensional, unless $\mathcal{A} = 0$.

From now on we consider the general case that $\theta \ker \omega$ is nontrivial. Then $\theta \ker \omega$ must contain $H^n(\mathcal{A}_\omega)$, because for nontrivial $\theta x \in \theta \ker \omega$ there exists $y \in H^*(\mathcal{A}_\omega)$ such that

$$\alpha_H(\theta x y) = (\alpha_H \frown \theta x)y \neq 0.$$

To make $\mathcal{T} = 0$, we will change γ to $\gamma' = \gamma + \eta$ for some $\eta: E^* \rightarrow \mathcal{Z}^{*-1}(\mathcal{A}_\omega)$ satisfying $\text{im } \eta \subset I(\theta)$.

Consider the map

$$\gamma \alpha^2: E^* \otimes H^*(\mathcal{A}_\omega) \otimes H^*(\mathcal{A}_\omega) \rightarrow \mathcal{A}_\omega^{*-1}, \quad e \otimes x \otimes y \mapsto \gamma(e)\alpha(x)\alpha(y).$$

It takes closed values on $K[E^* \otimes H^*(\mathcal{A}_\omega) \otimes H^*(\mathcal{A}_\omega)]$, and factors through the projection

$$E^* \otimes H^*(\mathcal{A}_\omega) \otimes H^*(\mathcal{A}_\omega) \rightarrow E^* \otimes \mathcal{G}^2 H^*(\mathcal{A}_\omega).$$

Hence, $\gamma \alpha^2$ induces another map

$$\mu: K[E^* \otimes \mathcal{G}^2 H^*(\mathcal{A}_\omega)] \rightarrow H^{*-1}(\mathcal{A}_\omega).$$

μ depends on the choice of γ . Our goal is to find some γ' such that the corresponding $\mu' = 0$ when acting on the degree $n+1$ part of $K[E^* \otimes \mathcal{G}^2 H^*(\mathcal{A}_\omega)]$.

First consider the restriction of μ to $K[E^* \otimes E^*]$. Note that here $K[E^* \otimes E^*]$ means the kernel of full graded symmetrization $E^* \otimes E^* \rightarrow \mathcal{G}^4 H^*(\mathcal{A}_\omega)$. For arbitrary $e, e' \in E^*$,

$$\begin{aligned} \gamma(e)\alpha^2(e') - (-1)^{|e||e'|}\gamma(e')\alpha^2(e) &= \gamma(e)d\gamma(e') - (-1)^{|e||e'|+(|e'|-1)|e|}d\gamma(e)\gamma(e') \\ &= (-1)^{|e|-1}d(\gamma(e)\gamma(e')). \end{aligned}$$

So μ vanishes on graded antisymmetric tensors, and it factors through the projection $K[E^* \otimes E^*] \rightarrow \mathcal{B}^*(\mathcal{A}_\omega)$. Moreover, the induced map $\mathcal{B}^*(\mathcal{A}_\omega) \rightarrow H^{*-1}(\mathcal{A}_\omega)$ is exactly the Bianchi–Massey tensor, which is 0 when acting on $\mathcal{B}^{n+1}(\mathcal{A}_\omega)$, by assumption. Therefore, $\mu \equiv 0$ on $K[E^* \otimes E^*]$, in particular on its degree $n+1$ part. As $\eta(e)\alpha^2(e')$ is exact for any $e, e' \in E^*$, $\mu' = 0$ on the degree $n+1$ part of $K[E^* \otimes E^*]$ for any choice of γ' .

Let $q: \mathcal{G}^2 H^*(\mathcal{A}_\omega) \rightarrow \mathcal{G}^2(\text{coker } \omega)$ be the morphism induced by the projection $H^*(\mathcal{A}_\omega) \rightarrow \text{coker } \omega$ with kernel $\theta \ker \omega$, and L^* be the preimage of $\theta \ker \omega$ under the multiplication map $c: \mathcal{G}^2 H^*(\mathcal{A}_\omega) \rightarrow H^*(\mathcal{A}_\omega)$. Then $c \circ q(L^*) = 0$ because its image is the projection of $c(L^*)$. Hence, $q(L^*) \subset E^*$. So there exists an induced map

$$Q: E^* \otimes L^* \rightarrow E^* \otimes E^*, \quad e \otimes a \mapsto qe \otimes qa.$$

Moreover, Q sends $K[E^* \otimes L^*]$ into $K[E^* \otimes E^*]$. Here $K[E^* \otimes L^*]$ also means the kernel of full graded symmetrization $E^* \otimes L^* \rightarrow \mathcal{G}^4 H^*(\mathcal{A}_\omega)$.

Let $e \in E^*$ and $a \in L^*$ with $|e| + |a| = n + 1$. Then every term of $a - qa$ has a factor in $\theta \ker \omega$. Thus, $\alpha^2(a - qa) \in \mathbf{I}(\theta)$. As $\gamma(e)$ is also in $\mathbf{I}(\theta)$, we have $\gamma(e)\alpha^2(a) = \gamma(e)\alpha^2(qa)$. For the same reason,

$$\begin{aligned} \gamma(e)\alpha^2(qa) &= (-1)^{|e|} d(\gamma(e)\gamma(qa)) - (-1)^{|e|}\alpha^2(e)\gamma(qa) \\ &= (-1)^{|e|} d(\gamma(e)\gamma(qa)) - (-1)^{|e|}\alpha^2(qe)\gamma(qa) \\ &= (-1)^{|e|} d(\gamma(e)\gamma(qa)) - (-1)^{|e|} d(\gamma(qe)\gamma(qa)) + \gamma(qe)\alpha^2(qa). \end{aligned}$$

Therefore, $\mu(e \otimes a) = [\gamma(e)\alpha^2(a)] = [\gamma(qe)\alpha^2(qa)] = \mu \circ Q(e \otimes a)$. As discussed above, μ vanishes on $K[E^* \otimes E^*]$. So $\mu(K[E^* \otimes L^*]) = \mu \circ Q(K[E^* \otimes L^*]) \subset \mu(K[E^* \otimes E^*]) = 0$.

Now let $p = \text{id} \otimes c: K[E^* \otimes \mathcal{G}^2 H^*(\mathcal{A}_\omega)] \rightarrow E^* \otimes H^*(\mathcal{A}_\omega)$ sending $e \otimes (x \cdot y)$ to $e \otimes xy$. Then $\ker p$ is exactly $K[E^* \otimes E^*]$. Hence, p induces a morphism $\bar{\mu}: \text{im } p \rightarrow H^*(\mathcal{A}_\omega)$ of degree -1 . As μ vanishes on $K[E^* \otimes L^*]$, $\bar{\mu} = 0$ on $\text{im } p \cap (E^* \otimes \theta \ker \omega)$. So we can extend $\bar{\mu}$ to all $E^* \otimes H^*(\mathcal{A}_\omega)$ such that it vanishes on $E^* \otimes \theta \ker \omega$.

For each $e \in E^i$, $\bar{\mu}$ induces a morphism

$$H^{n+1-i}(\mathcal{A}_\omega) \rightarrow \mathbb{K}, \quad x \mapsto -\alpha_H \circ \bar{\mu}(e \otimes x)$$

where α_H is the Poincaré class of $H^*(\mathcal{A}_\omega)$. Let $\delta(e)$ be the unique element in $H^{i-1}(\mathcal{A}_\omega)$ such that this morphism equals to $\alpha_H \frown \delta(e)$. We claim that $\delta(e) \in \theta \ker \omega$.

Write $\delta(e) = x_0 + \theta y_0$ for some $x_0 \in \text{coker } \omega$ and $y_0 \in \ker \omega$. If $x_0 \neq 0$, there exist some $x_1 \in \text{coker } \omega$ and $y_1 \in \ker \omega$ such that $(\alpha_H \frown x_0)(x_1 + \theta y_1) \neq 0$, ie $x_0(x_1 + \theta y_1)$ is nontrivial in $H^n(\mathcal{A}_\omega)$. By the discussion at the beginning of the proof, $H^n(\mathcal{A}_\omega) \subset \theta \ker \omega$. So $x_0 x_1$ has to be 0. It follows that $\delta(e)\theta y_1 = (x_0 + \theta y_0)\theta y_1 = x_0 \theta y_1 \neq 0$. On the other hand, $\delta(e)\theta y_1 = -\alpha_H \circ \bar{\mu}(e \otimes \theta y_1) = 0$ as $\bar{\mu} = 0$ on $E^* \otimes \theta \ker \omega$, which is a contradiction. Hence, x_0 must be 0 and $\delta(e) \in \theta \ker \omega$.

Let $\eta = \alpha \circ \delta: E^* \rightarrow \mathcal{Z}^{*-1}(\mathcal{A}_\omega)$. It represents δ and its image is in $\mathbf{I}(\theta)$. Then $\gamma'\alpha^2 = (\gamma + \eta)\alpha^2$ induces a map $\mu': K[E^* \otimes \mathcal{G}^2 H^*(\mathcal{A}_\omega)] \rightarrow H^{*-1}(\mathcal{A}_\omega)$. For arbitrary $\sum e_j \otimes a_j \in K[E^* \otimes \mathcal{G}^2 H^*(\mathcal{A}_\omega)]$,

$$\mu'\left(\sum e_j \otimes a_j\right) = \mu\left(\sum e_j \otimes a_j\right) + \sum [\eta(e_j)\alpha^2(a_j)] = \bar{\mu}\left(\sum e_j \otimes c(a_j)\right) + \sum \delta(e_j)c(a_j) = 0.$$

Finally, we prove that the uniform Massey product $\mathcal{T}'$ defined by γ' is identically 0. This follows from $K[E^* \otimes H^*(\mathcal{A}_\omega)] \otimes H^*(\mathcal{A}_\omega) \subset K[E^* \otimes H^*(\mathcal{A}_\omega) \otimes H^*(\mathcal{A}_\omega)]$ and

$$\mathcal{T}'\left(\sum e_j \otimes x_j\right)y = \left[\sum \gamma(e_j)\alpha(x_j)\alpha(y)\right] = \mu'\left(\sum e_j \otimes (x_j \cdot y)\right) = 0$$

for $e_j \in E^*$, $x_j, y \in H^*(\mathcal{A}_\omega)$ and $|e_j| + |x_j| + |y| = n + 1$. Thus, $\alpha_H \frown \mathcal{T}'(\sum e_j \otimes x_j) = 0$ and then $\mathcal{T}'(\sum e_j \otimes x_j) = 0$. As $\text{im } \gamma' \subset \mathbf{I}(\theta)$ by construction, the discussion in the first paragraph of the proof shows that $\gamma'\alpha$ has to be 0 on $K[E^* \otimes H^*(\mathcal{A}_\omega)]$. $\square$

Lemma 3.4 Let $\mathcal{A}$ be any connected CDGA. Suppose that there are choices of $\alpha: H^*(\mathcal{A}) \rightarrow \mathcal{A}^*$ and $\gamma: E^*(\mathcal{A}) \rightarrow \mathcal{A}^{*-1}$ such that the restriction of $\gamma\alpha$ to $K[E^*(\mathcal{A}) \otimes H^*(\mathcal{A})]$ and

$$\gamma^2: E^*(\mathcal{A}) \otimes E^*(\mathcal{A}) \rightarrow \mathcal{A}^{*-2}, \quad e \otimes e' \mapsto \gamma(e)\gamma(e')$$

both vanish. Then $\mathcal{A}$ is formal.

Proof For simplicity, in this proof we will use H^* and E^* to denote $H^*(\mathcal{A})$ and $E^*(\mathcal{A})$, respectively.

Construct a Sullivan model $\phi: \Lambda V \rightarrow H^*$ as follows. Start from a d -closed space V_0^* of lower index 0. Then construct V_p^* inductively such that $dV_p^k \subset (\Lambda V_{<p}^*)^{k+1}$. Here $V_{<p}^*$ denotes $V_0^* \oplus V_1^* \oplus \cdots \oplus V_{p-1}^*$, and $(\Lambda V_{<p}^*)_{p-1}$ is the subspace of elements of lower index $p-1$, where the lower index of an element in $V_{i_1}^* \otimes \cdots \otimes V_{i_k}^*$ is $i_1 + \cdots + i_k$.

Precisely, let $V_0^* = H^+$ as a graded vector space, ie for each $x \in H^+$, there exists a unique $v_x \in V_0^*$ such that $\phi(v_x) = x$. Then let V_1^k be spanned by $\{w_{(x \cdot y)} \mid (x \cdot y) \in (\mathcal{G}^2 H^+)^{k+1}\}$, and set $dw_{(x \cdot y)} = v_{xy} - v_x v_y$. Thus, ϕ induces an isomorphism from the quotient space ΛV_0^* over the ideal generated by dV_1^* to H^* .

For $p \geq 2$, define V_p^* such that $d: V_p^k \rightarrow (\Lambda V_{<p}^*)^{k+1} \cap \ker d$ is an isomorphism. Let $V = \bigoplus_{p \geq 0} V_p^*$ and $\phi(V_p^*) = 0$ for all $p \geq 1$. Then $\phi: \Lambda V \rightarrow H^*$ is a quasiisomorphism.

Now we construct a CDGA morphism $f: \Lambda V \rightarrow \mathcal{A}^*$ as follows. For $v_x \in V_0^*$, let $f(v_x) = \alpha(x)$. For $w_{(x \cdot y)} \in V_1^*$, let $f(w_{(x \cdot y)}) = \gamma((1 \cdot xy) - (x \cdot y))$. Then f is well-defined on $\Lambda(V_0^* \oplus V_1^*)$. We claim that $f = 0$ when acting on the closed elements of lower index 1, so that we can make $f = 0$ on V_2^* .

To prove the claim, we will use the fact that for $x, y, z \in H^+$,

$$(-1)^{|x|} v_x w_{(y \cdot z)} - w_{(x \cdot y)} v_z - w_{(xy \cdot z)} + w_{(x \cdot yz)}$$

is a closed element in $\ker f$. A straight forward calculation shows that it is closed. And f acting on this element is $\gamma\alpha$ acting on

$$\Phi = (-1)^{|x|(|y|+|z|)} (1 \cdot yz - y \cdot z) \otimes x - (1 \cdot xy - x \cdot y) \otimes z - (1 \cdot xyz - xy \cdot z) \otimes 1 + (1 \cdot xyz - x \cdot yz) \otimes 1.$$

Φ is in $E^* \otimes H^*$ and its full symmetrization is

$$(-1)^{|x|(|y|+|z|)} (1 \cdot yz \cdot x - y \cdot z \cdot x) - (1 \cdot xy \cdot z - x \cdot y \cdot z) - (1 \cdot xyz \cdot 1 - xy \cdot z \cdot 1) + (1 \cdot xyz \cdot 1 - x \cdot yz \cdot 1) = 0.$$

Therefore, $\Phi \in K[E^* \otimes H^*]$ and $\gamma\alpha(\Phi) = 0$ by hypothesis.

Now let Ψ be a closed element in $\Lambda(V_0^* \oplus V_1^*)$ of lower index 1. So $\Psi \in V_1^* \cdot \Lambda V_0^*$. Choose a basis $\{x_i\}$ of H^+ , then Ψ can be written as a linear combination of $w_{(x_{i_1} \cdot x_{i_2})} v_{x_{i_3}} \cdots v_{x_{i_l}}$, $l \geq 2$. Suppose k is the largest integer such that Ψ has a term in $V_1 \cdot \Lambda^k V_0^*$, and there are m_k terms in $V_1^* \cdot \Lambda^k V_0^*$ under this writing.

When $k \geq 1$, without loss of generality we may assume that Ψ has a term $w_{(x_1 \cdot x_2)} v_{x_3} \cdots v_{x_{k+2}}$. As Ψ is closed, there must be another term such that its differential contains $v_{x_1} \cdots v_{x_{k+2}}$. Moreover, this term is

also in $V_1 \cdot \Lambda^k V_0$ since $dV_1 \cdot \Lambda^j V_0 \subset \Lambda^{\leq j+2} V_0$ for all $j \geq 0$. So it is of the form $Cw_{(x_{i_1} \cdot x_{i_2})} v_{x_{i_3}} \cdots v_{x_{i_{k+2}}}$, where $\{i_1, \dots, i_{k+2}\}$ is a reshuffle of $\{1, \dots, k+2\}$ and $C \in \mathbb{K}$ is a constant. Then either one of 1, 2 is in $\{i_1, i_2\}$, or $1, 2 \notin \{i_1, i_2\}$.

Case 1 $\{1, 2\} \cap \{i_1, i_2\} \neq \emptyset$.

We may assume that the other term is $Cw_{(x_2 \cdot x_3)} v_{x_1} v_{x_4} \cdots v_{x_{k+2}}$. Set

$$\Psi' = \Psi + ((-1)^{|x_1|} v_{x_1} w_{(x_2 \cdot x_3)} - w_{(x_1 \cdot x_2)} v_{x_3} - w_{(x_1 x_2 \cdot x_3)} + w_{(x_1 \cdot x_2 x_3)}) v_{x_4} \cdots v_{x_{k+2}}.$$

Then $d\Psi' = d\Psi = 0$ and $f(\Psi') = f(\Psi)$. Moreover, Ψ' has at most $m_k - 1$ terms in $V_1 \cdot \Lambda^k V_0$.

Case 2 $\{1, 2\} \cap \{i_1, i_2\} = \emptyset$.

We may assume that the other term is $Cw_{(x_3 \cdot x_4)} v_{x_1} v_{x_2} v_{x_5} \cdots v_{x_{k+2}}$. Set

$$\begin{aligned} \Psi' = \Psi + & ((-1)^{|x_1|} v_{x_1} w_{(x_2 \cdot x_3)} - w_{(x_1 \cdot x_2)} v_{x_3} - w_{(x_1 x_2 \cdot x_3)} + w_{(x_1 \cdot x_2 x_3)}) v_{x_4} \cdots v_{x_{k+2}} \\ & + (-1)^{|x_1|} v_{x_1} ((-1)^{|x_2|} v_{x_2} w_{(x_3 \cdot x_4)} - w_{(x_2 \cdot x_3)} v_{x_4} - w_{(x_2 x_3 \cdot x_4)} + w_{(x_2 \cdot x_3 x_4)}) v_{x_5} \cdots v_{x_{k+2}}. \end{aligned}$$

Then we also have $d\Psi' = d\Psi = 0$, $f(\Psi') = f(\Psi)$, and Ψ' has at most $m_k - 1$ terms in $V_1 \cdot \Lambda^k V_0$.

Continue this process we can finally find some closed $\Psi'' \in V_1^*$ such that $f(\Psi'') = f(\Psi)$. But by construction the only closed element in V_1^* is 0. So $\Psi'' = 0$ and $f(\Psi) = 0$. The claim is proved.

Therefore, we can set $f(V_2^*) = 0$. On the other hand, $f(V_1^*) \subset \text{im } \gamma$. By hypothesis $f(\Lambda^2 V_1^*) \subset \text{im } \gamma^2 = 0$. Hence, $f = 0$ on $\Lambda(V_{\leq 2}^*)_2$ and we can set $f(V_3^*) = 0$. Inductively we have $f = 0$ on $\Lambda(V_{\leq p}^*)_p$ and set $f(V_{p+1}^*) = 0$ for $p \geq 2$.

It remains to verify that f is a quasiisomorphism. For each $x \in H^+$, there exists some closed $v_x \in V_0^*$ such that the cohomology class $[f(v_x)] = [\alpha(x)] = x$. Together with $f(1) = 1$ we have proved the surjectivity.

For injectivity, suppose that $\Omega \in \Lambda V$ is closed and $f(\Omega)$ is exact. Observe that d decreases the lower index by 1 homogeneously and all the closed elements of positive lower index are exact. So the cohomology class $[\Omega]$ has a representative in ΛV_0^* . Moreover, by construction for any $x, y \in H^+$, the cohomology class $[v_{xy}] = [v_x v_y]$. It follows that $[\Omega]$ has a representative in either V_0^* or $\mathbb{K}$. But f acting on $V_0^* \oplus \mathbb{K}$ is nonexact except $f(0)$. So this representative is 0 and Ω is exact.

Therefore, ϕ and f give a CDGA equivalence between H^* and $\mathcal{A}^*$. □

Remark 3.5 We can also construct an A_∞ -quasiisomorphism from H^* to $\mathcal{A}^*$, by setting $f_1 = \alpha$, $f_2(x, y) = \gamma(1 \cdot xy - x \cdot y)$, and $f_p = 0$ for all $p \geq 3$.

The CDGA $\mathcal{A}_\omega$ in Lemma 3.3, satisfies the hypothesis of Lemma 3.4. So $\mathcal{A}_\omega$ is formal and we have the following theorem.

Theorem 3.6 Let $\mathcal{A}$ be a formal CDGA, and $\mathcal{A}_\omega = \mathcal{A} \otimes \Lambda\theta$ be a Poincaré algebra where $\theta^2 = 0$ and $d\theta = \omega \in \mathcal{A}$. Then $\mathcal{A}_\omega$ is formal if and only if its Bianchi–Massey tensor vanishes.

Let $\pi: X \rightarrow M$ be an orientable S^k -bundle. We have the following CDGA equivalence:

$$\Omega^*(X) \simeq \begin{cases} \Omega^*(M) \otimes \Lambda(\theta), & d\theta = e & \text{if } k \text{ is odd,} \\ \Omega^*(M) \otimes \Lambda(\theta, \theta'), & d\theta = 0, \quad d\theta' = \theta^2 + \frac{1}{4}p & \text{if } k \text{ is even.} \end{cases}$$

Here $|\theta| = k$, and $|\theta'| = 2k - 1$. The Euler class is $[e] \in H^{k+1}(M)$ and $[p] \in H^{2k}(M)$ is the rational $2k^{\text{th}}$ Pontryagin class of the sphere bundle.

When M is simply connected, the proof can be found in [6, Example 4, page 202]. For general manifolds, the case that k is odd is proved in [11, the appendix], and a similar proof works when k is even. Moreover, this equivalence still holds when X is a spherical Serre fibration, as long as $\pi_1(M)$ acts nilpotently on $H^*(S^k)$ [7, Theorem 20.3], which is equivalent to the holonomy action of $\pi_1(M)$ on the fiber S^k preserves the orientation.

Therefore, when M is compact, orientable and formal, $\Omega^*(X)$ is equivalent to a CDGA $\mathcal{A}_\omega$ satisfying the hypothesis of Theorem 3.6. When k is odd, $\mathcal{A} = H^*(M)$ and $\omega = [e]$. When k is even,

$$\mathcal{A} = H^*(M) \otimes \Lambda(\theta), \quad \mathcal{A}_\omega = \mathcal{A} \otimes \Lambda(\theta'), \quad \text{and} \quad \omega = \theta^2 + \frac{1}{4}[p].$$

In both cases the formality of X is determined by the Bianchi–Massey tensor of $\mathcal{A}_\omega$.

Moreover, if k is even, the kernel of multiplying by $\omega = \theta^2 + \frac{1}{4}[p]$ in $\mathcal{A} = H^*(M) \otimes \Lambda(\theta)$ is 0. By the discussion in the third paragraph of the proof of Lemma 3.3, the uniform Massey product $\mathcal{T}$ of $\mathcal{A}_\omega = H^*(M) \otimes \Lambda(\theta, \theta')$ is trivial as $\ker \omega = 0$. Then it is formal by Lemma 3.4. Thus, we have the following statement.

Theorem 3.7 Suppose M is a compact orientable formal manifold, and $\pi: X \rightarrow M$ is an orientable S^k -bundle. Then X is formal if and only if the Bianchi–Massey tensor of $\Omega^*(X)$ vanishes. Moreover, when k is even, X is always formal.

3.2 A special case: Euler class is of the top degree

In this subsection, we will consider the case that the Euler class is a top degree cohomology class of a formal manifold. An example of such bundle is the unit tangent bundle. Let M be a compact orientable formal manifold. Equip the tangent bundle TM with a metric, then the vectors of norm 1 form a sphere bundle UTM , whose Euler class $\chi(M)[\omega]$. Here $\chi(M)$ is the Euler characteristic and ω is a volume form. UTM is called the unit tangent bundle of M .

We will explore when such bundles are formal. When the base is of odd dimension, the fiber is an even dimensional sphere. By the discussion of the previous subsection, we already know that the bundle is formal. So the nontrivial case only happens on even dimensional base manifolds.

Lemma 3.8 Suppose $\mathcal{A}$ is a $2n$ -dimensional Poincaré CDGA with trivial differential, and $\omega \in \mathcal{A}^{2n}$ be nonzero. We also assume that $\mathcal{A}^i = 0$ if $i < 0$, and $\mathcal{A}^0 = \mathbb{K}$. If $\mathcal{A}_\omega$ is formal, then the multiplication map $\mathcal{A}^i \otimes \mathcal{A}^j \rightarrow \mathcal{A}^{i+j}$ is injective for all $i, j \leq n$.

Proof If i or $j = 0$, the multiplication map is an isomorphism. Assume that the multiplication map has a nontrivial kernel in $\sum_{r=1}^k x_r \otimes y_r \in \mathcal{A}^i \otimes \mathcal{A}^j$, where $0 < i, j \leq n$. $\{x_r\}$ can be chosen linearly independent with $k \geq 1$. A similar procedure can make $\{y_r\}$ linearly independent: If $y_r = c_1 y_1 + \cdots + c_{r-1} y_{r-1}$, $x_1 \otimes y_1 + \cdots + x_r \otimes y_r$ can be rewritten as $(x_1 + c_1 x_r) \otimes y_1 + \cdots + (x_{r-1} + c_{r-1} x_r) \otimes y_{r-1}$. And $\{x_1 + c_1 x_r, \dots, x_{r-1} + c_{r-1} x_r\}$ is also linearly independent.

Take $x_1^* \in \mathcal{A}^{2n-i}$, $y_1^* \in \mathcal{A}^{2n-j}$ such that $x_r x_1^* = y_r y_1^* = \delta_{1r} \omega$. Then $n \leq x_1^*, y_1^* < 2n$. Hence, x_1^*, y_1^* are nonexact in $\mathcal{A}_\omega$, but the products $x_r x_1^*, y_r y_1^*$ are either 0 or ω , which are both exact in $\mathcal{A}_\omega$. So we can set $\alpha: H^*(\mathcal{A}_\omega) \rightarrow \mathcal{A}_\omega$ such that

$$\alpha([x_r]) = x_r, \quad \alpha([y_r]) = y_r, \quad \alpha([x_1^*]) = x_1^*, \quad \alpha([y_1^*]) = y_1^*,$$

and $\gamma: E^*(\mathcal{A}_\omega) \rightarrow \mathcal{A}_\omega$ such that

$$\gamma\left(\sum_{r=1}^k [x_r] \cdot [y_r]\right) = 0, \quad \gamma([x_r] \cdot [x_1^*]) = \gamma([y_r] \cdot [y_1^*]) = \delta_{1r} \theta.$$

Then

$$\left(\left(\sum_{r=1}^k [x_r] \cdot [y_r]\right) \cdot ([x_1^*] \cdot [y_1^*])\right) - (-1)^{j(2n-i)} \sum_{r=1}^k ([x_r] \cdot [x_1^*]) \cdot ([y_r] \cdot [y_1^*])$$

is in $K[\mathcal{G}^2 \mathcal{G}^2 H^*(\mathcal{A}_\omega)]$, and

$$\begin{aligned} \mathcal{F}\left(\left(\left(\sum_{r=1}^k [x_r] \cdot [y_r]\right) \cdot ([x_1^*] \cdot [y_1^*])\right) - (-1)^{j(2n-i)} \sum_{r=1}^k ([x_r] \cdot [x_1^*]) \cdot ([y_r] \cdot [y_1^*])\right) \\ = -(-1)^{ij} [\gamma([x_1] \cdot [x_1^*]) y_1 y_1^*] - (-1)^{ij} (-1)^{2n-2n} [\gamma([y_1] \cdot [y_1^*]) x_1 x_1^*] \\ = -(-1)^{ij} 2[\theta \omega]. \end{aligned}$$

Therefore, $\mathcal{A}_\omega$ is nonformal. □

Lemma 3.9 Suppose H^* is a $2n$ -dimensional Poincaré graded algebra, $H^0 = \mathbb{K}$, and H^i is nontrivial only when $0 \leq i \leq 2n$. If the multiplication map $H^i \otimes H^j \rightarrow H^{i+j}$ is injective for all $i, j \leq n$, then $H^* = \mathbb{K}[x]/(x^p)$ is a quotient of the polynomial ring with a single variable.

Proof First observe that $\dim H^{2i} \leq 1$ and $\dim H^{2i+1} = 0$. If $x, y \in H^{2i}$ are linearly independent for some $0 < 2i \leq n$, then $x \otimes y - y \otimes x$ will be a nontrivial element in the kernel of the multiplication map $H^{2i} \otimes H^{2i} \rightarrow H^{4i}$. If $z \in H^{2i+1}$ is nonzero for some $0 < 2i+1 \leq n$, then $z \otimes z$ will be a nontrivial element in the kernel of the multiplication map $H^{2i+1} \otimes H^{2i+1} \rightarrow H^{4i+2}$. As H^* is $2n$ -dimensional Poincaré, for $n < 2i \leq 2n$ we have $\dim H^{2i} = \dim H^{2n-2i} \leq 1$, and for $n < 2i+1 \leq 2n$ we have $\dim H^{2i+1} = \dim H^{2n-2i-1} = 0$.

Let $S = \{i \in \mathbb{Z} \mid \dim H^i = 1\}$, k be the smallest positive integer in S , and T be the image of S under the projection $\mathbb{Z} \rightarrow \mathbb{Z}/k\mathbb{Z}$. For $i, j \in S$ with $i, j \leq n$, the hypothesis that the multiplication $H^i \otimes H^j \rightarrow H^{i+j}$ is injective implies $i + j \in S$. In particular, $ik \in S$ for $0 \leq i \leq 2\lfloor \frac{n}{k} \rfloor$. Let x be a generator of the vector space H^k , then x^i generates H^{ik} for such i .

We claim that every $a \in T$ has a representative $m \leq n$ in S . Indeed, when $m > n$ is a representative of a with $m \in S$, we have $2n - m < n$ and $2n - m \in S$ as H^* is Poincaré. On the other hand, $\lfloor \frac{n}{k} \rfloor k \leq n$ is also in S as discussed in last paragraph, so $(2n - m) + \lfloor \frac{n}{k} \rfloor k \in S$. By the definition of k , we have $2n - m \geq k$. Then $(2n - m) + \lfloor \frac{n}{k} \rfloor k \geq (1 + \lfloor \frac{n}{k} \rfloor)k > n$. Thus, we can set $m' = 2n - ((2n - m) + \lfloor \frac{n}{k} \rfloor k) = m - \lfloor \frac{n}{k} \rfloor k$, which is also a representative of a . Moreover, $m' \in S$ and $m' < n$.

For arbitrary $a_1, a_2 \in T$, there are respective representatives $m_1, m_2 \leq n$ in S . Then $m_1 + m_2 \in S$ and $a_1 + a_2 \in T$. Since the finite subset T of $\mathbb{Z}/k\mathbb{Z}$ is closed under addition, it is a subgroup. So $T = \mathbb{Z}/c\mathbb{Z}$ for some positive integer c , and $\frac{k}{c}$ is the number of elements in T .

Choose the largest $m \leq n$ in S . We claim that $n - m < c$. Otherwise, the class of $m + c$ in $\mathbb{Z}/k\mathbb{Z}$ is also in T , and it has a representative $m' \leq n$ in S . By assumption, $m' \leq m < m + c$. So $m + c - m'$ is a positive integer divisible by k . Moreover, $m + c - m' \leq m + c < n < (1 + \lfloor \frac{n}{k} \rfloor)k \leq 2\lfloor \frac{n}{k} \rfloor k$ as $\lfloor \frac{n}{k} \rfloor \geq 1$ clearly. Hence, $m + c - m' \in S$. It follows that $m + c = m' + (m + c - m') \in S$, which is a contradiction.

If $T \neq \{0\}$, the number of its elements $\frac{k}{c} \geq 2$. Together with $n - m < c$ we have $2n - 2m < k$. On the other hand, $2m \in S$ as $m < n$ is in S , then $2n - 2m \in S$, which contradicts to the hypothesis of k . So T has to be $\{0\}$ and k divides $2n$.

When k divides n , $2\lfloor \frac{n}{k} \rfloor k = 2n$. The discussion in the second paragraph shows that

$$H^* = \langle x^i \mid 0 \leq |x^i| \leq 2n \rangle.$$

When k does not divide n , $(2\lfloor \frac{n}{k} \rfloor + 1)k = 2n$. In the special case that $k = 2n$, $H^* = H^0 \oplus H^{2n}$. So any generator of x the space H^{2n} makes $H^* = \mathbb{K}[x]/(x^2)$. In the general case, $H^k, H^{2k}, \dots, H^{2n-k}$ are all generated by some x^i , and it remains to verify that the generator of H^{2n} is also a power of x . Since H^* is Poincaré, this generator is the product of x and a nontrivial element in H^{2n-k} . But the only generator of H^{2n-k} is x^i with $i = 2\lfloor \frac{n}{k} \rfloor$, so x^{i+1} is a generator of H^{2n} . This completes the proof. $\square$

Combining the lemmas above, we have the following theorem.

Theorem 3.10 Suppose that $\mathcal{A}$ is a $2n$ -dimensional Poincaré CDGA, which is formal and connected. Let $\omega \in \mathcal{A}^{2n}$ such that $[\omega]$ is nonzero. Set $\mathcal{A}_\omega = \mathcal{A} \otimes \Lambda\theta$ with $d\theta = \omega$. If $\mathcal{A}_\omega$ is formal, then $H^*(\mathcal{A}) = \mathbb{K}[x]/(x^p)$ is a quotient of the polynomial ring with a single variable.

Corollary 3.11 Let M be a compact orientable formal manifold. Its unit tangent bundle UTM is formal if and only if one of the following statements hold:

- (1) The Euler characteristic $\chi(M) = 0$.
- (2) $H^*(M) = \mathbb{K}[x]/(x^p)$ is a quotient of the polynomial ring with a single variable.

Proof If $\chi(M) \neq 0$, then M is an even dimensional manifold and the Euler class of UTM is nontrivial. [Theorem 3.10](#) states that $H^*(M)$ has to be a quotient of the polynomial ring with a single variable.

Conversely, when $\chi(M) = 0$, the Euler class of UTM is trivial. It follows that UTM is a trivial bundle and formal. When $H^*(M) = \mathbb{K}[x]/(x^p)$ and $\chi(M) \neq 0$, $\Omega^*(UTM)$ is equivalent to $(\mathbb{K}[x]/(x^p)) \otimes \Lambda\theta$ with $d\theta = x^{p-1}$. $H^*(UTM)$ is spanned by $1, x, \dots, x^{p-2}, \theta x, \theta x^2, \dots, \theta x^{p-1}$. Let $\mathcal{M} = \Lambda(u, v, w)$ such that $|u| = |x|$, $|v| = |\theta|$, $|w| = |\theta x|$, $du = dw = 0$ and $dv = u^{p-1}$. Then we can define quasiisomorphisms $\phi: \mathcal{M} \rightarrow H^*(UTM)$ and $f: \mathcal{M} \rightarrow (\mathbb{K}[x]/(x^p)) \otimes \Lambda(\theta)$ such that

$$\phi(u) = f(u) = x, \quad \phi(v) = 0, \quad f(v) = \theta, \quad \phi(w) = f(w) = \theta x.$$

This gives the CDGA equivalence between $H^*(UTM)$ and $\Omega^*(UTM)$. $\square$

Corollary 3.12 *Let M be a compact orientable formal manifold. If the Euler characteristic $\chi(M) < 0$, then UTM is nonformal.*

Example 3.13 As a simple example, the circle bundles over Riemann surfaces distinguish the different cases above:

genus	unit tangent bundle	Euler class is volume form
0	formal	formal
1	formal	nonformal
≥ 2	nonformal	nonformal

4 An obstruction to the formality of general sphere bundles

If the base manifold M is compact and formal, we have established when the sphere bundle X is formal. One may consider the case for nonformal M . We will discuss it in this section, and give an obstruction to the formality of X .

Let $\mathcal{M}$ be the minimal Sullivan model of a CDGA $\mathcal{A}$. For a closed element $\omega \in \mathcal{A}$, there exists a closed element in $\mathcal{M}$ whose cohomology class is corresponding to $[\omega]$. We will also write this element in $\mathcal{M}$ as ω . Regardless of the choice of ω , we have the quasiisomorphism

$$\mathcal{M}_\omega = \mathcal{M} \otimes \Lambda\theta \simeq \mathcal{A} \otimes \Lambda\theta = \mathcal{A}_\omega,$$

where $d\theta = \omega$. In this section we assume that $|\omega|$ is even. So θ^2 is automatically 0.

Whether the Sullivan model $\mathcal{M}_\omega$ is minimal depends on whether $\omega \in \mathcal{M}$ is reducible. Note that the representatives of $[\omega] \in H^*(\mathcal{M})$ are either all reducible or all irreducible, since exact elements in $\mathcal{M}$ are reducible by the definition of minimal Sullivan algebra in [Section 2.2](#).

Proposition 4.1 *When $\omega \in \mathcal{M}$ is reducible, $\mathcal{M}_\omega$ is a minimal Sullivan model of $\mathcal{A}_\omega$.*

Proof Suppose $\mathcal{M} = \Lambda V^*$ and $V^* = \langle v_\alpha \rangle$. Write $\|v_\alpha\| = \alpha$ as the index of v_α . Then set $\|\theta\| > \alpha$ for all $|v_\alpha| \leq |\theta|$, and $\|\theta\| < \alpha$ for all $|v_\alpha| < |\theta|$. So $\{v_\alpha\} \cup \{\theta\}$ is a well-ordered set. It is straightforward to verify that $\mathcal{M}_\omega = \Lambda(V^* \oplus \langle \theta \rangle)$ is a minimal Sullivan algebra. $\square$

Theorem 4.2 Suppose ω is a closed element in a minimal Sullivan algebra $\mathcal{M}$ satisfying the following conditions.

- (1) $|\omega| = 2r$ for some odd integer r , and $[\omega]$ has a representative that can be written as

$$\sum_{i=1}^k x_i y_i,$$

where x_i, y_i are all closed in $\mathcal{M}^r$.

- (2) There exists some $s \geq 0$ such that $H^s(\mathcal{M})$ is nontrivial. Moreover, the morphism

$$\omega: H^s(\mathcal{M}) \rightarrow H^{s+2r}(\mathcal{M})$$

multiplying by $[\omega]$ is an isomorphism, and $\omega: H^{s-r}(\mathcal{M}) \rightarrow H^{s+r}(\mathcal{M})$ is injective.

Then $\mathcal{M}_\omega = \mathcal{M} \otimes \Lambda \theta$ is nonformal, where $d\theta = \omega$.

Proof Assume that $\mathcal{M}_\omega$ is formal. By Theorem 2.6 we can write $\mathcal{M}_\omega = \Lambda V^*$ and $V^* = C^* \oplus N^*$ such that all closed elements in $I(N^*)$ are exact.

Since $\omega: H^s(\mathcal{M}) \rightarrow H^{s+2r}(\mathcal{M})$ is an isomorphism, $[\omega] \in H^{2r}(\mathcal{M})$ cannot be a trivial class. So $\mathcal{M}_\omega$ satisfies the hypothesis of Lemma 3.1, which allows us to reset ω and θ . Then without loss of generality, we can assume that ω itself has the form $\sum_{i=1}^k x_i y_i$ in condition (1), and $\theta \in I(N^*)$.

We will first prove the following claim that will be used in proof.

Claim Suppose $\mathcal{M}$ satisfies the hypothesis of the theorem, and $\mathcal{M}_\omega$ is formal. Let $\beta \in \mathcal{M}^{s+r}$ be a closed element. If $\omega\beta = \sum w_j z_j$, where $w_j \in \mathcal{M}^{s+2r}$ are all exact in $\mathcal{M}$ and $z_j \in \mathcal{M}^r$ are all closed, then β is exact in $\mathcal{M}$.

Since w_j is exact in $\mathcal{M}$, it is also exact in $\mathcal{M}_\omega$ and can be written as $w_j = d(\xi_j + \theta\eta_j)$, where $\xi_j \in \mathcal{M}^{s+2r-1}$, $\eta_j \in \mathcal{M}^r$ and $\xi_j + \theta\eta_j \in I(N^*)$. As $w_j = d\xi_j + \omega\eta_j + \theta\eta_j$ is in $\mathcal{M}$, we have $d\eta_j = 0$ and $\omega\eta_j = w_j - d\xi_j$. It follows that $[\omega][\eta_j] = 0$ in $H^{s+2r}(\mathcal{M})$. By the hypothesis of condition (2), $[\eta_j]$ has to be 0.

Now we consider $\theta\beta - \sum(\xi_j + \theta\eta_j)z_j$. It is in $I(N^*)$ as we assume θ and all $\xi_j + \theta\eta_j$ are in this ideal. Moreover,

$$d\left(\theta\beta - \sum(\xi_j + \theta\eta_j)z_j\right) = \omega\beta - \sum w_j z_j = 0.$$

Hence, the closed element $\theta\beta - \sum(\xi_j + \theta\eta_j)z_j$ in $I(N^*)$ has to be exact in $\mathcal{M}_\omega$, and we can write it as $d(\xi + \theta\lambda)$ with $\xi, \lambda \in \mathcal{M}$. Comparing the coefficient of θ , we have $\beta - \sum \eta_j z_j = -d\lambda$. As we have shown that z_j are all exact in $\mathcal{M}$, so is β . This proves the [claim](#).

Next we prove by induction that for any i , there exists a closed but nonexact $a_i \in \mathcal{M}^s$ such that $[a_i x_j] = 0$ in $H^{r+s}(\mathcal{M})$ for all $j \leq i$. The $i = 0$ case is immediate from condition (2), as $H^s(\mathcal{M}) \neq 0$ and any representative of a nontrivial cohomology class can be taken as a_0 .

Suppose we have found $a_0, \dots, a_{i-1}$ for $i > 0$. If $[a_{i-1} x_i] = 0$ in $H^{s+r}(\mathcal{M})$, we can simply set $a_i = a_{i-1}$. Otherwise, since $\omega: H^s(\mathcal{M}) \rightarrow H^{s+2r}(\mathcal{M})$ is an isomorphism according to condition (2), for all $j > i$ we can write $[a_{i-1} x_i x_j] = [\omega b_j]$ with $b_j \in \mathcal{M}^s$ closed. Let $\beta = a_{i-1} x_i - \sum_{j>i} b_j y_j$. Then

$$\omega\beta = \sum_{j \geq 1} a_{i-1} x_i x_j y_j - \sum_{j>i} \omega b_j y_j = \sum_{1 \leq j \leq i-1} (a_{i-1} x_i x_j) y_j + \sum_{j>i} (a_{i-1} x_i x_j - \omega b_j) y_j.$$

The last equation follows from $x_i^2 = 0$ as the degree of x_i is odd. Since $a_{i-1} x_j$ are all exact in $\mathcal{M}^{s+r}$ for $j < i$, so are the degree $s + 2r$ elements $a_{i-1} x_i x_j$. As $a_{i-1} x_i x_j - \omega b_j$ are also exact in $\mathcal{M}^{s+2r}$ for $j > i$ and all y_j are closed, we can apply the above [claim](#) to β . So β is exact and $[a_{i-1} x_i] = \sum_{j>i} [b_j y_j]$ in $H^*(\mathcal{M})$.

$[a_{i-1} x_i]$ is assumed to be nonzero in $H^*(\mathcal{M})$, so there is some $j > i$ such that $[b_j] \neq 0$. Write $\omega b_j = a_{i-1} x_i x_j + d\eta$ for some $\eta \in \mathcal{M}$. Then for any $l \leq i$,

$$\omega(b_j x_l) = a_{i-1} x_i x_j x_l + (d\eta)x_l = (a_{i-1} x_l x_i)x_j + (d\eta)x_l.$$

When $l < i$, by the hypothesis of induction $a_{i-1} x_l$ is exact in $\mathcal{M}$, and hence so is $a_{i-1} x_l x_i$. When $l = i$, $a_{i-1} x_l x_i = 0$. In either case $a_{i-1} x_l x_i$ and $d\eta$ are exact elements in $\mathcal{M}$ of degree $s + 2r$, and $x_j, x_l \in \mathcal{M}^s$ are closed. Hence, we can apply the above [claim](#) to $b_j x_l$ to deduce that they are all exact in $\mathcal{M}$ for $l \leq i$. Therefore, we can set $a_i = b_j$.

By induction we can find some nonzero $[a_k] \in H^s(\mathcal{M})$ such that $[a_k x_j] = 0$ in $H^{r+s}(\mathcal{M})$ for all j . It follows that $[a_k \omega] = 0$ in $H^{r+2s}(\mathcal{M})$. But by condition (2), $\omega: H^r(\mathcal{M}) \rightarrow H^{r+2s}(\mathcal{M})$ is an isomorphism, which is a contradiction. So $\mathcal{M}_\omega$ cannot be formal. $\square$

Remark 4.3 The condition that $|\omega| \equiv 2 \pmod{4}$ is necessary. For $|\omega| = 4$, let $\mathcal{M} = \Lambda\langle x, \xi \rangle$ be the minimal Sullivan model of $\mathbb{C}P^2$, where $|x| = 2$, $|\xi| = 5$, $dx = 0$, $d\xi = x^3$. Then $\omega = x^2$ induces an isomorphism $H^0(\mathcal{M}) \rightarrow H^4(\mathcal{M})$. However, $\mathcal{M}_\omega$ is formal because we can set $C^* = \langle x, \theta x - \xi \rangle$ and $N^* = \langle \theta \rangle$.

On a symplectic manifold (M, ω) , if $[\omega]$ is an integral cohomology class, there exists a circle bundle whose Euler class is $[\omega]$. This circle bundle is called the Boothby–Wang fibration.

Corollary 4.4 *Let (M, ω) be a connected symplectic manifold satisfying the hard Lefschetz property. Suppose $[\omega]$ is an integral and reducible cohomology class, ie there exists some $x_i, y_i \in H^1(M)$ such that $[\omega] = \sum x_i y_i$. Then the Boothby–Wang fibration of M is nonformal.*

Proof Let $\mathcal{M}$ be the minimal Sullivan model of M , then there are $\omega, x_i, y_i \in \mathcal{M}$ whose cohomology classes are same as the corresponding elements in $H^*(M)$, and they satisfy $\omega = \sum x_i y_i$.

Suppose $\dim M = 2n$. Then $\omega: H^{n-1}(\mathcal{M}) \rightarrow H^{n+1}(\mathcal{M})$ is isomorphic and $\omega: H^{n-2}(\mathcal{M}) \rightarrow H^n(\mathcal{M})$ is injective by the hard Lefschetz property. As $[\omega]^n \neq 0$, there exists some $[x_{i_1} y_{i_1} \cdots x_{i_n} y_{i_n}] \neq 0$. Hence $[x_{i_1} \cdots x_{i_{n-1}}] \neq 0$. Then we can apply [Theorem 4.2](#) to prove that $\mathcal{M}_\omega$ and the Boothby–Wang fibration of M are nonformal. $\square$

When the base manifold M is formal, the condition that $[\omega]$ is reducible in $H^*(M)$ is equivalent to having a reducible representative ω_0 in $\mathcal{M}$, the minimal Sullivan model of M . The reason is that $\mathcal{M}$ can be generated by some $C^* \oplus N^*$ and closed elements in $I(N^*)$ are all exact. So the ΛC^* part of ω_0 is also a representative of $[\omega]$. However, the sufficiency of this weakened condition for general base manifolds remains unknown.

Besides, it is uncertain whether the above corollary for symplectic manifolds still holds without the hard Lefschetz property.

Finally, when ω is irreducible, the minimal Sullivan model of $\mathcal{M}_\omega$ is slightly different. In this case, ω can be chosen a generator of $\mathcal{M}$, ie $\mathcal{M} = \Lambda V^*$, $V^* = \langle v_\alpha \rangle$ and $\omega = v_\alpha$ for some α . Let V^*/ω be a subspace spanned by all v_α except ω , where the order of v_α is preserved. Then let $\mathcal{M}/\omega = \Lambda(V^*/\omega)$ and $\Pi: \mathcal{M} \rightarrow \mathcal{M}/\omega$ be the natural projection. So $\Pi \circ d$ can be taken as the differential of $\mathcal{M}/\omega$.

Proposition 4.5 *$\mathcal{M}/\omega$ is a minimal Sullivan model of $\mathcal{M}_\omega$. The inclusion is a quasiisomorphism.*

Thus, the way of proving formality or finding obstructions for irreducible Euler classes may be quite different than the reducible case. Algebraically, Amann and Kapovitch [[1](#), page 21] constructed a formal S^3 -fibration with a nonformal base. Although this example is infinite dimensional, it may provide a potential insight into the construction of a formal sphere bundle over a nonformal manifold. Additionally, it is interesting to investigate the condition under which there exists a formal sphere bundle over a nonformal manifold.

References

- [1] **M Amann, V Kapovitch**, *On fibrations with formal elliptic fibers*, Adv. Math. 231 (2012) 2048–2068 [MR](#) [Zbl](#)
- [2] **I Biswas, M Fernández, V Muñoz, A Tralle**, *On formality of Sasakian manifolds*, J. Topol. 9 (2016) 161–180 [MR](#) [Zbl](#)

- [3] **G R Cavalcanti**, *The Lefschetz property, formality and blowing up in symplectic geometry*, Trans. Amer. Math. Soc. 359 (2007) 333–348 [MR](#) [Zbl](#)
- [4] **D Crowley**, **J Nordström**, *The rational homotopy type of $(n-1)$ -connected manifolds of dimension up to $5n-3$* , J. Topol. 13 (2020) 539–575 [MR](#) [Zbl](#)
- [5] **P Deligne**, **P Griffiths**, **J Morgan**, **D Sullivan**, *Real homotopy theory of Kähler manifolds*, Invent. Math. 29 (1975) 245–274 [MR](#) [Zbl](#)
- [6] **Y Félix**, **S Halperin**, **J-C Thomas**, *Rational homotopy theory*, Grad. Texts in Math. 205, Springer (2001) [MR](#) [Zbl](#)
- [7] **S Halperin**, *Lectures on minimal models*, Mém. Soc. Math. France 9-10, Soc. Math. France, Paris (1983) [MR](#) [Zbl](#)
- [8] **G Lupton**, *Variations on a conjecture of Halperin*, from “Homotopy and geometry”, Banach Center Publ. 45, Polish Acad. Sci. Inst. Math., Warsaw (1998) 115–135 [MR](#) [Zbl](#)
- [9] **T J Miller**, *On the formality of $(k-1)$ -connected compact manifolds of dimension less than or equal to $4k-2$* , Illinois J. Math. 23 (1979) 253–258 [MR](#)
- [10] **D Sullivan**, *Infinitesimal computations in topology*, Inst. Hautes Études Sci. Publ. Math. 47 (1977) 269–331 [MR](#) [Zbl](#)
- [11] **H L Tanaka**, **L-S Tseng**, *Odd sphere bundles, symplectic manifolds, and their intersection theory*, Camb. J. Math. 6 (2018) 213–266 [MR](#) [Zbl](#)
- [12] **J Zhou**, *A_∞ -minimal model on differential graded algebras*, preprint (2019) [arXiv 1904.10143](#)

School of Mathematics and Computer Sciences, Nanchang University
Nanchang, China

Beijing Institute of Mathematical Sciences and Applications
Beijing, China

jiaweizhou90@ncu.edu.cn

Received: 23 July 2023 Revised: 4 February 2024

A Quillen stability criterion for bounded cohomology

CARLOS DE LA CRUZ MENGUAL

TOBIAS HARTNICK

We provide a version of Quillen’s homological stability criterion for continuous bounded cohomology. In the companion paper we exploit this criterion to derive new bounded cohomological stability results for various families of classical groups.

[20J06](#); [55T25](#), [57T10](#)

1 Introduction

An infinite chain $(G_r)_{r \geq 0} = (G_0 < G_1 < G_2 < \cdots)$ of groups is called *homologically stable* if there exists a function $r : \mathbb{N} \rightarrow \mathbb{N}$ such that the respective inclusions induce isomorphisms

$$H_q(G_{r(q)}) \cong H_q(G_{r(q)+1}) \cong H_q(G_{r(q)+2}) \cong \cdots$$

in group homology for all $q \in \mathbb{N}$.¹ Any such function r is then called a *stability range* for the family. Homological stability (with a linear stability range) has been established for many families of classical linear algebraic groups, but also for various families of nonlinear groups of interest, such as mapping class groups or automorphism groups of free groups.

A particularly successful method to establish homological stability, employed for example by Van der Kallen [14], Harer [12], Hatcher and Vogtmann [13], Essert [9], and Sprehn and Wahl [22], is based on *Quillen’s stability criterion*, which can be stated as follows. Suppose that, for every $r \in \mathbb{N}$, we are given a Δ -complex $X(r)$ — also known as a semisimplicial set — endowed with a simplicial G_r -action, and two natural numbers $\gamma(r)$ and $\tau(r)$ with the following properties:

- (Q1) $X(r)$ is $\gamma(r)$ -acyclic, ie the reduced homology $\tilde{H}_\bullet(X(r))$ vanishes up to degree $\gamma(r)$.
- (Q2) $X(r)$ is $\tau(r)$ -transitive, ie there is only one G_r -orbit of l -simplices for $l \in \{0, \dots, \tau(r)\}$.
- (Q3) The complexes are $\tau(r)$ -compatible, ie the stabilizer of an l -simplex in $X(r)$ is isomorphic to G_{r-l-1} for every $l < \tau(r)$, and these isomorphisms are compatible with inclusions of stabilizers in the sense of condition (MQ3a) below.

Then $(G_r)_{r \geq 0}$ is homologically stable provided $\min\{\gamma(r), \tau(r)\} \rightarrow \infty$ as $r \rightarrow \infty$. In addition, a stability range can be computed explicitly from the functions γ and τ . See Quillen’s unpublished notes [21], or the more recent treatments by Bestvina [1] and Sprehn and Wahl [22] for details.

¹As usual in homological algebra, we enumerate starting from 0. Accordingly, we use the convention $0 \in \mathbb{N}$, and for any $k \in \mathbb{N}$, we denote by $[k]$ the $(k+1)$ -element set $\{0, \dots, k\} \subset \mathbb{N}$. We also set $[\infty] := \mathbb{N}$.

Our purpose here is to establish a similar stability criterion in *continuous bounded cohomology*. The theory of bounded cohomology for discrete groups goes back to Johnson, Trauber, and Gromov [11], and was later extended to the realm of topological groups by Burger and Monod [4]; we refer to the monographs by Frigerio [10] and Monod [16] for background and numerous applications.

By analogy with the classical situation, we say that an infinite family $G_0 < G_1 < G_2 < \dots$ of locally compact groups is *bc-stable* (short for *bounded-cohomologically stable*) with *stability range* $r: \mathbb{N} \rightarrow \mathbb{N}$ if, for every $q \geq 0$, the respective inclusions induce isomorphisms

$$H_b^q(G_{r(q)}) \cong H_b^q(G_{r(q)+1}) \cong H_b^q(G_{r(q)+2}) \cong \dots$$

in continuous bounded cohomology. We will give conditions analogous to (Q1)–(Q3) and show that these imply bc-stability. For technical reasons we will assume that G is either second-countable or discrete. This setting is broad enough to deal with both almost-connected Lie groups (as considered in our article [8]) and with uncountable discrete groups (as considered by Monod and Nariman [19]).

Compared to classical (co)homology, (continuous) bounded cohomology has a more functional analytic flavor—in fact, it can be defined as a derived functor in some suitable exact category of Banach modules; see work of Bühler [3]. While it is sometimes possible to extend classical cohomological results to this setting, this usually requires additional effort in order to take certain functional analytic peculiarities into account. These peculiarities are why we need to assume second-countability in the nondiscrete case.

1.1 General setting

Throughout this article let G denote a locally compact group. We will assume throughout that G is either second-countable or discrete. To formulate analogues of Quillen's conditions in the setting of bounded cohomology, we need to consider G -actions on complexes of a suitable type.

- In case G is discrete, the associated complexes will be, as in standard (co)homology, Δ -complexes with a simplicial G -action (semisimplicial G -sets); their sets of simplices will be endowed with the counting measure.
- In case G is second-countable and nondiscrete, we will need to replace Δ -complexes by a class that is better adapted to a measurable setting, namely, semisimplicial objects in the category of *Lebesgue G -spaces*. Here by a Lebesgue G -space we mean a standard Borel space with a Borel G -action and a G -invariant probability measure class.

In order to unify the treatment of the two cases we will refer to the corresponding complexes as *Lebesgue G -complexes* in either case. To a Lebesgue G -complex $X = (X_q)_{q \geq 0}$, we can associate a complex

$$0 \rightarrow \mathbb{R} \rightarrow L^\infty(X_0) \rightarrow L^\infty(X_1) \rightarrow L^\infty(X_2) \rightarrow \dots$$

of Banach spaces, where the codifferentials are given by alternating sums of dual face maps. Note that in the discrete case, this complex reduces to

$$0 \rightarrow \mathbb{R} \rightarrow \ell^\infty(X_0) \rightarrow \ell^\infty(X_1) \rightarrow \ell^\infty(X_2) \rightarrow \dots$$

We then say that X is *boundedly γ_0 -acyclic* if the cohomology of this complex vanishes up to degree γ_0 . We also say that X is *essentially τ_0 -transitive* if G acts essentially transitively (ie with a conull orbit) on q -simplices for $q \in \{0, 1, \dots, \tau_0\}$. In this case, we can choose simplices $o_0 \in X_0, \dots, o_{\tau_0} \in X_{\tau_0}$ such that o_i is a face of o_{i+1} for all $i \in \{0, \dots, \tau_0 - 1\}$, and such that the orbits of $o_0, \dots, o_{\tau_0}$ are conull in the respective spaces of simplices (see [Construction 3.3](#)). We then refer to $(o_0, \dots, o_{\tau_0})$ as a *generic flag* in X .

1.2 A spectral sequence in bounded cohomology

Our main tool will be the following spectral sequence, which connects the continuous bounded cohomology ring of a locally compact group G (always assumed to be second-countable or discrete) to those of the stabilizers of simplices of a generic flag in a highly boundedly acyclic and highly essentially transitive Lebesgue- G -complex.

Theorem A *Let X be a Lebesgue G -complex which is boundedly γ_0 -acyclic and essentially τ_0 -transitive for some $\gamma_0, \tau_0 \in \mathbb{N} \cup \{\infty\}$. Moreover, let $(o_0, \dots, o_{\tau_0})$ be a generic flag in X with stabilizers $H_i := \text{Stab}_G(o_i)$ for $i \in \{0, \dots, \tau_0\}$, and set $H_{-1} := G$.*

- (1) *For every $p \in \{0, \dots, \tau_0 - 1\}$ and $i \in \{0, \dots, p + 1\}$, there exists an element $w_{p,i} \in G$ such that $\delta_i(o_{p+1}) = w_{p,i}^{-1} o_p$, that induces an injective homomorphism*

$$\text{Int}(w_{p,i}): H_{p+1} \rightarrow H_p, \quad h \mapsto w_{p,i} h w_{p,i}^{-1}.$$

- (2) *There exists a first-quadrant spectral sequence $E_{\bullet, \bullet}^{\bullet}$ such that:*

- (i) $E_{\infty}^t = 0$ for every $t \in \{0, \dots, \gamma_0 + 1\}$.
- (ii) $E_1^{p,q} = H_b^q(H_{p-1})$ for all $p \in \{0, \dots, \tau_0 + 1\}$ and $q \geq 0$.
- (iii) $E_2^{p,0} = 0$ for all $p \in \{0, \dots, \tau_0 + 1\}$.
- (iv) *The first page differentials $d_1^{0,q}: H_b^q(G) \rightarrow H_b^q(H_0)$ are induced by the inclusion $H_0 \hookrightarrow G$ for all $q \geq 0$.*
- (v) *The first page differentials $d_1^{p,q}: H_b^q(H_{p-1}) \rightarrow H_b^q(H_p)$ are the alternating sums*

$$d_1^{p,q} = \sum_{i=0}^p (-1)^i H_b^q(\text{Int}(w_{p-1,i}))$$

for all $p \in \{1, \dots, \tau_0\}$, $q \geq 0$, and any choice of elements $w_{p,i} \in G$ as in (1).

Remark 1.1 A version of this spectral sequence for product complexes (complexes where $X_r = X_0^{r+1}$ with the usual differential) of locally compact second-countable (lcsc) groups was first studied by Monod in [\[17\]](#). The version for uncountable discrete groups played a key role in work by Monod and Nariman [\[19\]](#) — specifically their Theorem 3 — on the computation of the bounded cohomology of various homeomorphism and diffeomorphism groups of manifolds.

1.3 Measurable Quillen families

Assume that we are given an infinite family $(G_r)_{r \geq 0} = (G_0 < G_1 < G_2 < \dots)$ of locally compact groups which are either all second-countable or all discrete, and a Lebesgue G_r -complex $X(r)$ for every $r \geq 0$. We can then define counterparts to the Quillen conditions (Q1) and (Q2) as follows. We assume that we are given functions $\gamma, \tau: \mathbb{N} \rightarrow \mathbb{N} \cup \{\pm\infty\}$ such that

(MQ1) $X(r)$ is boundedly $\gamma(r)$ -acyclic,

(MQ2) $X(r)$ is essentially $\tau(r)$ -transitive.

These assumptions imply by [Theorem A](#) that, for each $r \geq 0$, we obtain elements $w_{r,p,i} \in G_r$ for all $p \in \{0, \dots, \tau(r)\}$ and $i \in \{0, \dots, p+1\}$ as in [Theorem A\(1\)](#), and a spectral sequence as in (2), relating the continuous bounded cohomology of G_r to the continuous bounded cohomology of the stabilizers $G_r =: H_{r,-1} > H_{r,0} > \dots > H_{r,\tau(r)}$ of a generic flag $(o_{r,0}, \dots, o_{r,\tau(r)})$ in $X(r)$. To obtain a stability result in the vein of Quillen's, we need to relate the bounded cohomology of these stabilizers to the one of the previous groups in the sequence. We offer three different compatibility conditions which are useful in different situations and are all sufficient to obtain bc-stability:

(MQ3a) For all $r \geq 0$ and $p \in \{-1, 0, \dots, \tau(r)\}$, we have $H_{r,p} = G_{r-p-1}$ (we use the convention that G_k is the trivial group for $k < 0$), and the conjugation $\text{Int}(w_{r,p,i}): H_{r,p+1} \rightarrow H_{r,p}$ equals the inclusion $\iota_{r-p-2}: G_{r-p-2} \hookrightarrow G_{r-p-1}$ for all r , all $p \leq \tau(r) - 1$, and all i .

(MQ3b) For all $r \geq 0$ and $p \in \{-1, 0, \dots, \tau(r)\}$, there is an epimorphism $\pi_{r,p}: H_{r,p} \twoheadrightarrow G_{r-p-1}$ with amenable kernel such that

$$\begin{array}{ccc} H_{r,p+1} & \xrightarrow{\text{Int}(w_{r,p,i})} & H_{r,p} \\ \pi_{r,p+1} \downarrow & & \downarrow \pi_{r,p} \\ G_{r-p-2} & \xrightarrow{\iota_{r-p-2}} & G_{r-p-1} \end{array}$$

is a commutative diagram for all r , all $p \leq \tau(r) - 1$, and all i .

(MQ3c) For all $r \geq 0$ and $p \in \{-1, 0, \dots, \tau(r)\}$, there is an epimorphism $\pi_{r,p}: H_{r,p} \twoheadrightarrow G_{r-p-1}$ with amenable kernel and a continuous homomorphic section $\sigma_{r,p}$ of $\pi_{r,p}$ such that

$$\begin{array}{ccc} H_{r,p+1} & \xrightarrow{\text{Int}(w_{r,p,i})} & H_{r,p} \\ \sigma_{r,p+1} \uparrow & & \downarrow \pi_{r,p} \\ G_{r-p-2} & \xrightarrow{\iota_{r-p-2}} & G_{r-p-1} \end{array}$$

is a commutative diagram for all r , all $p \leq \tau(r) - 1$, and all i .

Here, (MQ3a) is Quillen's original condition; indeed, the equality $\text{Int}(w_{r,p,i}) = \iota_{r-p-2}$ means that the element $w_{r,p,i} \in G_r$ centralizes $G_{r-p-2} < G_r$. The conditions (MQ3b) and (MQ3c) are two different relaxations, which take advantage of the fact that continuous surjections with amenable kernel induce

isomorphisms in continuous bounded cohomology. The latter is the most technical, but also most useful condition, as witnessed by [Example 1.4](#) below.

Definition 1.2 We say that $(G_r, X(r))_{r \geq 0}$ is a *measurable (γ, τ) -Quillen family* if conditions (MQ1), (MQ2), and one of (MQ3a), (MQ3b), or (MQ3c) hold.

The following is our main qualitative result:

Theorem B (Quillen stability for continuous bounded cohomology) *Assume that $(G_r, X(r))_{r \geq 0}$ is a measurable (γ, τ) -Quillen family. If both γ and τ are proper, then $(G_r)_{r \geq 0}$ is bc-stable.*

Theorem B will be established by comparing the various spectral sequences associated to the Lebesgue complexes $X(r)$ by means of [Theorem A](#). This will actually provide a quantitative version of [Theorem B](#): we will obtain an explicit stability range which depends only on the functions γ and τ . This quantitative version actually also makes sense for finite Quillen families. We will state the most general version of our main theorem as [Theorem 4.6](#) below.

For *product complexes*, ie complexes of the form $X(r)_q = X(r)_0^{q+1}$ with the forgetful face maps, [Theorem B](#) and its quantitative version were essentially established by Monod in [\[17\]](#), and our proof is an extension of Monod's original proof. As we will explain after [Example 1.4](#), our more general version is crucial if one wants to generalize Monod's results concerning bc-stability of general linear groups to other classes of classical groups.

Remark 1.3 The isomorphisms produced by [Theorem B](#) are generally not isometric with respect to the Gromov norms. This fact is evidenced, for instance, by the norms of the bounded Borel classes in complex simple Lie groups of classical type, which were computed in works by Bucher, Burger, and Iozzi [\[2\]](#) and the first author [\[6; 7\]](#).

1.4 The discrete case

Recall that [Theorem B](#) applies to the case in which $(G_r)_{r \geq 0}$ are discrete groups. In this case, the statement simplifies considerably, and we record this special case for ease of reference. The sequence $(X(r))_{r \geq 0}$ is just a family of Δ -complexes, and the sets of simplices of $X(r)$ are all endowed with the respective counting measures. To prove rational homological stability of $(G_r)_{r \geq 0}$ using Quillen's criterion, it suffices to show that G_r acts $\tau(r)$ -transitively on $X(r)$ with stabilizers $H_{r,p} = G_{r-p-1}$ and that $X(r)$ is rationally $\gamma(r)$ -acyclic for two proper functions $\tau, \gamma: \mathbb{N} \rightarrow \mathbb{N} \cup \{\pm\infty\}$. One way to do this is by exhibiting a partial contracting chain homotopy h_* for each of the complexes

$$\cdots \rightarrow \mathbb{Q} X(r)_1 \xrightarrow{d_0} \mathbb{Q} X(r)_0 \xrightarrow{d_{-1}} \mathbb{Q} \rightarrow 0$$

so that $d_k \circ h_k + h_{k-1} \circ d_{k-1} = 1$. These chain homotopies will then induce a partial dual contracting chain homotopy of the dual cocomplexes

$$0 \rightarrow \mathbb{R} \xrightarrow{d^{-1}} \ell^\infty(X(r)_0) \xrightarrow{d^0} \ell^\infty(X(r)_1) \rightarrow \cdots,$$

provided that for every $k \in \mathbb{N}$ there is a constant $C_k > 0$ such that for every simplex $\sigma \in X_k$ the ℓ^1 -norm $\|h_k(\sigma)\|_1$ is bounded by C_k , ie

$$h_k(\sigma) = \sum_{i=1}^{n_\sigma} \alpha_i \tau_i \implies \sum_{i=1}^{n_\sigma} |\alpha_i| \leq C_k.$$

In this case we refer to h_* as a *rational ℓ^1 -homotopy*.

Corollary C *Let $(G_r)_{r \geq 0}$ be a family of discrete groups. For every $r \geq 0$, let $X(r)$ be a $\gamma(r)$ -acyclic Δ -complex with a $\tau(r)$ -transitive G_r -action and stabilizers $H_{r,q} = G_{r-q-1}$ which are compatible in the sense of (MQ3a).*

- (i) *If $\gamma, \tau: \mathbb{N} \rightarrow \mathbb{N} \cup \{\pm\infty\}$ are proper, then $(G_r)_{r \geq 0}$ is homologically stable.*
- (ii) *If, moreover, rational $\gamma(r)$ -acyclicity of $X(r)$ is witnessed by a rational ℓ^1 -homotopy, then $(G_r)_{r \geq 0}$ is also bc-stable.*

Again, the stability range can be computed explicitly in terms of γ and τ ; see [Theorem 4.6](#). **Corollary C** provides a clear strategy to establish bc-stability for all families of countable groups for which homological stability has been established using Quillen's method: One just has to check whether the underlying homotopy is rationally ℓ^1 . Unfortunately, the homotopies in question are usually only given implicitly. We thus leave it to future work to determine for which countable groups an explicit rational ℓ^1 -homotopy can be constructed.

1.5 The case of reductive Lie groups

Our main interest in establishing [Theorem B](#) in its present generality was to provide a framework for establishing bc-stability of the classical families of reductive Lie groups. For all of these families, bc-stability is predicted by the (notoriously open) isomorphism conjecture in continuous bounded cohomology. However, prior to this work, this predicted stability had only been established for general and special linear groups over $\mathbb{R}$ and $\mathbb{C}$ by Monod [\[17\]](#) using a special case of [Theorem B](#). Let us briefly explain how the results for general linear groups² fit into our present context:

Example 1.4 (general linear groups) Let $G_r := \mathrm{GL}_r(\mathbb{K})$ with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$. These groups constitute an infinite family $(G_0 < G_1 < G_2 < \dots)$ of lscg groups with the block inclusions

$$(1-1) \quad \iota_r: G_r \hookrightarrow G_{r+1}, \quad A \mapsto 1 \times A := \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}.$$

Let $X(r)_q := \mathbb{P}(\mathbb{K}^r)^{q+1}$, where $\mathbb{P}(\mathbb{K}^r)$ denotes the projective space of $\mathbb{K}^r$ and the face maps are the usual forgetful maps. If we equip each of these spaces with its canonical Lebesgue measure class, then each $X(r)$ becomes a boundedly acyclic Lebesgue $G(r)$ -complex. Moreover, the action of G_r on

²As explained by Monod in [\[18, Note before Lemma 10\]](#), the stability range given in [\[17\]](#) is not correct due to an inaccuracy in the induction step. We use this opportunity to state the correct range; see [\[6, Sections 6.3 and 6.4\]](#) for details. Also, the quaternionic case is absent from [\[17\]](#), but it does not require any new ideas.

$X(r)$ is essentially r -transitive. If $e_1, \dots, e_r$ denotes the standard basis of $\mathbb{K}^r$ and $e_{r+1} = \sum_{i=1}^r e_i$, then a generic flag in $X(r)$ is given by $(o_0, \dots, o_r)$, where $o_q := ([e_1], \dots, [e_{q+1}])$. The corresponding stabilizers $H_{r,q} = \text{Stab}_{G_r}(o_q)$ are

$$H_{r,q} = \left\{ \begin{pmatrix} D & V \\ 0 & A \end{pmatrix} \mid D \in G_{q+1} \text{ diagonal}, A \in G_{r-q-1}, V \in M_{(q+1) \times (r-q-1)}(\mathbb{K}) \right\}$$

for $q \in [r-1]$, and $H_{r,r} = (\mathbb{K}^\times) \cdot I_r$; we also set $H_{r,-1} = G_r$. For $q \in \{0, \dots, r-2\}$ and $i \in \{0, \dots, q+1\}$, we let $w_{r,q,i} \in G_r$ be the matrix that permutes the standard basis vectors via the cycle $(q+2, q+1, \dots, i+1)$, and for $q = r-1$ we fix an arbitrary choice of $w_{r,r-1,i} \in G_r$ according to [Theorem A\(1\)](#). With this notation, condition (MQ3a) is violated, and while the obvious projections $\pi_{r,q}: H_{r,q} \rightarrow G_{r-q-1}$ do have amenable kernel, they do not satisfy condition (MQ3b) since the corresponding diagram does not commute. In turn, condition (MQ3c) is satisfied with the homomorphic sections

$$\sigma_{r,q}: G_{r-q-1} \rightarrow H_{r,q}, \quad A \mapsto I_{q+1} \times A,$$

of $\pi_{r,q}$. This is the reason why we insist on this slightly technical compatibility condition.

We may now deduce that $(G_r, X(r))_{r \geq 0}$ is a measurable (∞, r) -Quillen family, and hence (G_r) is bc-stable. An explicit stability range can be read off from [Theorem 4.6](#) below. In the notation of that theorem, we have $q_0 = 2$, $\gamma(r) = \infty$, and $\tau(r) = r$ (since $H_b^2(G_r) = 0$ for all $r \geq 0$; this follows from work of Burger and Monod [\[5\]](#) and Monod [\[16, Corollary 8.8.6 and Example 9.9.3\]](#)), and hence

$$\min\{\tilde{\gamma}(q, r), \tilde{\tau}(q, r) - 1\} = \min_{j=q_0}^q \{\tau(r+1-2(q-j)) - j\} - 1 = r - 2q + q_0 = r - (2q - 2).$$

We deduce that for every $q \geq q_0 + 1 = 3$, the inclusions ι_r induce isomorphisms and injections

$$\dots \cong H_b^q(\text{GL}_{2q-1}(\mathbb{K})) \cong H_b^q(\text{GL}_{2q-2}(\mathbb{K})) \hookrightarrow H_b^q(\text{GL}_{2q-3}(\mathbb{K})) \hookrightarrow H_b^q(\text{GL}_{2q-4}(\mathbb{K})).$$

For the general linear groups, we thus obtain bc-stability with slope two.

In [Example 1.4](#), the complexes $X(r)$ can be chosen to be product complexes, since GL_r acts essentially r -transitively on projective space. It is known that reductive Lie groups whose semisimple part is of type other than A_n do not admit highly transitive actions on generalized flag manifolds; see work by Popov [\[20\]](#). Thus the product version of [Theorem B](#) is insufficient to establish bc-stability for other classical groups. This was our main motivation in stating [Theorem B](#) in its present form.

Example 1.5 (symplectic, orthogonal, and unitary groups) For $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and $d \in \mathbb{N}$, let $(G_r)_{r \geq 0}$ denote one of the following classical families of Lie groups:

$$\begin{aligned} \{1\} &< \text{Sp}_2(\mathbb{K}) < \text{Sp}_4(\mathbb{K}) < \dots < \text{Sp}_{2r}(\mathbb{K}) < \dots, \\ O_{d,0}(\mathbb{K}) &< O_{d+1,1}(\mathbb{K}) < O_{d+2,2}(\mathbb{K}) < \dots < O_{d+r,r}(\mathbb{K}) < \dots, \\ U_{d,0}(\mathbb{C}) &< U_{d+1,1}(\mathbb{C}) < U_{d+2,2}(\mathbb{C}) < \dots < U_{d+r,r}(\mathbb{C}) < \dots. \end{aligned}$$

In our article [\[8\]](#), we apply [Theorem B](#) to establish bc-stability for all of these families (G_r) . Our proof is based on the fact that all of the groups above arise as automorphism groups of sesquilinear forms.

With each of the groups G_r above, one can thus associate a corresponding *Stiefel complex* in the sense of Vogtmann [23], and as in the case of the general linear groups, it is easy to see that the actions are $(r-1)$ -transitive and compatible (in the sense of (MQ3c)). In order to establish bc-stability, one thus only has to show that the Stiefel complex X_r is measurably $\gamma(r)$ -acyclic for some proper function γ . This is established in our article [8] using probabilistic methods, leading to an acyclicity range of the form $\gamma(r) \sim \log_2(r)$.

1.6 The need for finite measurable Quillen families

Theorem B does not apply directly to the families of special linear (or special orthogonal or special unitary) groups. The problem is that while $G_r = \mathrm{SL}_r(\mathbb{K})$ acts highly transitively on the product complex $X(r)$ defined as in [Example 1.4](#), the quotient of the stabilizer $H_{r,q}$ by its radical is isomorphic to $\mathrm{GL}_{r-q-1}(\mathbb{K})$ instead of $\mathrm{SL}_{r-q-1}(\mathbb{K})$.

To deal with this problem, we introduce in [Section 4](#) below the notion of a finite measurable Quillen family. It then turns out that (for $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$), the family

$$\{1\} < \mathrm{GL}_1(\mathbb{K}) < \mathrm{GL}_2(\mathbb{K}) < \cdots < \mathrm{GL}_{R-1}(\mathbb{K}) < \mathrm{SL}_R(\mathbb{K})$$

is a finite measurable Quillen family in this sense, and our main [Theorem 4.6](#) also applies. Assuming the theorem and its notation, we recover Monod's results for SL_n .

Example 1.6 (special linear groups) Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ and $R \geq 2$. We set $G_r := \mathrm{GL}_r(\mathbb{K})$ for any $r \in [R-1]$ and $G_R := \mathrm{SL}_R(\mathbb{K})$. The inclusions $\iota_r : G_r \hookrightarrow G_{r+1}$ are defined as in [\(1-1\)](#) for $r < R-1$, and as

$$A \mapsto (\det A)^{-1} \times A$$

for $r = R-1$. We consider the same acyclic product complexes $X(r)$ as in [Example 1.4](#); then the action of G_r on $X(r)$ is $\tau(r)$ -transitive, where $\tau(r) = r$ if $r < R$ and $\tau(R) = R-1$. The group elements $w_{r,q,i} \in G_r$ can be chosen as in [Example 1.4](#) for $r < R$, but for $r = R$ one requires a sign adjustment to guarantee that the matrices lie in G_R . In the notation of [Theorem 4.6](#) the initial condition is again given by $q_0 = 2$ if $R \geq 3$ and $q_0 = 1$ if $R = 2$. Assume first that $R \geq 3$. Then

$$\begin{aligned} \min\{\tilde{\gamma}(q, R-1), \tilde{\tau}(q, R-1) - 1\} &= \min_{j=q_0}^q \{\tau(R-2(q-j)) - j\} - 1 \\ &= \min \left\{ \min_{j=q_0}^{q-1} \{R-2q+j\}, R-1-q \right\} - 1 \\ &= \min\{R-2q+2, R-q-1\} - 1 = R-(2q-2), \end{aligned}$$

and, hence the map ι_R induces

- (i) an isomorphism $H_b^q(\mathrm{SL}_R(\mathbb{K})) \cong H_b^q(\mathrm{GL}_{R-1}(\mathbb{K}))$ if $R \geq 2q-2$,
- (ii) an injection $H_b^q(\mathrm{SL}_R(\mathbb{K})) \hookrightarrow H_b^q(\mathrm{GL}_{R-1}(\mathbb{K}))$ if $R \geq 2q-4$.

Combining this information with [Example 1.4](#), we see that for $q \geq 3$ the natural inclusions induce isomorphisms/injections³

$$\cdots \cong H_b^q(\mathrm{SL}_{2q}(\mathbb{K})) \cong H_b^q(\mathrm{SL}_{2q-1}(\mathbb{K})) \cong H_b^q(\mathrm{SL}_{2q-2}(\mathbb{K})) \hookrightarrow H_b^q(\mathrm{GL}_{2q-3}(\mathbb{K})).$$

The case $q = 2$, not covered by the discussion above, is well known (see [\[5; 16, Example 9.9.3\]](#)):

$$\cdots \cong H_b^2(\mathrm{SL}_4(\mathbb{K})) \cong H_b^2(\mathrm{SL}_3(\mathbb{K})) \cong \{0\} \hookrightarrow H_b^2(\mathrm{SL}_2(\mathbb{K})).$$

Here the inclusion is an isomorphism for $\mathbb{K} = \mathbb{C}$ or $\mathbb{H}$, but not for $\mathbb{K} = \mathbb{R}$.

Acknowledgments We thank Marc Burger and Nicolas Monod for insightful conversations, and the referee for valuable feedback on the first version of this article. This work was completed during a postdoctoral fellowship of De la Cruz Mengual at the Weizmann Institute of Science, Israel. He was supported by the Swiss National Science Foundation, grants 169106 and 188010. Hartnick was supported by the Deutsche Forschungsgemeinschaft, grant HA 8094/1-1 within the Schwerpunktprogramm SPP 2026 (Geometry at Infinity).

2 Spectral sequences associated to Lebesgue complexes

Throughout this section, G denotes a locally compact group. We will assume that G is either second-countable or discrete. The goal of this section is to associate a first-quadrant spectral sequence in continuous bounded cohomology with every partially boundedly acyclic Lebesgue G -complex. We explain the terminology used in this statement and recall the necessary background on bounded cohomology, and refer the reader to Monod's book [\[16\]](#) for further details.

2.1 Coefficient G -modules and Lebesgue G -spaces

Following [\[16\]](#), we discuss the class of modules which can be used as coefficients in continuous bounded cohomology. In order to be able to deal with uncountable discrete groups, we have to modify the definitions slightly.

A *Banach G -module* is a Banach space over the field of real numbers equipped with a G -action by linear isometries, and a *G -morphism* is a G -equivariant bounded (but not necessarily isometric) operator between Banach G -modules. A Banach G -module is said to be *separable* if the underlying Banach space is separable, and *continuous* if the G -action is jointly continuous. Given a Banach G -module E , we denote by $E^\#$ the topological dual of E , considered as a Banach G -module endowed with the contragredient G -action. Note that $E^\#$ need not be separable nor continuous, even if E has either of those properties. The following notion is due to Monod [\[16\]](#) in the second-countable case:

Definition 2.1 (i) If G is nondiscrete and second-countable, then a *coefficient G -module* is a pair $(E, E^\#)$, where E is a separable, continuous Banach G -module.

³Note that the stability range given by Monod [\[17\]](#) is not correct, since it uses the inaccurate stability range for $\mathrm{GL}_r(\mathbb{K})$.

(ii) If G is discrete, then a *coefficient G -module* is a pair $(E, E^\#)$, where E is a continuous Banach G -module (not necessarily separable).

In either case, a *morphism* $(E, E^\#) \rightarrow (F, F^\#)$ of coefficient G -modules is a pair $(\Phi, \Phi^\#)$, where $\Phi: E \rightarrow F$ is a bounded operator with a G -equivariant dual $\Phi^\#: F^\# \rightarrow E^\#$. We write $\mathbf{Coef}_G$ for the category of coefficient G -modules.

Remark 2.2 If we consider a coefficient G -module (E^b, E) , then the only role of the specified predual E^b is to define a weak-* topology on E . We will thus usually drop the predual E^b from notation and simply refer to E as a coefficient G -module. With this abuse of notation, a morphism between coefficient G -modules E and F is simply a G -equivariant map $E \rightarrow F$, which is continuous with respect to the weak-* topologies defined by the omitted preduals.

The next notion allows us to produce examples of coefficient G -modules:

Definition 2.3 (i) If G is nondiscrete and second-countable, we say that X is a *Lebesgue G -space* if it is a standard Borel space endowed with a Borel G -action and a G -invariant Borel measure class. For economy, we will say that the pair (X, μ) is a Lebesgue G -space whenever X is a Lebesgue G -space and μ is a fixed G -quasi-invariant probability measure within the measure class.

(ii) If G is discrete, we will say that X is a *Lebesgue G -space* if it is a G -set, endowed with the measure class of its counting measure.

We denote by $\mathbf{Leb}_G$ the category whose objects are Lebesgue G -spaces and G -equivariant measure-class-preserving Borel maps.

Example 2.4 The space G^n is a Lebesgue G -space (with respect to the diagonal G -action) for every $n \in \mathbb{N}$. For this to be the case it is crucial that we do not assume a Lebesgue- G -space to be countably separated if G is an uncountable discrete group.

Example 2.5 Let X be a Lebesgue G -space. If G is nondiscrete, then we denote by μ a fixed probability measure in the given measure class on X , and if G is discrete, then we denote by μ the counting measure on X . The group G acts on $L^1(X, \mu)$ via

$$g \cdot \phi(x) := \rho_\mu(g^{-1}, x) \cdot \phi(g^{-1}x),$$

where $\rho_\mu(g, x) := d(g_*\mu)/d\mu(x)$ denotes the Radon–Nikodym cocycle of μ . In the nondiscrete case this action is continuous, as proven by Bühler [3, Appendix D]. Moreover, $L^1(X, \mu)$ is separable in this case (but not necessarily in the discrete case). Either way, the contragredient module is then given by $L^1(X, \mu)^\# = L^\infty(X)$ with action given by $g \cdot f(x) := f(g^{-1}x)$, and our definitions have been set up in such a way that the pair $(L^1(X, \mu), L^\infty(X))$ is a coefficient G -module which up to isomorphism does not depend on the choice of μ .

Remark 2.6 The assignment $L^\infty(-): \mathbf{Leb}_G^{\text{op}} \rightarrow \mathbf{Coef}_G$ is functorial. In fact, if $T: Y \rightarrow X$ is a morphism of Lebesgue G -spaces, then the induced operator

$$L^\infty(T): L^\infty(X) \rightarrow L^\infty(Y), \quad L^\infty(T)(\phi) := \phi \circ T,$$

is a morphism of coefficient G -modules. Choosing a G -quasi-invariant measure ν on Y as above and letting $\mu := T_*\nu$, the predual operator to $L^\infty(T)$ is

$$L^1(T): L^1(Y, \nu) \rightarrow L^1(X, \mu), \quad (L^1(T)(\psi))(x) := d(T_*(\psi \cdot \nu))/d\mu.$$

We will need the following generalization:

Construction 2.7 Let X be a Lebesgue G -space and let E be a coefficient G -module. We equip E with the Borel σ -algebra associated with the weak- $*$ topology on E . Given $n \in \mathbb{N}$ we then define

$$L^\infty(X; E) := \{\phi: X \rightarrow E \mid \phi \text{ is measurable and essentially bounded}\} / \sim,$$

where $\sim$ denotes almost-everywhere equality. This defines a Banach G -module when equipped with the essential supremum norm and the G -action given by the formula $g.f(x) = g.f(g^{-1}x)$. In particular, for all $n \in \mathbb{N}$ we obtain a Banach G -module

$$L^\infty(G^n; E) := \{\phi: G^n \rightarrow E \mid \phi \text{ is measurable and essentially bounded}\} / \sim.$$

Remark 2.8 Our definition of coefficient G -module is made in such a way that if X is a Lebesgue G -space and E is a coefficient G -module, then $L^\infty(X; E)$ is again a coefficient G -module. Indeed, it is the dual of the Bochner space $L^1(X; E^b)$, where E^b denotes the underlying predual of E . Note that in the discrete case the module $L^1(X; E^b)$ need not be separable.

The use of the weak- $*$ Borel structure of E in the definition of $L^\infty(X; E)$ is necessary since $L^\infty(X; E)$ is not well defined if measurability is taken with respect to its norm-structure. If G is discrete, measurability is a void assumption and $L^\infty(X; E)$ can in fact be defined for *any* Banach G -module E .

As in the case of the trivial coefficient module $\mathbb{R}$, one shows that the assignment $X \mapsto L^\infty(X; E)$ is functorial for any fixed coefficient module E .

2.2 Background on continuous bounded cohomology

We collect the definition and a series of useful facts on continuous bounded cohomology.

Definition 2.9 If E is Banach G -module and $n \in \mathbb{N}$, then we denote by $C_b(G^{n+1}; E)$ the Banach space of continuous bounded functions $f: G^{n+1} \rightarrow E$ with the supremum norm. The group G acts diagonally on G^{n+1} and hence by isometries on $C_b(G^{n+1}; E)$ via the left-regular representation, and we denote by $C_b(G^{n+1}; E)^G$ the subspace of G -invariant functions. Then the *continuous bounded cohomology* of G with coefficients in E is defined as

$$H_b^\bullet(G; E) := H^\bullet(C_b(G^{\bullet+1}; E)^G).$$

If $E = \mathbb{R}$ with trivial G -action, we will drop E from the notation and write $H_b^\bullet(G) := H_b^\bullet(G; \mathbb{R})$.

Proposition 2.10 *If E is a coefficient G -module, then*

$$H_b^\bullet(G; E) = H^\bullet(0 \rightarrow L^\infty(G; E)^G \rightarrow L^\infty(G^2; E)^G \rightarrow L^\infty(G^3; E)^G \rightarrow \cdots).$$

Proof In the discrete case, this is just the definition, since $L^\infty(G^{n+1}; E)^G = \ell^\infty(G^{n+1}; E)^G = C_b(G^{n+1}; E)^G$. In the second-countable case this is nontrivial, and appears as [16, Proposition 7.5.1] in Monod's book. $\square$

To provide a more general class of resolutions, we need to recall the notion of an amenable Lebesgue G -space. In the discrete case, its definition is given in [10, Section 4.9] in Frigerio's book. In the nondiscrete case, the definition is more technical; see [16, Section 5.3].

Proposition 2.11 [10, Theorem 4.23; 16, Theorem 7.5.3] *If E is a coefficient G -module and S is an amenable Lebesgue G -space, then*

$$H_b^\bullet(G; E) = H^\bullet(0 \rightarrow L^\infty(S; E)^G \rightarrow L^\infty(S^2; E)^G \rightarrow L^\infty(S^3; E)^G \rightarrow \cdots).$$

Example 2.12 The group G acts amenably on $S := G$ via left-multiplication; see [16, Example 2.1.2.(i)]. If $H < G$ is a closed subgroup and S is an amenable G -space, then it is also an amenable H -space; see [16, Example 2.1.2.(iv)]. In particular, G is an amenable H -space, and we deduce from Proposition 2.11 that for any coefficient H -module E we have

$$(2-1) \quad H_b^\bullet(H; E) = H^\bullet(0 \rightarrow L^\infty(G; E)^H \rightarrow L^\infty(G^2; E)^H \rightarrow L^\infty(G^3; E)^H \rightarrow \cdots).$$

The following lemma is an essential technical tool for us. Let **Vect** denote the category of real vector spaces. We say that a sequence $0 \rightarrow E^0 \rightarrow E^1 \rightarrow E^2 \rightarrow \cdots$ of Banach G -modules is a *cochain complex* (resp. *exact*) if the underlying sequence of vector spaces has the corresponding property.

Lemma 2.13 *For any $n \in \mathbb{N}$, the functor $\mathbf{Coef}_G \rightarrow \mathbf{Vect}$ given by $E \mapsto L^\infty(G^n, E)^G$ is exact.*

Proof For discrete groups, this is contained in [16, Lemma 8.2.4], and, for nondiscrete lcsc groups, in [16, Lemma 8.2.5]. $\square$

More concretely, this lemma states that if $0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0$ is a short exact sequence of coefficient G -modules, then

$$0 \rightarrow L^\infty(G^n; A)^G \rightarrow L^\infty(G^n; B)^G \rightarrow L^\infty(G^n; C)^G \rightarrow 0$$

is an exact sequence of vector spaces.

2.3 Lebesgue G -complexes

From Lebesgue G -spaces, we can build associated complexes using the following notion:

Definition 2.14 (see Weibel [24, Definition 8.1.9]) A *semisimplicial object* X over a concrete category $\mathcal{C}$ is a sequence of objects $(X_k)_{k \in \mathbb{N}}$, together with morphisms $\delta_{i,k}: X_{k+1} \rightarrow X_k$ for all k and $i \in [k]$, called *face maps*, such that

$$(2-2) \quad \delta_{i,k-1} \circ \delta_{j,k} = \delta_{j-1,k-1} \circ \delta_{i,k} \quad \text{whenever } i < j.$$

If X and Y are semisimplicial objects over a category $\mathcal{C}$, then a *semisimplicial morphism* $f: X \rightarrow Y$ is a collection $(f_k: X_k \rightarrow Y_k)_{k \in \mathbb{N}}$ of morphisms such that $f_k \circ \delta_{i,k-1} = \delta_{i,k} \circ f_{k+1}$ for all $k \in \mathbb{N}$ and all $i \in [k+1]$.

We will usually omit the index k in $\delta_{i,k}$, writing simply δ_i for all i^{th} face maps.

Definition 2.15 A semisimplicial object in $\mathbf{Leb}_G$ will be called a *Lebesgue G -complex*.

Remark 2.16 If X is a Lebesgue G -space, then we refer to points in $\bigsqcup_{k \in \mathbb{N}} X_k$ as *simplices* and to points of X_k as k -*simplices* of X . Given two simplices $x, y \in X$, we say that x is a *face* of y , denoted by $x < y$, if x can be obtained by applying finitely many face maps to y . A (finite or infinite) sequence $x_0 < x_1 < x_2 < \dots$ of simplices is called a *flag*.

If $x, y \in X$ are such that $x < y$, then the reversed inclusion $\text{Stab}_G(y) < \text{Stab}_G(x)$ of stabilizer subgroups holds, and hence every flag $x_0 < x_1 < x_2 < \dots$ gives rise to a chain of closed subgroups

$$G > \text{Stab}_G(x_0) > \text{Stab}_G(x_1) > \text{Stab}_G(x_2) > \dots$$

The closedness of these inclusions in the case that G is nondiscrete follows from [25, Corollary 2.1.20] in Zimmer's book, since standard Borel spaces are countably separated.

Construction 2.17 Let X be a Lebesgue G -complex. We define X_{-1} to be a singleton, so that $L^\infty(X_{-1}) \cong \mathbb{R}$. There is a unique “face map” $\delta_0: X_0 \rightarrow X_{-1}$ which induces the inclusion of constants $\delta^0: \mathbb{R} \rightarrow L^\infty(X_0)$.

Applying the L^∞ -functor from Remark 2.6 to the face maps $\delta_i: X_{l+1} \rightarrow X_l$ provides a family of morphisms $\delta^i: L^\infty(X_l) \rightarrow L^\infty(X_{l+1})$ of coefficient G -modules. If we define

$$d^l: L^\infty(X_l) \rightarrow L^\infty(X_{l+1}), \quad d^l := \sum_{i=0}^{l+1} (-1)^i \delta^i \quad \text{for } l \geq -1,$$

then $d^l \circ d^{l-1} = 0$ for all $l \geq 0$, and hence we obtain a cochain complex

$$(2-3) \quad 0 \rightarrow \mathbb{R} \xrightarrow{d^{-1}} L^\infty(X_0) \xrightarrow{d^0} L^\infty(X_1) \xrightarrow{d^1} L^\infty(X_2) \xrightarrow{d^2} L^\infty(X_3) \rightarrow \dots$$

of coefficient G -modules.

Definition 2.18 Given a Lebesgue G -complex X , the cochain complex of coefficient G -modules (2-3) is called the *augmented L^∞ -complex* associated to X .

We will say that X is *boundedly acyclic* if its augmented L^∞ -complex is exact, and *boundedly γ_0 -acyclic* for some $\gamma_0 \in \mathbb{N} \cup \{\infty\}$ if $\ker d^l = \text{im } d^{l-1}$ for every $l \in [\gamma_0]$.

If a Lebesgue G -complex X is boundedly γ_0 -acyclic for some $\gamma_0 \in \mathbb{N}$ that we do not want to specify, then we will say that X is *partially boundedly acyclic*. We will also use the convention that every Lebesgue G -complex is boundedly $(-\infty)$ -acyclic.

2.4 The spectral sequence of a partially boundedly acyclic Lebesgue G -complex

With every partially boundedly acyclic Lebesgue G -complex, we can now associate a spectral sequence; our notation concerning spectral sequences follows McCleary's book [15].

Proposition 2.19 *Let $X = (X_q)_{q \in \mathbb{N}}$ be a Lebesgue G -complex which is boundedly γ_0 -acyclic for some $\gamma_0 \in \mathbb{N} \cup \{\infty\}$. Then there exists a first-quadrant spectral sequence $E_{\bullet, \bullet}^*$ with first page terms and differentials*

$$(2-4) \quad E_1^{p,q} = H_b^q(G; L^\infty(X_{p-1})) \quad \text{and} \quad d_1^{p,q} = H_b^q(G; d^{p-1}) \quad \text{for all } p, q \geq 0$$

that converges to $E_\infty^t = 0$ for all $t \in [\gamma_0 + 1]$, where d^q are the coboundary operators in (2-3).

Proof Consider the first-quadrant double complex $(L^{\bullet, \bullet}, d_H, d_V)$, whose terms and differentials for $p, q \geq 0$ are given by

$$\begin{aligned} L^{p,q} &:= L^\infty(G^{p+1} \times X_{q-1})^G \cong L^\infty(G^{p+1}; L^\infty(X_{q-1}))^G, \\ d_H^{p,q}: L^{p,q} &\rightarrow L^{p+1,q}, \quad d_H^{p,q} f(g_0, \dots, g_p) := \sum_{i=0}^p (-1)^i f(g_0, \dots, \hat{g}_i, \dots, g_p), \\ d_V^{p,q}: L^{p,q} &\rightarrow L^{p,q+1}, \quad d_V^{p,q} f(g_0, \dots, g_{p-1}) := d^{q-1}(f(g_0, \dots, g_{p-1})). \end{aligned}$$

Here d^{q-1} denotes the coboundary operator of the augmented L^∞ -complex of X , as in (2-3). We let ${}^I E_{\bullet, \bullet}^*$ and ${}^{II} E_{\bullet, \bullet}^*$ be the two spectral sequences associated with the horizontal and vertical filtrations of $L^{\bullet, \bullet}$, respectively, both of which converge to the cohomology of the total complex of $L^{\bullet, \bullet}$ and whose first-page terms and differentials are given by

$$\begin{aligned} {}^I E_1^{p,q} &= H^q(L^{p, \bullet}, d_V^{p, \bullet}), \quad {}^I d_1^{p,q} = H^q(d_H^{p, \bullet}): {}^I E_1^{p,q} \rightarrow {}^I E_1^{p+1,q}, \\ {}^{II} E_1^{p,q} &= H^q(L^{\bullet, p}, d_H^{\bullet, p}), \quad {}^{II} d_1^{p,q} = H^q(d_V^{\bullet, p}): {}^{II} E_1^{p,q} \rightarrow {}^{II} E_1^{p+1,q}. \end{aligned}$$

For a proof of the existence and convergence properties of these two spectral sequences, see McCleary [15, Theorem 2.15]. By the γ_0 -acyclicity of X and exactness of the functor $L^\infty(G^{p+1}, -)^G$ for every $p \geq 0$ (see Lemma 2.13), the complex

$$0 \rightarrow L^{p,0} \rightarrow L^{p,1} \rightarrow \dots \rightarrow L^{p,\gamma_0} \rightarrow L^{p,\gamma_0+1} \rightarrow L^{p,\gamma_0+2} \rightarrow \dots$$

is exact up to degree $\gamma_0 + 1$. Thus ${}^I E_1^{p,q} = 0$ for all $p \geq 0$ and $q \in [\gamma_0 + 1]$, and therefore ${}^I E_\infty^t \cong \bigoplus_{p+q=t} {}^I E_\infty^{p,q} = 0$ for every $t \in [\gamma_0 + 1]$. We deduce that also the spectral sequence $E_{\bullet, \bullet}^* := {}^{II} E_{\bullet, \bullet}^*$ converges to 0 in degrees $t \in [\gamma_0 + 1]$. Finally, it follows from Proposition 2.10 that the first page of $E_{\bullet, \bullet}^*$ is given by (2-4). $\square$

3 The highly essentially transitive case

In this section, G continues to denote a locally compact group, either second-countable or discrete, and X denotes a partially boundedly acyclic Lebesgue G -complex. In the discrete case, G -sets are always endowed with the counting measure. We are going to discuss how certain transitivity properties of the G -action on X affect the spectral sequence constructed in the previous section. This will lead us to a proof of [Theorem A](#) from [Section 1](#).

Definition 3.1 Let X be a Lebesgue G -complex and $\tau_0 \in \mathbb{N} \cup \{\infty\}$. We say that X is *essentially τ_0 -transitive* if X_l admits a conull G -orbit for all $l \in [\tau_0]$. We use the convention that any Lebesgue G -complex is essentially $(-\infty)$ -transitive.

For the remainder of this section we assume that X is a Lebesgue G -complex which is boundedly γ_0 -acyclic and essentially τ_0 -transitive for some $\gamma_0, \tau_0 \in \mathbb{N} \cup \{\infty\}$. We denote by $E_{\bullet}^{\bullet, \bullet}$ the associated spectral sequence constructed in [Proposition 2.19](#).

Remark 3.2 The bottom row of the first page of the spectral sequence $E_{\bullet}^{\bullet, \bullet}$ is given by

$$E_1^{p,0} = H_b^0(G; L^\infty(X_{p-1})) = L^\infty(X_{p-1})^G \quad \text{and} \quad d_1^{p,0} = H_b^0(G; d^{p-1}) \quad \text{for all } p \geq 0.$$

Since the constant functions are always contained in $L^\infty(X_{p-1})^G$ we have an embedding of constants $c_p: \mathbb{R} \hookrightarrow E_1^{p,0}$ for every $p \geq 0$, such that all of the diagrams

$$\begin{array}{ccc} E_1^{p,0} & \xrightarrow{\delta^i} & E_1^{p+1,0} \\ \swarrow c_p & & \searrow c_{p+1} \\ & \mathbb{R} & \end{array}$$

commute. Since G acts essentially τ_0 -transitive, we have

$$E_1^{p,0} = L^\infty(X_{p-1})^G = c_p(\mathbb{R}) \cong \mathbb{R} \quad \text{for all } p \in [\tau_0 + 1].$$

It thus follows from the commuting diagram above that each of the dual face maps $\delta^i: E_1^{p,0} \rightarrow E_1^{p+1,0}$ equals the identity if $p \in [\tau_0]$ and is an injection if $p = \tau_0 + 1$. We deduce that the face map $d_1^{p,0}$ is the zero map if $p \in [\tau_0 + 1]$ is odd, an isomorphism if $p \in [\tau_0]$ is even, and injective if $p = \tau_0 + 1$ is even. Consequently

$$(3-1) \quad E_2^{p,0} = 0 \quad \text{for all } p \in [\tau_0 + 1].$$

Construction 3.3 By assumption, we can find $o_{\tau_0} \in X_{\tau_0}$ such that the orbit $G.o_{\tau_0} \subset X_{\tau_0}$ is conull. Since the face maps are G -equivariant and measure-class-preserving, any codimension-1 face $o_{\tau_0-1} \in X_{\tau_0-1}$ has a conull orbit in X_{τ_0-1} . Thus we construct a flag $o_{-1} < o_0 < \dots < o_{\tau_0}$ inductively with $o_q \in X_q$ such that o_q has a conull orbit in X_q for all $q \in \{-1\} \cup [\tau_0]$ —recall our convention that X_{-1} is a singleton,

and thus o_{-1} denotes its only element. We refer to such a flag as a *generic flag* in X . Given a generic flag, we define

$$H_q := \text{Stab}_G(o_q) \quad \text{for } q \in \{-1\} \cup [\tau_0].$$

Note that according to our convention, G acts trivially on X_{-1} , so $H_{-1} = G$.

By [Remark 2.16](#), we have a chain of closed subgroups

$$G = H_{-1} > H_0 > H_1 > \cdots > H_{\tau_0}.$$

Now let $q \in \{-1\} \cup [\tau_0 - 1]$ and $i \in [q + 1]$, so that $\delta_i(o_{q+1}) \in X_q$. By transitivity, there exists $w_{q,i} \in G$ such that

$$\delta_i(o_{q+1}) = w_{q,i}^{-1} \cdot o_q,$$

and we fix a choice of such elements of G once and for all.

Lemma 3.4 *For all $q \in \{-1\} \cup [\tau_0 - 1]$ and $i \in [q + 1]$, we have the inclusion $w_{q,i} H_{q+1} w_{q,i}^{-1} < H_q$. In particular, there is a well-defined map*

$$\text{Int}(w_{q,i}): H_{q+1} \rightarrow H_q, \quad h \mapsto w_{q,i} h w_{q,i}^{-1}.$$

Proof If $h \in H_{q+1}$, then

$$w_{q,i}^{-1} \cdot o_q = \delta_i(o_{q+1}) = \delta_i(h \cdot o_{q+1}) = h \cdot \delta_i(o_{q+1}) = h \cdot w_{q,i}^{-1} \cdot o_q \implies w_{q,i} h w_{q,i}^{-1} \in H_q. \quad \square$$

Construction 3.5 For $p \in [\tau_0 + 1]$, we define the $(p-1)^{\text{th}}$ induction module $\text{Ind}^{p-1} := L^\infty(G)^{H_{p-1}}$, where the H_{p-1} -invariance is taken with respect to the restriction of the G -action by left-translation. Equipped with the right-translation G -action ρ , the space Ind^{p-1} is a Banach G -module. We now define a family of induction maps

$$\text{Ind}: L^\infty(G^{q+1})^{H_{p-1}} \rightarrow L^\infty(G^{q+1}, \text{Ind}^{p-1})^G, \quad (\text{Ind } \varphi)(\vec{g})(x) := \varphi(x \vec{g}) \quad \text{for } q \geq 0,$$

which produces a morphism of cochain complexes; see [\[16, Proposition 10.1.3\]](#). We also define

$$\Psi_0: \text{Ind}^{p-1} \rightarrow L^\infty(X_{p-1}), \quad (\Psi_0 \varphi)(g \cdot o_{p-1}) := \varphi(g^{-1}),$$

and define a morphism of cochain complexes $\Psi := L^\infty(G^{q+1}; \Psi_0)^G$, so that

$$\Psi: L^\infty(G^{q+1}; \text{Ind}^{p-1})^G \rightarrow L^\infty(G^{q+1}; L^\infty(X_{p-1}))^G \quad \text{for } q \geq 0.$$

Note that by [Proposition 2.11](#), the cohomology of $L^\infty(G^{q+1})^{H_{p-1}}$ is $H_b^q(H_{p-1})$, and by [Proposition 2.10](#), the cohomology of $L^\infty(G^{q+1}; L^\infty(X_{p-1}))^G$ is $H_b^q(G; L^\infty(X_{p-1})) = E_1^{p,q}$.

Proposition 3.6 *The composition $\Psi \circ \text{Ind}: L^\infty(G^{q+1})^{H_{p-1}} \rightarrow L^\infty(G^{q+1}; L^\infty(X_{p-1}))^G$ induces an isomorphism*

$$I_{p,q}: H_b^q(H_{p-1}) \rightarrow E_1^{p,q} \quad \text{for every } p \in [\tau_0 + 1] \text{ and every } q \geq 0.$$

Proof It suffices to show that Ind and Ψ induce isomorphisms in cohomology. The former is just the Eckmann–Shapiro lemma for L^∞ -modules (see Monod [16, Proposition 10.1.3] and Monod and Nariman [19, Lemma 3.5]) and the latter is immediate from the fact Ψ_0 is an isomorphism for any fixed $q \geq 0$ by essential transitivity. $\square$

Having computed the terms on the first page of $E_{\bullet,\bullet}^*$, we now compute the differentials. The maps $\text{Int}(w_{p-1,i}): H_p \rightarrow H_{p-1}$ from Lemma 3.4 induce a family of maps

$$H_b^q(\text{Int}(w_{p-1,i})): H_b^q(H_{p-1}) \rightarrow H_b^q(H_p),$$

and we claim:

Lemma 3.7 For every $p \in [\tau_0]$, $q \geq 0$, and $i \in [p]$, we have a commuting diagram

$$\begin{array}{ccc} E_1^{p,q} & \xrightarrow{H_b^q(G;\delta^i)} & E_1^{p+1,q} \\ I_{p,q} \uparrow & & \uparrow I_{p+1,q} \\ H_b^q(H_{p-1}) & \xrightarrow{H_b^q(\text{Int}(w_{p-1,i}))} & H_b^q(H_p) \end{array}$$

Proof For all i , p and q as in the lemma we consider the commutative diagram

$$\begin{array}{ccccccc} \mathbb{R} & \longrightarrow & L^\infty(G) & \longrightarrow & L^\infty(G^2) & \longrightarrow & \dots \longrightarrow L^\infty(G^{q+1}) \longrightarrow \dots \\ \text{Id} \downarrow & & \Psi_{i,p,0} \downarrow & & \Psi_{i,p,1} \downarrow & & \Psi_{i,p,q} \downarrow \\ \mathbb{R} & \longrightarrow & L^\infty(G) & \longrightarrow & L^\infty(G^2) & \longrightarrow & \dots \longrightarrow L^\infty(G^{q+1}) \longrightarrow \dots \end{array}$$

where $\Psi_{i,p,q}(\varphi)(g_0, \dots, g_q) := \varphi(w_{p-1,i}g_0, \dots, w_{p-1,i}g_q)$. For all $h \in H_p$ and $\vec{g} \in G^{q+1}$

$$\begin{aligned} \Psi_{i,p,q}(\text{Int}(w_{p-1,i})(h) \cdot \varphi)(\vec{g}) &= \text{Int}(w_{p-1,i})(h) \cdot \varphi(w_{p-1,i} \cdot \vec{g}) = \varphi(w_{p-1,i} h^{-1} \cdot \vec{g}) \\ &= \Psi_{i,p,q}(\varphi)(h^{-1} \vec{g}) = h \cdot \Psi_{i,p,q}(\varphi)(\vec{g}), \end{aligned}$$

and hence $\Psi_{i,p,q}$ restricts to a map $\Psi_{i,p,q}: L^\infty(G^{q+1})^{H_{p-1}} \rightarrow L^\infty(G^{q+1})^{H_p}$, which by [16, Proposition 8.4.2] induces the map $H_b^q(\text{Int}(w_{p-1,i}))$ in cohomology. It thus remains to show only that the diagram

$$\begin{array}{ccc} L^\infty(G^{q+1}; L^\infty(X_{p-1}))^G & \xrightarrow{\delta^i} & L^\infty(G^{q+1}; L^\infty(X_p))^G \\ \Psi \circ \text{Ind} \uparrow & & \uparrow \Psi \circ \text{Ind} \\ L^\infty(G^{q+1})^{H_{p-1}} & \xrightarrow{\Psi_{i,p,q}} & L^\infty(G^{q+1})^{H_p} \end{array}$$

commutes. Now if $\varphi \in L^\infty(G^{q+1})^{H_{p-1}}$, $\vec{g} \in G^{q+1}$, and $x \in G$, then

$$\begin{aligned} (\delta^i \circ \Psi \circ \text{Ind})(\varphi)(\vec{g})(x \cdot o_p) &= (\Psi \circ \text{Ind})(\varphi)(\vec{g})(x w_{p-1,i}^{-1} \cdot o_{p-1}) = \varphi(w_{p-1,i} x^{-1} \vec{g}) \\ &= \Psi_{i,p,q}(x^{-1} \vec{g}) = (\Psi \circ \text{Ind} \circ \Psi_{i,p,q})(\varphi)(\vec{g})(x \cdot o_p), \end{aligned}$$

and hence $\delta^i \circ \Psi \circ \text{Ind} = \Psi \circ \text{Ind} \circ \Psi_{i,p,q}$, which proves the lemma. $\square$

Combining this fact with Propositions 2.19 and 3.6, Lemma 3.4, and (3-1), we arrive at Theorem A.

4 Measurable Quillen families

In [Definition 1.2](#), we have introduced the notion of an infinite measurable Quillen family. Since the quantitative version of [Theorem B](#) (see [Theorem 4.6](#) below) also applies to certain finite families of groups, we modify the definitions accordingly.

Setting 4.1 Let $R \in \mathbb{N} \cup \{\infty\}$ be a *length parameter* and let

$$\gamma, \tau: [R] \rightarrow \mathbb{N} \cup \{\pm\infty\}$$

be two functions, called the *acyclicity range* and *transitivity range*, respectively. We assume that we are given

- a locally compact group G_r for every $r \in [R]$, all either second-countable or discrete,
- an embedding $\iota_r: G_r \hookrightarrow G_{r+1}$ for every $r < R$,
- a Lebesgue G_r -complex $X(r)$ for every $r \in [R]$.

We set $G_r := \{1\}$ for all $r < 0$. Furthermore, we keep our convention from last section that $X(r)_{-1}$ is a singleton, and we denote by $o_{r,-1}$ its only element.

Generalizing [Definition 1.2](#), we declare:

Definition 4.2 We say that $(G_r, X(r))_{r \in [R]}$ is a *measurable (R, γ, τ) -Quillen family* provided:

- (MQ1) $X(r)$ is a boundedly $\gamma(r)$ -acyclic Lebesgue G_r -complex for every $r \in [R]$.
- (MQ2) $X(r)$ is an essentially $\tau(r)$ -transitive Lebesgue G_r -complex for every $r \in [R]$.
- (MQ3) For every $r \in [R]$, there exists a generic flag $o_{r,-1} < o_{r,0} < o_{r,1} < \dots < o_{r,\tau(r)}$ in $X(r)$ such that the corresponding stabilizers $H_{r,q} := \text{Stab}_{G_r}(o_{r,q}) < G_r$ are compatible with $G_{r-\tau(r)}, \dots, G_r$ in the sense of one of (MQ3a), (MQ3b), or (MQ3c) from [Section 1](#) for elements $w_{r,q,i} \in G_r$ with $r \in [R]$, $q \in [\tau(r) - 1]$, and $i \in [q + 1]$.

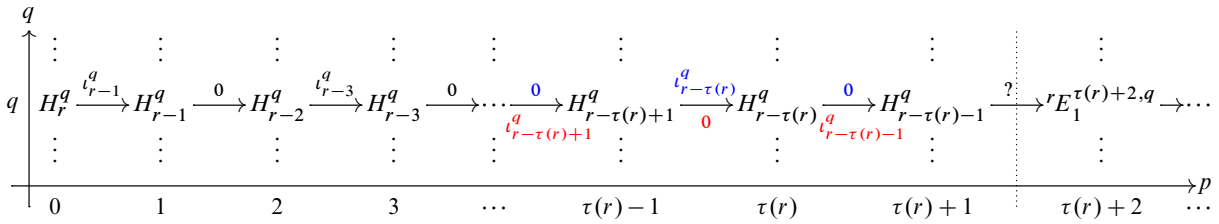
In particular, a measurable (∞, γ, τ) -Quillen family is the same as a measurable (γ, τ) -Quillen family in the sense of [Definition 1.2](#).

Applying [Theorem A](#) to each of the complexes $X(r)$ in a measurable Quillen family, we obtain:

Corollary 4.3 Assume that $(G_r, X(r))_{r \in [R]}$ is a measurable (R, γ, τ) -Quillen family. Then for every $r \in [R]$, there exists a spectral sequence ${}^r E_{\bullet, \bullet}^{\bullet}$ such that:

- (i) ${}^r E_{\bullet, \bullet}^{\bullet}$ converges and ${}^r E_{\infty}^t = 0$ for every $t \in [\gamma(r) + 1]$.
- (ii) ${}^r E_1^{p,q} = H_b^q(G_{r-p})$ for all $p \in [\tau(r) + 1]$ and $q \geq 0$.
- (iii) ${}^r E_2^{p,0} = 0$ for all $p \in [\tau(r) + 1]$ and $q \geq 0$.
- (iv) For all $p \in [\tau(r)]$ and $q \geq 0$, the map $d_1^{p,q}: H_b^q(G_{r-p}) \rightarrow H_b^q(G_{r-p-1})$ is given by

$$d_1^{p,q} = \begin{cases} H_b^q(\iota_{r-p-1}) & \text{if } p \text{ is even,} \\ 0 & \text{if } p \text{ is odd.} \end{cases}$$


Figure 1: The first page ${}^rE_1^{\bullet,\bullet}$.

Proof Since the inflation maps in bounded cohomology whose inducing epimorphisms have amenable kernels are isomorphisms (see Monod [16, Corollary 8.5.2]), condition (MQ3) produces the following commutative diagram for all $r \in [R]$, $p \in [\tau(r)]$, $i \in [p]$, and $q \geq 0$:

$$\begin{array}{ccc} H_b^q(H_{r,p-1}) & \xrightarrow{H_b^q(\text{Int}(w_{r,p-1,i}))} & H_b^q(H_{r,p}) \\ \uparrow \cong & & \uparrow \cong \\ H_b^q(\pi_{r,p}) & \xrightarrow{H_b^q(\iota_{r,p-1})} & H_b^q(\pi_{r,p-1}) \end{array}$$

The corollary follows now from Theorem A. $\square$

If for any $s \in [R-1]$ and any $q \geq 0$ we abbreviate

$$H_s^q := H_b^q(G_s) \quad \text{and} \quad \iota_s^q := H_b^q(\iota_s),$$

then the first page ${}^rE_1^{\bullet,\bullet}$ is given by Figure 1. The terms to the left of the dotted line are explained by Corollary 4.3(ii), while the behavior of the terms to the right of the dotted line is a priori not understood. According to Corollary 4.3(iii), the parity of $\tau(r)$ determines the pattern of the arrows just to the left of the dotted line: the labels above the arrows correspond to the case in which $\tau(r)$ is odd, and the labels underneath correspond to the even case.

Remark 4.4 Let $(G_r, X(r))_{r \in [R]}$ be a measurable (R, γ, τ) -Quillen family. We say that $q_0 \in \mathbb{N}$ is an *initial parameter* for this measurable Quillen family if

$$H_b^q(\iota_r): H_b^q(G_{r+1}) \rightarrow H_b^q(G_r) \text{ is an isomorphism for all } q \leq q_0 \text{ and } r < R.$$

In theory, one would always like to work with the optimal initial parameter

$$q_0 := \sup\{q \mid H_b^q(\iota_r): H_b^q(G_{r+1}) \rightarrow H_b^q(G_r) \text{ is an isomorphism for all } r < R\},$$

but in practice, the optimal initial parameter is not known. If no a priori information about the bounded cohomology of the group G_r is available, then one can always work with $q_0 := 1$, since $H_b^1(G)$ vanishes for lsc groups G . In some cases of interest, better initial parameters are available. Notably, if all G_r are connected semisimple Lie groups with finite center and either all of the groups G_r are of Hermitian type or all are of non-Hermitian type, then $q_0 \geq 2$.

To state the quantitative version of our bc-stability criterion, we introduce the following:

Definition 4.5 Let $(G_r, X(r))_{r \in [R]}$ be a measurable (R, γ, τ) -Quillen family with initial parameter $q_0 \in \mathbb{N}$. We define the associated *dual acyclicity range* and *dual transitivity function* to be functions $\tilde{\gamma}, \tilde{\tau}: \mathbb{N} \times [R-1] \rightarrow \mathbb{N} \cup \{\pm\infty\}$ defined by the formulae

$$\tilde{\gamma}(q, r) := \min_{j=q_0}^q \{\gamma(r+1-2(q-j)) - j\} \quad \text{and} \quad \tilde{\tau}(q, r) := \min_{j=q_0}^q \{\tau(r+1-2(q-j)) - j\}$$

for $r+1-2(q-q_0) \geq 0$, and $\tilde{\gamma}(q, r) = \tilde{\tau}(q, r) = -\infty$ otherwise.

We state now our main result:

Theorem 4.6 Let $(G_r, X(r))_{r \in [R]}$ be a measurable (R, γ, τ) -Quillen family with initial parameter $q_0 \geq 1$, associated dual acyclicity range $\tilde{\gamma}$, and dual transitivity function $\tilde{\tau}$. Then the inclusion ι_r induces an isomorphism (resp. an injection)

$$H_b^q(\iota_r): H_b^q(G_{r+1}) \xrightarrow{\cong} H_b^q(G_r) \quad (\text{resp. } H_b^{q+1}(\iota_r): H_b^{q+1}(G_{r+1}) \hookrightarrow H_b^{q+1}(G_r)),$$

provided $r \in [R-1]$ and $q \geq 0$ satisfy the condition

$$(4-1) \quad \min\{\tilde{\gamma}(q, r), \tilde{\tau}(q, r) - 1\} \geq 0.$$

An immediate corollary of this quantitative bc-stability result is [Theorem B](#), as stated in [Section 1](#).

Proof of Theorem B from Theorem 4.6 We may assume without loss of generality that $q_0 = 1$. Since $\gamma(r) \rightarrow \infty$ and $\tau(r) \rightarrow \infty$, we find for every $q \geq q_0$ some $r(q) \in \mathbb{N}$ such that for all $j \in \{q_0, \dots, q\}$ and all $r \geq r(q)$, we have

$$\gamma(r+1-2(q-j)) \geq j \quad \text{and} \quad \tau(r+1-2(q-j)) \geq j+1.$$

Then [Theorem 4.6](#) implies the chain of isomorphisms

$$H_b^q(G_{r(q)}) \xleftarrow{\cong} H_b^q(G_{r(q)+1}) \xleftarrow{\cong} H_b^q(G_{r(q)+2}) \xleftarrow{\cong} \dots,$$

which is the desired bc-stability. □

The functions $\tilde{\gamma}(q, r)$ and $\tilde{\tau}(q, r)$ have been defined so that they satisfy the following lemma, which will allow us to make inductive arguments:

Lemma 4.7 Let $r, q \in \mathbb{N}$ for $q \geq q_0$, and assume that $\min\{\tilde{\gamma}(q, r), \tilde{\tau}(q, r) - 1\} \geq 0$. Then the following hold:

- (i) $\gamma(r+1) \geq q$ and $\tau(r+1) \geq q+1$.
- (ii) $\min\{\tilde{\gamma}(q-1, r), \tilde{\tau}(q-1, r) - 1\} \geq 0$.
- (iii) If $p \in [q-q_0]$ is odd, then $\min\{\tilde{\gamma}(q-p, r-p-1), \tilde{\tau}(q-p, r-p-1) - 1\} \geq 0$.
- (iv) If $p \in [q-q_0]$ is even, then $\min\{\tilde{\gamma}(q-p-1, r-p-2), \tilde{\tau}(q-p-1, r-p-2) - 1\} \geq 0$.

Proof The assumption means that for all $j \in \{q_0, \dots, q\}$ we have

$$(4-2) \quad r + 1 - 2(q - j) \geq 0,$$

$$(4-3) \quad \gamma(r + 1 - 2(q - j)) \geq j, \quad \tau(r + 1 - 2(q - j)) \geq j + 1.$$

(i) Choose $j = q$ in (4-3).

(ii) If $q = q_0$, then $r + 1 - 2(q - 1 - q_0) \geq 0$ and $\tilde{\gamma}(q - 1, r) = \tilde{\tau}(q - 1, r) = \infty$ for any r . Assume now that $q > q_0$, and let $j \in \{q_0, \dots, q - 1\}$. Observe that

$$r + 1 - 2(q - 1 - j) = r + 1 - 2(q - (j + 1)).$$

Since $j + 1 \in \{q_0, \dots, q\}$, from this equality combined with (4-2) and (4-3) we deduce

$$r + 1 - 2(q - 1 - j) \geq 0, \quad \gamma(r + 1 - 2(q - 1 - j)) \geq j + 1 > j, \quad \tau(r + 1 - 2(q - 1 - j)) \geq j + 2 > j + 1,$$

which establishes the claim.

(iii) Observe that for any odd $p \in [q - q_0]$ and any $j \in \{q_0, \dots, q - p\}$, we have

$$r - p - 2(q - p - j) = r + 1 - 2(q - (j + \frac{1}{2}(p - 1))),$$

and $j + \frac{1}{2}(p - 1) \in \{q_0 + \frac{1}{2}(p - 1), \dots, q - \frac{1}{2}(p + 1)\} \subset \{q_0, \dots, q\}$. Just as in (ii), the claim follows now from (4-2) and (4-3).

(iv) For any even $p \in [q - q_0]$ and any $j \in \{q, \dots, q - p\}$, we have

$$r - p - 1 - 2(q - p - 1 - j) = r + 1 - 2(q - (j + \frac{1}{2}p)),$$

and $j + \frac{1}{2}p \in \{q_0 + \frac{1}{2}p, \dots, q - \frac{1}{2}p\} \subset \{q_0, \dots, q\}$, and conclude by (4-2) and (4-3). $\square$

5 Proof of the main theorem

The remainder of this article is devoted to deriving Theorem 4.6 from Corollary 4.3. We consider a measurable (R, γ, τ) -Quillen family $(G_r, X(r))_{r \in [R]}$ with initial parameter $q_0 \geq 1$, dual acyclicity range $\tilde{\gamma}$, and dual transitivity function $\tilde{\tau}$, and abbreviate

$$H_s^q := H_b^q(G_s) \quad \text{and} \quad \iota_s^q := H_b^q(\iota_s),$$

as before. We consider the following statement for $q \in \mathbb{N}$ and $r < R$:

$(S_{q,r})$ $\iota_r^q: H_{r+1}^q \rightarrow H_r^q$ is an isomorphism and $\iota_r^{q+1}: H_{r+1}^{q+1} \rightarrow H_r^{q+1}$ is an injection.

The conclusion of Theorem 4.6 then says that for all $q \in \mathbb{N}$, the following statement (S_q) holds true:

(S_q) If $r < R$ satisfies $\min\{\tilde{\gamma}(q, r), \tilde{\tau}(q, r) - 1\} \geq 0$, then $(S_{q,r})$ holds.

If $q < q_0$ then $r + 1 - 2(q - q_0) \geq 0$, and hence by definition $\tilde{\gamma}(q, r) = \tilde{\tau}(q, r) = \infty$. Since q_0 is an initial parameter, both ι_r^q and ι_r^{q+1} are isomorphisms for all $r < R$. Thus (S_q) holds for all $q < q_0$. We now proceed to prove (S_q) for $q \geq q_0$ by induction on q . For this purpose, assume that $q \geq q_0$ and that $(S_{q'})$ holds for all $q' < q$. In view of Lemma 4.7, we have:

Lemma 5.1 For $r < R$ with $\min\{\tilde{\gamma}(q, r), \tilde{\tau}(q, r) - 1\} \geq 0$, the following hold:

- (A) The map ι_r^q is injective.
- (B) The map ι_{r-p-1}^{q-p} is an isomorphism for all odd numbers $p \in [q]$.
- (C) The map ι_{r-p-2}^{q-p} is an injection for all even numbers $p \in [q]$.

Proof (A) By Lemma 4.7(ii), the inequality $\min\{\tilde{\gamma}(q-1, r), \tilde{\tau}(q-1, r) - 1\} \geq 0$ holds. Thus we may apply (S_{q-1}) to deduce the claim.

(B) If $p > q - q_0$, then $q - p < q_0$ and hence ι_{r-p-1}^{q-p} is an isomorphism by the initial condition. If $p \leq q - q_0$, then by Lemma 4.7(iii) we have $\min\{\tilde{\gamma}(q-p, r-p-1), \tilde{\tau}(q-p, r-p-1) - 1\} \geq 0$. We may thus use (S_{q-p}) to show that $(S_{q-p, r-p-1})$ is true, implying the claim.

(C) The case $p > q - q_0$ is again covered by the initial condition, and if $p \leq q - q_0$, then by Lemma 4.7(iv) we have $\min\{\tilde{\gamma}(q-p-1, r-p-2), \tilde{\tau}(q-p-1, r-p-2) - 1\} \geq 0$. Thus, by (S_{q-p-1}) , the statement $(S_{q-p-1, r-p-2})$ applies. $\square$

To establish (S_q) , we fix a natural number $r < R$ such that $\min\{\tilde{\gamma}(q, r), \tilde{\tau}(q, r) - 1\} \geq 0$ and consider the spectral sequence $E_{\bullet, \bullet}^{\bullet, \bullet} := {}^{r+1}E_{\bullet, \bullet}^{\bullet, \bullet}$ from Corollary 4.3. By Lemma 4.7(i), we have

$$(5-1) \quad \gamma(r+1) \geq q \quad \text{and} \quad \tau(r+1) \geq q+1,$$

and thus Corollary 4.3 gives

$$(5-2) \quad \begin{aligned} E_1^{p', q'} &= H_{r+1-p'}^{q'} && \text{for } p' \in [q+2] \text{ and } q' \geq 0, \\ E_2^{p', 0} &= 0 && \text{for } p' \in [q+2], \\ d_1^{p', q'} &= \begin{cases} \iota_{r-p'}^{q'} & \text{if } p' \text{ is even,} \\ 0 & \text{if } p' \text{ is odd,} \end{cases} && \text{for } p' \in [q+1] \text{ and } q' \geq 0, \\ E_\infty^t &= 0 && \text{for } t \in [q+1]. \end{aligned}$$

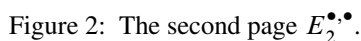
The second equality is (3-1). We are going to show now that

$$(5-3) \quad E_2^{0, q+1} \cong \ker(\iota_r^{q+1}) \quad \text{and} \quad E_2^{1, q} \cong \operatorname{coker}(\iota_r^q),$$

$$(5-4) \quad E_2^{p', q'} = 0 \quad \text{for all } (p', q') \in \{0, \dots, q\}^2 \text{ with } p' + q' = q+2.$$

These computations will be sufficient to establish (S_q) : As indicated in Figure 2, the differentials emanating from $E_2^{0, q+1}$ and $E_2^{1, q}$ as well as all those emanating from the positions $(0, q+1)$ and $(1, q)$ on any of the following pages end up on the diagonal $p' + q' = q+2$. Since this diagonal contains only zeros up to row q , we conclude that

$$E_\infty^{0, q+1} \cong E_2^{0, q+1} \cong \ker(\iota_r^{q+1}) \quad \text{and} \quad E_\infty^{1, q} \cong E_2^{1, q} \cong \operatorname{coker}(\iota_r^q).$$



To prove (5-3) and (5-4), we only need to consider the few arrows on the first page of the spectral sequence depicted in Figure 3. The statement (5-3) is immediate from the fact that the leftmost maps in the q^{th} and $(q+1)^{\text{th}}$ row of $E_1^{\bullet, \bullet}$ are given by

Figure 3: The first page ${}^{r+1}E_1^{\bullet,\bullet}$.

respectively. As for (5-4), we first consider $q' \in \{1, \dots, q\}$ and define $p' := q - q' \in [q - 1]$ so that $p' + q' = q$. By (5-2), the two maps $E_1^{p'+1, q'} \rightarrow E_1^{p'+2, q'} \rightarrow E_1^{p'+3, q'}$ are given by

$$\begin{cases} H_{r-p'}^{q-p'} \xrightarrow{0} H_{r-p'-1}^{q-p'} \xrightarrow{\iota_{r-p'-2}^{q-p'}} H_{r-p'-2}^{q-p'} & \text{if } p' \text{ is even,} \\ H_{r-p'}^{q-p'} \xrightarrow{\iota_{r-p'-1}^{q-p'}} H_{r-p'-1}^{q-p'} \xrightarrow{0} H_{r-p'-2}^{q-p'} & \text{if } p' \text{ is odd.} \end{cases}$$

From Lemma 5.1(B)–(C), we deduce that

$$E_2^{p'+2, q'} = 0 \quad \text{for all } (p', q') \in [q]^2 \text{ with } p' + q' = q,$$

where the case $(p', q') = (q, 0)$ is covered in (5-2). This concludes the proof of (5-4). □

References

- [1] **M Bestvina**, *Homological stability of $\text{Aut}(F_n)$ revisited*, from “Hyperbolic geometry and geometric group theory”, Adv. Stud. Pure Math. 73, Math. Soc. Japan, Tokyo (2017) 1–11 [MR](#) [Zbl](#)
- [2] **M Bucher, M Burger, A Iozzi**, *The bounded Borel class and 3-manifold groups*, Duke Math. J. 167 (2018) 3129–3169 [MR](#) [Zbl](#)
- [3] **T Bühler**, *On the algebraic foundations of bounded cohomology*, Mem. Amer. Math. Soc. 1006, Amer. Math. Soc., Providence, RI (2011) [MR](#) [Zbl](#)
- [4] **M Burger, N Monod**, *Bounded cohomology of lattices in higher rank Lie groups*, J. Eur. Math. Soc. 1 (1999) 199–235 [MR](#) [Zbl](#)
- [5] **M Burger, N Monod**, *On and around the bounded cohomology of SL_2* , from “Rigidity in dynamics and geometry” (M Burger, A Iozzi, editors), Springer (2002) 19–37 [MR](#) [Zbl](#)
- [6] **C De la Cruz Mengual**, *On bounded-cohomological stability for classical groups*, PhD thesis, ETH Zürich (2019) Available at <https://doi.org/10.3929/ethz-b-000373321>
- [7] **C De la Cruz Mengual**, *The degree-three bounded cohomology of complex Lie groups of classical type*, preprint (2023) [arXiv 2304.00607](#)
- [8] **C De la Cruz Mengual, T Hartnick**, *Stabilization of bounded cohomology for classical groups*, preprint (2022) [arXiv 2201.03879](#)
- [9] **J Essert**, *Homological stability for classical groups*, Israel J. Math. 198 (2013) 169–204 [MR](#) [Zbl](#)
- [10] **R Frigerio**, *Bounded cohomology of discrete groups*, Math. Surv. Monogr. 227, Amer. Math. Soc., Providence, RI (2017) [MR](#) [Zbl](#)
- [11] **M Gromov**, *Volume and bounded cohomology*, Inst. Hautes Études Sci. Publ. Math. 56 (1982) 5–99 [MR](#) [Zbl](#)
- [12] **J L Harer**, *Stability of the homology of the mapping class groups of orientable surfaces*, Ann. of Math. 121 (1985) 215–249 [MR](#) [Zbl](#)
- [13] **A Hatcher, K Vogtmann**, *Homology stability for outer automorphism groups of free groups*, Algebr. Geom. Topol. 4 (2004) 1253–1272 [MR](#) [Zbl](#)
- [14] **W van der Kallen**, *Homology stability for linear groups*, Invent. Math. 60 (1980) 269–295 [MR](#) [Zbl](#)

- [15] **J McCleary**, *A user's guide to spectral sequences*, 2nd edition, Cambridge Stud. Adv. Math. 58, Cambridge Univ. Press (2001) [MR](#) [Zbl](#)
- [16] **N Monod**, *Continuous bounded cohomology of locally compact groups*, Lecture Notes in Math. 1758, Springer (2001) [MR](#) [Zbl](#)
- [17] **N Monod**, *Stabilization for SL_n in bounded cohomology*, from “Discrete geometric analysis”, Contemp. Math. 347, Amer. Math. Soc., Providence, RI (2004) 191–202 [MR](#) [Zbl](#)
- [18] **N Monod**, *Vanishing up to the rank in bounded cohomology*, Math. Res. Lett. 14 (2007) 681–687 [MR](#) [Zbl](#)
- [19] **N Monod**, **S Nariman**, *Bounded and unbounded cohomology of homeomorphism and diffeomorphism groups*, Invent. Math. 232 (2023) 1439–1475 [MR](#) [Zbl](#)
- [20] **V L Popov**, *Generically multiple transitive algebraic group actions*, from “Algebraic groups and homogeneous spaces”, Tata Inst. Fund. Res. Stud. Math. 19, Tata Inst. Fund. Res., Mumbai (2007) 481–523 [MR](#) [Zbl](#)
- [21] **D Quillen**, *Notebook 1974-1*, personal notes (1974) Available at <https://www.claymath.org/library/Quillen/Working-papers/quillen%201974/1974-1.pdf>
- [22] **D Sprehn**, **N Wahl**, *Homological stability for classical groups*, Trans. Amer. Math. Soc. 373 (2020) 4807–4861 [MR](#) [Zbl](#)
- [23] **K Vogtmann**, *A Stiefel complex for the orthogonal group of a field*, Comment. Math. Helv. 57 (1982) 11–21 [MR](#) [Zbl](#)
- [24] **C A Weibel**, *An introduction to homological algebra*, Cambridge Stud. Adv. Math. 38, Cambridge Univ. Press (1994) [MR](#) [Zbl](#)
- [25] **R J Zimmer**, *Ergodic theory and semisimple groups*, Monogr. Math. 81, Birkhäuser, Basel (1984) [MR](#) [Zbl](#)

Faculty of Electrical and Computer Engineering, Technion
Haifa, Israel

Institute of Algebra and Geometry, KIT
Karlsruhe, Germany

c.delacruz@technion.ac.il, tobias.hartnick@kit.edu

Received: 24 July 2023 Revised: 23 January 2024

T -equivariant motives of flag varieties

CAN YAYLALI

We use the construction of the stable homotopy category by Khan and Ravi to calculate the integral T -equivariant K -theory spectrum of a flag variety over an affine scheme, where T is a split torus associated to the flag variety. More precisely, we show that the T -equivariant K -theory ring spectrum of a flag variety is decomposed into a direct sum of K -theory spectra of the classifying stack BT indexed by the associated Weyl group. We also explain how to relate these results to the motivic world and deduce classical results for T -equivariant intersection theory and K -theory of flag varieties.

For this purpose, we analyze the motive of schemes stratified by affine spaces with group action, that preserves these stratifications. We work with cohomology theories, that satisfy certain vanishing conditions, which are satisfied for example by motivic cohomology and K -theory.

14F42; 14A20, 14L30, 14M15, 19E08

1. Introduction	2343
2. Motivic setup	2349
3. T -equivariant motivic homotopy theory of flag varieties	2355
References	2365

1 Introduction

Motivation

Let G be a split reductive group over a field k with split maximal torus T contained in a Borel subgroup B of G . The geometry of the flag variety G/B plays an important role in representation theory and the Langlands program. One of the aspects is to analyze the T -equivariant cycles of Schubert cells. There are various results on the T -equivariant intersection theory of a flag variety (see Brion [6]) or even on the T -equivariant K_0 of it (see Uma [43]). For example, $A_T^\bullet(G/B)$ has an $A_T^\bullet(k)$ -basis given by precisely the classes of the Schubert cells. Analogously the same is true for $K_0^T(G/B)$, ie the classes of the Schubert cells yield an $R(T)$ -basis. There is no canonical way to imply the former via the latter, as the equivariant Chern character map fails to be an isomorphism without completion along the augmentation ideal (see Krishna [32, Theorem 1.2]). But this result shows that after completion this was to be expected, also for higher K -theory.

One idea to bypass this problem is via passage to the motivic theory, which implies both results simultaneously. Motives were famously envisioned by Grothendieck. For any variety X one should be able to

encode the analogous behaviors of cohomology theories in an abelian category, the category of motives of X . In our context, the similar behavior of T -equivariant K -theory and intersection theory of a flag variety should be seen motivically. This is the starting point of this article.

Some motivic background

Defining a suitable abelian category of motives is a difficult task. One approach that has been studied over the years is to define the derived category of motives directly and attach a t -structure that recovers the (abelian) category of motives as the heart of this t -structure. There are many constructions of the derived category of motives by Voevodsky, Morel, Ayoub, Cisinski, Déglise, Spitzweck and more. Under certain assumptions, it is shown that the various definitions of the derived category of motives agree. Also, in the recent constructions, the category of motives comes equipped with a full 6-functor formalism which has become a powerful tool in analyzing functorial properties of cohomology theories.

We will in particular follow the construction of Morel and Voevodsky [37]. Roughly, for a scheme S they define the stable homotopy category $\mathrm{SH}(S)$ as the category of simplicial Nisnevich sheaves over Sm_S with coefficients in $\mathbb{Z}$, where for any smooth scheme X over S one inverts the structure map $\mathbb{A}_X^1 \rightarrow X$ (and the map induced on the associated representable sheaves) and inverts “tensoring” with $\mathbb{P}_S^1$ (on simplicial Nisnevich sheaves there is a closed monoidal structure given by the smash product; see [loc. cit.]). Ayoub has shown in his thesis (see [4, Scholie 1.4.2]) that the association $S \mapsto \mathrm{SH}(S)$ defines a functor that supports a full 6-functor formalism, ie $\mathrm{SH}(S)$ is closed monoidal and for finite type morphisms $f: S' \rightarrow S$ there exist adjunctions of functors $f_! \dashv f^!$, $f^* \dashv f_*$ between $\mathrm{SH}(S')$ and $\mathrm{SH}(S)$ with various compatibilities. If S is regular over a field and we work with étale sheaves with $\mathbb{Q}$ -coefficients (this is usually denoted by $\mathrm{SH}_{\mathbb{Q},\text{ét}}(S)$), this is equivalent to the construction of Cisinski and Déglise [11, 5.3.35, Theorem 16.1.4].

Voevodsky and Morel show that there exists an object $\mathrm{KGL} \in \mathrm{SH}(S)$, which we call the *motivic K -theory spectrum*, such that

$$\mathrm{Hom}_{\mathrm{SH}(S)}(1_S[n], \mathrm{KGL}) \cong K_n(S)$$

for any $n \in \mathbb{Z}$, where 1_S denotes the $\otimes$ -unit in $\mathrm{SH}(S)$. Further, by the work of Spitzweck — which relies on the moving lemma proved for Bloch’s cycle complex by Levine [34, Proposition 1.3] — we know that there exists an object $M\mathbb{Z}(n) \in \mathrm{SH}(S)$ such that

$$\mathrm{Hom}_{\mathrm{SH}(S)}(1_S, M\mathbb{Z}(n)[2n]) \cong A^n(S),$$

at least when S is a smooth scheme over a field (see Spitzweck [40, Corollary 7.19]). So, working with objects in the stable homotopy category enables us to deduce results in K -theory and intersection theory. In this way, one can also define a derived category of motives as modules over a chosen ring object in $\mathrm{SH}(S)$. For example Spitzweck defines a derived category of motives $\mathrm{DM}(S)$ as $\mathrm{SH}(S)$ -modules over $M\mathbb{Z}$ and Cisinski and Déglise define $\mathrm{DM}(S, \mathbb{Q})$ as $\mathrm{SH}(S, \mathbb{Q})$ -modules over $M\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Q}$ (here $\mathrm{SH}(S, \mathbb{Q})$ is the

stable homotopy category associated to sheaves with rational coefficients). Thus, by understanding $\mathrm{SH}(S)$ and the corresponding sheaves represented by smooth S -schemes, we can understand their cohomological behavior and their behavior in motivic categories.

To further generalize these constructions to the equivariant setting one needs to be careful. Usually, this is done by working with quotient stacks and imposing étale descent on SH . The idea then is to work with quotient stacks that can be smoothly covered by schemes and glue the corresponding motivic categories along the atlas. A drawback of the gluing process is that we lose information on the K -theoretic side. By the works of Carlsson and Joshua [8, Theorem 1.2] and Khan and Ravi [29, Corollary 7.1], one is able to see that the glued motivic K -theory spectrum does not represent genuine equivariant K -theory, but rather its completed version. This is due to the fact, that on algebraic stacks (in the sense of the Stacks project [41]) K -theory does not satisfy étale descent. For example, for a field k we have $K_{\mathbb{G}_m}(\mathrm{Spec}(k)) = \mathbb{Z}[t, t^{-1}]$ (the $\mathbb{G}_m$ -equivariant K -theory of $\mathrm{Spec}(k)$) but the simplicial limit along the smooth cover $\mathrm{Spec}(k) \rightarrow B \mathbb{G}_m$ yields

$$\lim_{[n] \in \Delta} K(\mathbb{G}_m^n) = \mathbb{Z}[[T]].$$

One can compute this limit as the completion of $\mathbb{Z}[t, t^{-1}]$ along $(1 - t)$, which is the ideal generated by the virtual rank 0 bundles in $K_{\mathbb{G}_m}(\mathrm{Spec}(k))$.

Motives of flag varieties

In this article we are interested in the T -equivariant motive of the flag variety G/B . Before we start to explain our results, let us recall some results in the nonequivariant setting. The scheme G/B classifies complete flags on an n -dimensional representation V of G , ie subspaces $V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n$, where each V_i has dimension i . We will recall in Section 3.1 a well-known fact, that G/B admits a finite stratification by affine spaces indexed by the Weyl group W of T in G . The strata are called Schubert cells, usually denoted by C_w for a $w \in W$, and the closure of C_w is called Schubert variety, usually denoted by X_w . The cohomology of these objects plays an important role in representation theory and enumerative geometry. In the former, the intersection cohomology of Schubert varieties can be related to so-called Kazhdan–Lusztig polynomials [27]. For the latter, one can see that the intersection ring of G/B has a basis consisting of Schubert cells (see Fulton [17, Section 14]). Fulton gives examples on how this result together with the multiplicative structure on Chow rings leads to solutions of enumerative problems — this is also known as Schubert calculus (see [17, Section 14.7; 39]). More generally, instead of the intersection ring, one can do Schubert calculus on the Grothendieck ring of G/B (see Brion [7] for an exposition).

In recent years the use of $\mathbb{A}^1$ -homotopy theory became more apparent in enumerative geometry (eg in the work of Levine and Pauli [35]). In that regard, one could ask if we can obtain a Schubert calculus on other $\mathbb{A}^1$ -cohomology theories. In fact, Hornbostel and Kiritchenko [20] show that on algebraic cobordism, we can find a similar Schubert calculus as in the Chow group case.

As hinted by Brion [7, Section 4], such results are also interesting in the T -equivariant setting. At least for the T -equivariant Grothendieck ring, this was studied by Griffeth and Ram [19]. In [loc. cit.] there is also a positivity conjecture (see [19, Conjecture 4.1]), which was proven later by Anderson, Griffeth and Miller [3, Corollary 5.1] generalizing the result of Graham [18, Corollary 4.1] on the Schubert calculus on the equivariant cohomology ring.

Back to our setting

As we are interested in equivariant K -theory as an $R(T)$ -module, where $R(T) = K_0^T(\mathrm{Spec}(k))$ is the representation ring of T , it suffices to look at our problem over the classifying stack of T , ie work with $p: [T \backslash G/B] \rightarrow BT$. The benefit of this viewpoint is that we only have to deal with representable maps. Further, as T is a split maximal torus, we know that the derived category of BT with quasicohherent cohomology is compactly generated. This fact allows for a genuine construction of the stable homotopy category SH with a six-functor formalism on BT (see Hoyois [22, Theorem 1.1]). This can be extended to relatively representable algebraic stacks over BT (see [30]). In [loc. cit.] Khan and Ravi show, using the works of Hoyois, that there is a K -theory spectrum in $\mathrm{SH}(BT)$ that represents genuine equivariant K -theory.

Using this version of the stable motivic homotopy category, we can analyze the “motive” of $[T \backslash G/B]$ relative to BT with coefficients in an E_∞ -ring spectrum $M_{BT} \in \mathrm{SH}(BT)$, ie $p_! p^! M_{BT}$. Because of technical reasons, we have to assume that M_{BT} satisfies some vanishing condition, namely for all $n \geq 0$ we have

$$(*) \quad \mathrm{Hom}_{\mathrm{SH}(BT)}(1_{BT} \langle n \rangle [-1], M_{BT}) = 0,$$

where $\langle n \rangle := (n)[2n]$ denotes the Tate-twist by n and shift by $2n$. This is not a drawback, as we will see in Example 2.14 that $(*)$ is satisfied for motivic cohomology and K -theory. As in the intersection theory case, we can use the Bruhat decomposition of G/B to stratify the flag variety via T -invariant affine cells and then compute $p_! p^! M_{BT}$ using this stratification.

The Khan–Ravi theory works even in the case where our base is not a field but an affine scheme. So, we will also prove our results in the most general case we are able to. Thus, from now on let S be an affine scheme and (G, B, T) be defined over S . First, we have to give a Bruhat decomposition in this setting (even though this is probably known to many people, we didn’t find a reference and proved it by ourselves).

Proposition 1 (Proposition 3.4) *Let S be a nonempty scheme (not necessarily affine). Let G be a split reductive S -group scheme with split maximal torus T and a Borel subgroup B containing T . Then the S -scheme G/B admits a cellular stratification indexed by the Weyl group of T in G .*

Afterward, we analyze the motive of a proper scheme X endowed with a group action of an S -group scheme H and an H -invariant cellular decomposition. In the special case of G/B with T -action this yields the structure of $p_! p^! M_{BT}$ as a M_{BT} -module with basis given by the classes of the Schubert cells.

Theorem 2 (Corollary 3.11) *Let G be a split reductive S -group scheme with maximal split torus T that is contained in a Borel subgroup B . Then*

$$p_! p^! M_{BT} \simeq \bigoplus_{w \in W} M_{BT} \langle l(w) \rangle,$$

where W is the Weyl group of T in G .

Applying this result with the representation of homotopy invariant K -theory in SH, we get the decomposition of homotopy invariant K -theory.

Theorem 3 (Corollary 3.12) *Let S be an affine scheme. Further, let G be a split reductive S -group scheme with maximal split torus T that is contained in a Borel subgroup B . Then we have*

$$\mathrm{KH}([T \backslash G/B]) \simeq \bigoplus_{w \in W} \mathrm{KH}(BT)$$

On the 0^{th} homotopy group we recover an integral version of the classical result, that $K_0^T(G/B)$ as an $R(T)$ -module is generated by the T -equivariant classes of the Schubert cells. For the higher equivariant K -groups, we get a similar statement.

Corollary 4 (Equation (1)) *Let S be a Noetherian regular affine scheme. Further, let G be a split reductive S -group scheme with maximal split torus T that is contained in a Borel subgroup B . Let $R(T)$ denote the integral representation ring of T . Then we have an isomorphism of $R(T)$ -modules:*

$$K_i^T(G/B) := K_i([T \backslash G/B]) \cong \bigoplus_{w \in W} K_i^T(S).$$

Now let us assume that $S = \mathrm{Spec}(k)$ is the spectrum of a field. On the higher homotopy groups, we similarly get an isomorphism of $R(T)_{\mathbb{Q}}$ -modules via tensoring with the higher K -groups of the ground field.

Corollary 5 (Equation (2)) *Let k be a field. Further, let G be a split reductive k -group scheme with maximal split torus T that is contained in a Borel subgroup B . Then we have an isomorphism of $R(T)_{\mathbb{Q}}$ -modules,*

$$K_i^T(G/B)_{\mathbb{Q}} \cong K_i(k)_{\mathbb{Q}} \otimes_{\mathbb{Q}} K_0^T(G/B)_{\mathbb{Q}}.$$

Using the universal property of SH, we can also extend these results to the lisse-extension of SH and the étale localized rational stable homotopy category $\mathrm{SH}_{\mathbb{Q}, \acute{\mathrm{e}}\mathrm{t}}$ (we have to assert some conditions on the base as seen in Section 3.3.2). In this way, we can extend Theorem 2 to the case of Beilinson motives and recover the analogous result on $A_T^*(G/B)$ and completed K_0 .

Proposition 6 (Equations (5) and (4)) *Let $S = \operatorname{Spec}(k)$ be the spectrum of a field. Further, let G be a split reductive S -group scheme with maximal split torus T that is contained in a Borel subgroup B . Then on completed equivariant K -theory, we have*

$$K_0^T(G/B)_{\mathbb{Q}}^{\wedge I_T} \cong K_0(G/B)_{\mathbb{Q}} \otimes_{\mathbb{Q}} K_0^T(S)_{\mathbb{Q}}^{\wedge I_T},$$

where I_T is the ideal generated by virtual rank 0-bundles in $R(T)_{\mathbb{Q}}$.

On Chow rings, we recover

$$A_T^*(G/B)_{\mathbb{Q}} \cong A^*(G/B)_{\mathbb{Q}} \otimes_{\mathbb{Q}} A_T^*(S)_{\mathbb{Q}}.$$

1.1 Notation

Categorical notation In this paper, we will without further mention use the language of ∞ -categories (see Lurie [36]). We will identify 1-categories with their Nerve and regard them as ∞ -categories. In particular, we will identify the category of sets with the full sub- ∞ -category of 0-truncated ∞ -groupoids and the category of groupoids with the full sub- ∞ -category of 1-truncated ∞ -groupoids. Likewise, when we say *full subcategory* of an ∞ -category, we will always mean a full sub- ∞ -category.

For the rest of this article, we fix an uncountable inaccessible regular cardinal κ and *small* will mean κ -small. Without further mention, if needed, we will assume smallness of the categories involved in this article. Indexing sets will always be small. We denote by $\mathbf{Cat}_{\infty}$ the ∞ -category of small ∞ -categories and by $\infty\text{-Grpd}$ the ∞ -category of small ∞ -groupoids.

A *presheaf* $\mathcal{F}$ on an ∞ -category $\mathcal{C}$ is a functor $\mathcal{F}: \mathcal{C}^{\text{op}} \rightarrow \infty\text{-Grpd}$. If $\mathcal{C}$ admits a Grothendieck topology τ , we will say that $\mathcal{F}$ is a τ -*sheaf* if it is a sheaf with respect to the topology τ .

Algebraic geometric notation Let S be a scheme. By an algebraic stack X over S , we mean an étale-sheaf of groupoids on S -schemes, such that the diagonal of X is representable by an algebraic space and there exists a scheme U and a smooth effective epimorphism $U \rightarrow X$. A morphism of algebraic stacks over S will always be an S -morphism.

Structure of this article

In [Section 2](#), we recall some facts about the stable homotopy category in our setting after Khan and Ravi. Afterward, we prove some basic facts, we need later on.

Our next step is to show the existence of the Bruhat decomposition of G/B over arbitrary schemes. We then continue to compute the motive of strict linear schemes with group action and apply this to $[T \backslash G/B]$.

We conclude this article, by applying our result to integral and rational homotopy invariant K -theory and their homotopy groups. Finally, we discuss how one extends these results to other motivic categories and get the classical results on Chow rings.

Acknowledgements

I would like to thank Torsten Wedhorn for his remarks on the earlier versions of this article and for sketching me his idea on the Bruhat decomposition. Also, I would like to thank Simon Pepin Lehalleur, for pointing out an error in the first version of this manuscript and the discussion with him afterward, that led to the corrected version. Further, I would also like to thank Timo Richarz and Thibaud van den Hove for various discussions concerning this paper.

This project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — project number 524431573, the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) TRR 326 *Geometry and Arithmetic of Uniformized Structures*, project number 444845124 and by the LOEWE grant “Uniformized Structures in Algebra and Geometry”.

2 Motivic setup

In this section, we fix an affine scheme S , a split reductive S -group scheme G together with a Borel pair (B, T) consisting of a split maximal torus T inside a Borel subgroup B of G . Further, any algebraic stack will be considered as an algebraic stack over S and any morphism will be relative over S .

In this article, we want to compute the motive of T -equivariant flag varieties. In particular, we are interested in motives of Artin stacks. There are several approaches on how to extend the theory of motives to Artin stacks. Recall from the introduction that one can right Kan extend $\mathrm{SH}_{\mathbb{Q}, \text{ét}}$ from schemes to Artin stacks and get a full 6-functor formalism. This works fine until one wants to compute the motivic cohomology in terms of K -theory. One can show that K -theory does not satisfy étale descent for Artin stacks.¹ For example, let $[X/G]$ be a smooth Artin stack over a field k with G split reductive. Then we have a map $K([X/G]) \rightarrow K^{\text{ét}}([X/G])$, where $K^{\text{ét}}$ is the right Kan extended K -theory from algebraic spaces to étale sheaves. This map is not an equivalence but realizes $K_0^{\text{ét}}([X/G])$ as the completion of $K_0^G(X)$ along the augmentation ideal $I_G \subseteq R(G) = K(\mathrm{Rep}(G))$ (note that $K_0^G(X)$ is in general not I_G -complete as seen for the $\mathbb{G}_m$ -equivariant K -theory of a point; see [Example 3.22](#)). This is an instance of the comparison between the Borel construction for K -theory and equivariant K -theory (see [\[33, Theorem 1.3\]](#)).

Also for nonrational coefficients, one has to be careful as étale descent is not even satisfied for schemes. Hence, one has to be careful to construct a full 6-functor formalism for SH with nonrational coefficients. This was done by Chowdhury in his thesis [\[10, Theorem 1.0.1\]](#) by gluing along smooth morphism with Nisnevich local sections. This is equivalent to gluing along smooth covers, the so-called *lissee-extension* (see [\[30, Corollary 12.28\]](#)). But again, computing the motivic cohomology along the *lissee-extended* K -theory spectrum yields the completion of K -theory along the augmentation ideal (see [\[30, Example 12.21\]](#)).

¹In fact, G -theory satisfies descent precisely for étale covers that are “isovariant” (see [\[26, Theorem 1.1\]](#)).

In the case of $\mathcal{X} := [T \backslash G/B]$ there is a construction by Khan and Ravi of a stable homotopy category $\mathrm{SH}(\mathcal{X})$ that admits a full 6-functor formalism and a motivic spectrum $\mathrm{KGL}_{\mathcal{X}} \in \mathrm{SH}(\mathcal{X})$ such that

$$\mathrm{Hom}_{\mathrm{SH}(\mathcal{X})}(1_{\mathcal{X}}, \mathrm{KGL}_{\mathcal{X}}) = \mathrm{KH}(\mathcal{X}),$$

where $\mathrm{KH}(\mathcal{X})$ denotes the homotopy invariant K -theory of $\mathcal{X}$ (see [30, Construction 10.3]). The quotient stack $\mathcal{X}$ belongs to a certain class of algebraic stacks, called *scalloped* (see below) for which the stable homotopy category is also defined.

In the end of this article, we will look at the implications on motivic cohomology in various frameworks (see Section 3.3.2).

2.1 Scalloped stacks

We recall the necessary definitions from [30, Section 2]. We will use the terminology of [loc. cit.].

Definition 2.1 Let H be a group scheme over S . We say that H is *nice* if it is an extension of an étale group scheme of order prime to the residue characteristics of S , by a group scheme of multiplicative type.

Example 2.2 (see [2, Remark 2.2]) An important example of a nice group scheme is a torus. One can show that any nice group scheme is linearly reductive. If S is the spectrum of a field of characteristic $p > 0$, then linearly reductive group schemes are also nice.

Let H be a nice S -group scheme and X a quasiasffine scheme with action by H . Hoyois constructs an equivariant version of the stable homotopy category $\mathrm{SH}^H(X)$ with full 6-functor formalism in this context (see [22, Theorem 1.1]). Khan and Ravi extend this construction via gluing along Nisnevich squares to a class of algebraic stacks, which they call *scalloped*. We don't want to give an explicit definition of a scalloped stack, as it is a bit technical, but give an important example and some properties of scalloped stacks, for the definition and details we refer to [30, Section 2].

Proposition 2.3 [30, Corollary 2.13, Theorem 2.14] (i) Let $f: X' \rightarrow X$ be a morphism of qcqs algebraic stacks. If X is scalloped and f is representable, then X' is scalloped.

(ii) Let X be a qcqs algebraic space over S with H -action, where H is a nice S -group scheme. Then $[X/H]$ is scalloped.

Throughout this article, we are interested in quotients of qcqs schemes by the torus T , such as $[T \backslash G/B]$. The above proposition tells us that these stacks are scalloped. In particular, we can work with the formalism of [30].

Notation 2.4 We set $\mathbf{Rep}_S$ to be the ∞ -category of morphisms $X' \rightarrow X$ of algebraic stacks over S that are representable. We denote by $\mathbf{Rep}_S^{\mathrm{ft}}$ the full subcategory of $\mathbf{Rep}_S$ consisting of morphisms of finite type over S . Further, we denote by $\mathbf{Sc}_S$ (resp. $\mathbf{Sc}_S^{\mathrm{ft}}$) the full subcategories of $\mathbf{Rep}_S$ (resp. $\mathbf{Rep}_S^{\mathrm{ft}}$) consisting of scalloped stacks.

2.2 The stable homotopy category for scalloped stacks

Let us quickly recall the construction of SH for scalloped stacks from [30]. For a scalloped stack X let us set $\mathbf{Sm}_X$ as the full subcategory of $(\mathbf{Rep}_S)_/X$ consisting of morphisms $X' \rightarrow X$ that are smooth and representable.

We define the homotopy category $\mathrm{H}(X)$ of X , as the ∞ -category of Nisnevich sheaves F from $\mathbf{Sm}_X$ to $\infty\text{-}\mathbf{Grpd}$ that are homotopy invariant, ie for any $X' \in \mathbf{Sm}_X$ and any vector bundle $p: V \rightarrow X'$, we have that the induced map $F(X') \rightarrow F(V)$ is an equivalence. In the classical motivic theory, for example of Cisinski and Déglise, one obtains the stable homotopy category by adjoining $\otimes$ -inverses of Thom-motives of finite locally free sheaves. In our case, we can associate to any finite locally free module $\mathcal{E}$ over X an object $\langle \mathcal{E} \rangle \in \mathrm{H}(X)$, called the *Thom-anima* (see [30, Section 4]). Now we obtain the stable homotopy category, by formally $\otimes$ -inverting these Thom-anima. One should note that formal $\otimes$ -inversion of objects in ∞ -categories is more delicate and we refer to [loc. cit.] for references and details.

Definition 2.5 [30] Let X be a scalloped algebraic stack. The *stable homotopy category of X* is defined as the ∞ -category

$$\mathrm{SH}(X) := \mathrm{H}(X)[\langle \mathcal{E} \rangle^{\otimes -1}],$$

where $[\langle \mathcal{E} \rangle^{\otimes -1}]$ denotes the formal inversion of all Thom-anima associated to any finite locally free module $\mathcal{E}$ over X .

The most important feature of the stable homotopy category for us is that the assignment $X \mapsto \mathrm{SH}(X)$ can be upgraded to a functor with a full 6-functor formalism, that satisfies homotopy invariance and yields a localization sequence. Further, homotopy invariant K -theory or motivic cohomology of X can be represented by objects in $\mathrm{SH}(X)$.

Let us quickly recall the 6-functor formalism for scalloped stacks on the stable homotopy category SH . We also recall the comparison with K -theory.

Theorem 2.6 [30] For any scalloped stack X there is an ∞ -category $\mathrm{SH}(X)$ with the following properties:

- (i) $\mathrm{SH}(X)$ is a stable, presentable, closed symmetric monoidal ∞ -category. The tensor product is colimit preserving and the inner Hom will be denoted by $\underline{\mathrm{Hom}}$. The $\otimes$ -unit will be denoted by 1_X .
- (ii) The assignment $X \mapsto \mathrm{SH}(X)$ can be upgraded to a presheaf of symmetric monoidal presentable ∞ -categories with colimit preserving functors on the site of scalloped stacks

$$\mathrm{SH}^*: (\mathbf{Sc}_S)^{\mathrm{op}} \rightarrow \mathbf{Cat}_{\infty}^{\otimes}, \quad X \mapsto \mathrm{SH}(X), \quad f \mapsto f^*.$$

For any morphism $f: X \rightarrow Y \in \mathbf{Sc}_S$, there is an adjunction

$$f^*: \mathrm{SH}(Y) \rightleftarrows \mathrm{SH}(X): f_*.$$

- (iii) (homotopy invariance) For every vector bundle $p: V \rightarrow X$ of scalloped stacks, the unit of the $*$ -adjunction

$$1 \rightarrow p_* p^*$$

is an equivalence.

- (iv) If $f \in \mathbf{Sc}_S$ is smooth morphism, then f^* has a left adjoint, denoted $f_{\#}$ that is a morphism of $\mathrm{SH}(Y)$ -modules.
- (v) The assignment $X \mapsto \mathrm{SH}(X)$ can be upgraded to a presheaf of presentable ∞ -categories,

$$\mathrm{SH}: (\mathbf{Sc}_S^{\mathrm{ft}})^{\mathrm{op}} \rightarrow \mathbf{Cat}_{\infty}, \quad X \mapsto \mathrm{SH}(X), \quad f \mapsto f^!,$$

from the ∞ -category of scalloped stacks with finite type representable morphisms. For each $f: X \rightarrow Y$ in $\mathbf{Sc}_S^{\mathrm{ft}}$, there is an adjunction

$$f_!: \mathrm{SH}(X) \rightleftarrows \mathrm{SH}(Y) : f^!.$$

For any factorization $f = p \circ j$ with j an open immersion and p a proper representable map, there is a natural equivalence $f_! \cong p_* j_{\#}$.

- (vi) (localization) If $i: Z \rightarrow X$ is a closed immersion of scalloped stacks with open complement $j: U \rightarrow X$, then we have the following cofiber sequences:

$$j_{\#} j^* \rightarrow \mathrm{id} \rightarrow i_* i^*, \quad i_! i^! \rightarrow \mathrm{id} \rightarrow j_* j^*.$$

- (vii) There is a map $K(X) \rightarrow \mathrm{Aut}(\mathrm{SH}(X))$, assigning for any $\alpha \in K(X)$ its twist $\langle \alpha \rangle$. If α is given by a finite locally free sheaf $\mathcal{E}$, then $\langle \mathcal{E} \rangle = p_{\#} s^* 1_X$ (this agrees with the previously considered Thom-anima), where $p: V(\mathcal{E}) \rightarrow X$ is the projection of the associated vector bundle and s its zero section. Further, any of the 6-operations commute with $\langle \alpha \rangle$ in a suitable sense (see [30, Remark 7.2]). We set $\langle n \rangle := \langle \mathcal{O}_X^n \rangle$.
- (viii) The canonical projection $p: \mathbb{G}_m \times X \rightarrow X$ yields a morphism $p_{\#} p^* 1_X[-1] \rightarrow 1_X[-1]$ whose fiber we denote $1_X(1)$. For an $M \in \mathrm{SH}(X)$, we denote its n -Tate twist by $M(n) := M \otimes 1_X(1)^{\otimes n}$. We have $\langle n \rangle \simeq (n)[2n]$.
- (ix) (purity) Let f be a smooth representable morphism of scalloped algebraic stacks with cotangent complex L_f , then

$$f^! \simeq f^* \langle L_f \rangle.$$

- (x) For a cartesian diagram in $\mathbf{Sc}_S^{\mathrm{ft}}$

$$\begin{array}{ccc} W & \xrightarrow{g'} & X \\ \downarrow f' & & \downarrow f \\ Y & \xrightarrow{g} & Z \end{array}$$

with $g \in \mathbf{Sc}_S^{\mathrm{ft}}$, we have

$$g^! f_* \xrightarrow{\simeq} f'_* g^!, \quad f^* g^! \xrightarrow{\simeq} g'_! f'^*.$$

- (xi) For $f: Z \rightarrow Y$ in $\mathbf{Sc}_S^{\text{ft}}$, the functor $f_!$ satisfies the projection formulas (see [30, Theorem 7.1]).
- (xii) There exists an E_∞ -ring spectrum $\mathbf{KGL}_X \in \mathbf{SH}(X)$ such that

$$\mathbf{KH}(X) = \underline{\mathbf{Hom}}_{\mathbf{SH}(X)}(1_X, \mathbf{KGL}_X),$$

that is functorial in smooth representable morphisms and satisfies Bott periodicity for twist by finite locally free sheaves (see [30, Theorem 10.7]).

In the rest of this article, we want to focus on modules over an E_∞ -ring spectrum $M \in \mathbf{SH}(X)$, where X is scalloped. The reason is that we will need a vanishing assumption (see below) that is satisfied for example for the homotopy invariant K -theory spectrum. As we are interested in T -equivariant K -theory of the flag variety, this is not a strong restriction.

For oriented cohomology theories, we can relax our situation from flag varieties to linear schemes. To be more precise, the flag variety is stratified by affine spaces. If we are interested in oriented cohomology theories, it is enough to consider objects that are stratified by vector bundles. The next example will show how this idea works.

But before we come to the example let us fix some notation.

Notation 2.7 Let X be a scalloped algebraic stack and $M_X \in \mathbf{SH}(X)$ an E_∞ -ring spectrum. Then we denote the ∞ -category of M_X -modules in $\mathbf{SH}(X)$ by $\mathbf{SH}(X)_M$. Further, for any representable morphism $f: Y \rightarrow X$ of scalloped algebraic stacks, we let $M_Y := f^*M_X$.

Remark 2.8 Let $f: X \rightarrow Y$ be a representable morphism of scalloped algebraic stacks. Further, let $M_Y \in \mathbf{SH}(Y)$ be an E_∞ -ring spectrum. Tensoring with M_Y (or M_X) induces a pullback functor $f_M^*: \mathbf{SH}(Y)_M \rightarrow \mathbf{SH}(X)_M$. As the $*$ -pullback for $\mathbf{SH}$ is monoidal, we see that its right adjoint is lax-monoidal. In particular, we get an adjunction

$$f_M^*: \mathbf{SH}(Y)_M \rightleftarrows \mathbf{SH}(X)_M : f_{*M}.$$

As remarked in Theorem 2.6(iv), if f is smooth, the left adjoint $f_\#$ of f^* is a morphism of $\mathbf{SH}(Y)$ -modules and in particular, induces a left adjoint $f_{\#M}$ of f_M^* (see [11, Section 7.2]).

By conservativity of the forgetful-functor $\mathbf{SH}_M \rightarrow \mathbf{SH}$ (here we see $\mathbf{SH}_M$ as a functor from scalloped algebraic stacks with representable morphisms to symmetric monoidal presentable ∞ -categories), we can use [30, Theorem 7.1] to obtain a 6-functor formalism on $\mathbf{SH}_M$ that satisfies properties (i)–(xi) of Theorem 2.6.

Notation 2.9 In the rest of this article, we will work with module spectra over some fixed E_∞ -ring spectrum. Thus, we will drop the subscript in the 6-functor formalism indicating the fixed E_∞ -ring spectrum as seen in Remark 2.8.

Definition 2.10 Let $f: X \rightarrow Y$ be a representable morphism of finite type of scalloped algebraic stacks. Further, let $M_Y \in \mathrm{SH}(Y)$ be an E_∞ -ring spectrum in $\mathrm{SH}(Y)$. Then we define *motive of X with values in M_Y* (resp. the *compactly supported motive of X with values in M_Y*) as $M_Y(X) := f_! f^! M_Y$ (resp. $M_Y^c(X) := f_* f^! M_Y$) in $\mathrm{SH}(Y)_M$.

Example 2.11 Let $\mathcal{E}$ be a finite locally free sheaf on a scalloped stack X and let $p: V(\mathcal{E}) \rightarrow X$ be the associated vector bundle. Further, let $M_X \in \mathrm{SH}(X)$ be an E_∞ -ring spectrum that admits an orientation, ie there is a functorial equivalence $M_X\langle \mathcal{F} \rangle \simeq M_X\langle n \rangle$ for any finite locally free $\mathcal{O}_X$ -module $\mathcal{F}$ of rank n . Note that by homotopy invariance, we have that $p_* p^* \simeq \mathrm{id} \simeq p_\# p^* \in \mathrm{SH}(X)_M$. In particular, by purity we have that $M_X(V(\mathcal{E}))$ and $M_X^c(V(\mathcal{E}))$ are equivalent to $M_X\langle \mathcal{E} \rangle$. Further, as we can orient the unit and by the Mayer–Vietoris sequence² we have that $M_X(V(\mathcal{E})) \simeq \bigoplus_{n \in \mathbb{Z}} M_X\langle \mathcal{E}_n \rangle$, where the $\mathcal{E}_n$ are the restrictions of $\mathcal{E}$ to the open and closed locus X_n such that $\mathcal{E}|_{X_n}$ has rank n . Further, as M_X admits an orientation, we have $M_X\langle \mathcal{E}_n \rangle \simeq M_X\langle n \rangle$.

Now let X be a qcqs algebraic space over S and $\mathcal{E}$ be a finite locally free sheaf on X . Further, let G be a nice S -group scheme acting on X such that $\mathcal{E}$ is G -equivariant. This yields a vector bundle over $[X/G]$, ie there is a finite locally free sheaf $\mathcal{E}_G$ on $[X/G]$ such that $[V(\mathcal{E})/G] \cong V(\mathcal{E}_G)$. As X is quasicompact, we can find a finite set $I_\mathcal{E} \subseteq \mathbb{N}_0$ and an open and closed cover $(X_i)_{i \in I_\mathcal{E}}$ of X such that for any $i \in I_\mathcal{E}$ the sheaf $\mathcal{E}|_{X_i}$ is finite locally free of rank i and $\mathcal{E} = \bigoplus_{i \in I_\mathcal{E}} \mathcal{E}_i$. As the trivial bundle $X \rightarrow [X/G]$ yields an atlas, we see that the same holds true for $\mathcal{E}_G$. In particular, as we have seen above this yields $M_{[X/G]}(V(\mathcal{E}_G)) \simeq M_{[X/G]}\langle \mathcal{E}_G \rangle = \bigoplus_{i \in I_\mathcal{E}} M_{[X/G]}\langle i \rangle$.

Lemma 2.12 (localization sequence) *Let $f: X \rightarrow Y$ be a representable morphism of scalloped algebraic stacks over S of finite type. Further, let $M_Y \in \mathrm{SH}(Y)$ be an E_∞ -ring spectrum. Let $i: Z \hookrightarrow X$ be a closed immersion over Y with open complement $j: U \rightarrow X$. Further, let us denote $f_0 := f \circ j$ and $\bar{f} := f \circ i$. Then for any M_X -module N in $\mathrm{SH}(X)$ there exists the fiber sequence in $\mathrm{SH}(Y)_M$*

$$\bar{f}_* \bar{f}^! N \rightarrow f_* f^! N \rightarrow f_{0*} f_0^! N.$$

Proof Applying the localization sequence

$$i_* i^! = i_! i^! \rightarrow \mathrm{id} \rightarrow j_* j^* = j_* j^!$$

to $f^! N$ yields

$$i_* \bar{f}^! N \rightarrow f^! N \rightarrow j_* f_0^! N.$$

Now applying f_* to this sequence yields the result. □

²Note that $f_* f^*$ and $f_! f^!$ satisfy Nisnevich descent and thus yield a Mayer–Vietoris sequence for M_X and M_X^c , ie for open substacks $U, U' \subseteq V(\mathcal{E})$, we have a fiber sequence of the form

$$M_X(U \cap U') \rightarrow M_X(U) \oplus M_X(U') \rightarrow M_X(V)$$

and similarly, for M_X^c .

We will need the following vanishing assumption later on in this article. In motivic cohomology, this is the analog of the vanishing of negative higher Chow groups. For the K -theory spectrum, this will follow from the vanishing of negative K -groups.

Assumption 2.13 Let $X = BH$ be the classifying stack of a nice group scheme H and $M_X \in \mathrm{SH}(X)$ be an E_∞ -ring spectrum. Further, let $n > 0$. Then we have

$$\mathrm{Hom}_{\mathrm{SH}(X)}(1_X \langle n \rangle [-1], M_X) = 0.$$

Assumption 2.13 is satisfied at least for the two cohomology theories, that are considered in this article. Namely, homotopy invariant K -theory and motivic cohomology.

Example 2.14 Assume S is Noetherian. Let $m < 0$ and $n \in \mathbb{Z}$ and $X = BH$ the classifying stack for a nice group scheme H . Further, let us consider the K -theory spectrum $\mathrm{KGL}_X \in \mathrm{SH}(X)$. Then we have

$$\mathrm{Hom}_{\mathrm{SH}(X)_{\mathrm{KGL}}}(\mathrm{KGL}_X \langle n \rangle [m], \mathrm{KGL}_X) \simeq \mathrm{Hom}_{\mathrm{SH}(X)_{\mathrm{KGL}}}(\mathrm{KGL}_X [m], \mathrm{KGL}_X) \simeq \pi_m \mathrm{KH}(X),$$

where the first equivalence follows from Bott-periodicity (see Theorem 2.6(xii)). As H is nice, the spectrum $\mathrm{KH}(X)$ is connective.³ In particular, since $m < 0$, we see that KGL_X satisfies Assumption 2.13.

Now let us assume that $S = \mathrm{Spec}(k)$ is the spectrum of a field and let us consider the motivic cohomology spectrum $M\mathbb{Z} \in \mathrm{SH}(X)$ (see [30, Construction 10.16]). Then

$$\mathrm{Hom}_{\mathrm{SH}(X)_{M\mathbb{Z}}}(M\mathbb{Z} \langle n \rangle [m], M\mathbb{Z}) = \mathrm{Hom}_{\mathrm{SH}(X)}(1_X, M\mathbb{Z} \langle -n \rangle [-m]) = A^{-n}(X, m),$$

which vanishes as m is negative.

The above example shows that the motivic K -theory spectrum and the motivic cohomology spectrum satisfy an even stronger condition than Assumption 2.13. Let us give this stronger assumption a number.

Assumption 2.15 Let $X = BH$ be the classifying stack of a nice group scheme H and $M_X \in \mathrm{SH}(X)$ be an E_∞ -ring spectrum. Further, let $n \in \mathbb{Z}$ and $m < 0$. Then we have

$$\mathrm{Hom}_{\mathrm{SH}(X)}(1_X \langle n \rangle [m], M_X) = 0.$$

3 T -equivariant motivic homotopy theory of flag varieties

Let S be a nonempty scheme, G be a split reductive S -group scheme and T a maximal split torus contained in a Borel subgroup $B \subseteq G$. We want to understand the motive of $[T \backslash G/B]$. In this case, the computations are rather straightforward, as G/B is cellular, ie has a stratification by vector bundles. So, one can filter G/B by T -invariant closed subschemes such that their successive differences are given by vector bundles

³Here we use that the derived category of complexes with quasicoherent cohomology $D_{\mathrm{qc}}(X)$ is compactly generated, as H is nice (combine [2, Proposition 6.14] and [1, Remark 12.2]). Then connectivity follows from [23, Theorem 5.7] (here we need that S is Noetherian).

(usually this property is called *linear* in the literature). The existence of such a stratification is well-known and referred to as the Bruhat decomposition of G/B . As we have only found references for split reductive groups over a field, we first recall the existence of an affine cell decomposition of G/B over S .

Afterward, we can analyze the motive of the scalloped stack $[T \backslash G/B]$ over BT . Note, however, that the 6-functor formalism of SH only works for representable morphisms of scalloped stacks, so we cannot compute the motive of $[T \backslash G/B]$ over S . We will however explain in [Section 3.3.2](#) how this extends to the lisse-extension of SH and Beilinson motives $\mathrm{DM}_{\mathbb{Q}}$, which both have 6-functor formalism for nonrepresentable morphisms.

3.1 Affine cell decomposition of G/B

In this section, we will show that G/B has an affine cell decomposition. We will use the Bruhat decomposition of G and pull back the induced stratification on $[B \backslash G/B]$ to G/B . This construction is compatible with base change and thus, we will reduce to the case $S = \mathrm{Spec}(\mathbb{Z})$. Then this is a classical statement on Schubert cells. This was communicated to us by Torsten Wedhorn.

Definition 3.1 A *stratification* of a scheme X is a map $\iota: \coprod_{i \in I} X_i \rightarrow X$, where I is a set, each X_i is a scheme, ι is a bijection on the underlying topological spaces, $\iota|_{X_i}$ is an immersion and the topological closure of $\iota(X_i)$ in X is the union of subsets of the form $\iota(X_j)$. The subschemes $\iota(X_i)$ of X are called *strata*.

Definition 3.2 An *S -cell* is an S -scheme isomorphic to a vector bundle. A *cellular S -scheme* X is a separated S -scheme of finite type which is smooth and admits a stratification whose strata are cells.

First, let us recall that G admits a Bruhat-decomposition indexed by the Weyl group W of a split maximal torus T inside G .

Lemma 3.3 (Bruhat-decomposition) *Let G be a split reductive S -group scheme and B a Borel subgroup of G containing a split maximal torus. Then G admits a stratification $\coprod_{w \in W} BwB \rightarrow G$, where W denotes the Weyl group of T in W .*

Proof See [\[12, exposé XXII, théorème 5.7.4\]](#). □

The Bruhat-decomposition yields a stratification $\coprod_{w \in W} [B \backslash BwB/B] \rightarrow [B \backslash G/B]$. The stack $[B \backslash G/B]$ can be identified with the quotient $G/B \times^G G/B$, where G acts diagonal via conjugation. For any S -scheme Q the set $G/B(Q)$ is in bijection with Borel-subschemas of G_Q (see [\[12, exposé XXII, corollaire 5.8.3\]](#)). Thus, we have a map

$$G/B \rightarrow [B \backslash G/B]$$

given on Q -valued points by $B' \mapsto (B', B)$. The pullback of the above stratification on $[B \backslash G/B]$ via this map yields a stratification $\coprod_{w \in W} C_w \rightarrow G/B$. We call the C_w *Schubert cells of G/B* .

On W we have a length function, which we denote by l (see [5, chapitre IV section 1.1]). Then we claim that $C_w \cong \mathbb{A}_S^{l(w)}$. In particular, G/B has a cellular stratification.

Proposition 3.4 *Let S be a nonempty scheme. Let G be a split reductive S -group scheme with split maximal torus T and a Borel subgroup B containing T . Then the Schubert cell C_w is isomorphic to $\mathbb{A}_S^{l(w)}$.*

In particular, the stratification of G/B by Schubert cells C_w is cellular.

Proof As the construction of the Schubert cells is compatible with base change, we may assume without loss of generality that $S = \text{Spec}(\mathbb{Z})$. Then the proposition follows from [24, II.13] $\square$

3.2 Equivariant motives of linear Artin stacks

In the following, we assume that S is an affine scheme and H an S -group scheme. Further, every scheme will be qcqs of finite type over S .

We fix an E_∞ -ring spectrum $M_{BH} \in \text{SH}(BH)$ that satisfies [Assumption 2.13](#).

Definition 3.5 A linear S -scheme $(X, (X_n)_{n \geq 0})$ consists of an S -scheme X and a filtration by closed subschemes

$$\emptyset = X_{-1} \hookrightarrow X_0 \hookrightarrow X_1 \hookrightarrow \dots \hookrightarrow X_n \hookrightarrow \dots \hookrightarrow X$$

such that each $X_{n-1} \rightarrow X_n$ is a closed immersion, each $X_n \setminus X_{n-1}$ is isomorphic to a coproduct vector bundles over S and the natural closed immersion $\text{colim}_n X_n \hookrightarrow X$ is an isomorphism on the reduced loci.

If $X_n \setminus X_{n-1}$ is isomorphic to a coproduct of affine spaces over S , we call $(X, (X_n)_{n \geq 0})$ *affinely linear*.

Definition 3.6 Let $(X, (X_n)_{n \geq 0})$ be a linear S -scheme such that X admits a H -action. We say that $(X, (X_n)_{n \geq 0})$ is H -equivariant each of the X_n is stabilized by H .

Remark 3.7 Let $(X, (X_n)_{n \geq 0})$ be a H -equivariant linear S -scheme and assume that X is quasicompact. Further, let us set $U_n := X_n \setminus X_{n-1} = \coprod_{j \in J_n} V(\mathcal{E}_{n,j})$. By H -invariance, we get an action of H on U_n . In particular, we can take the associated quotient stack U_n/H . As explained in [Example 2.11](#) this yields for each $j \in J_n$ a finite locally free H -equivariant sheaf $\mathcal{E}_{n,j,H}$, such $U_n/H \cong \coprod_{j \in J_n} V(\mathcal{E}_{n,j,H})$ together with a finite set $I_{n,j} \subseteq \mathbb{N}_0$ and a decomposition $S = \coprod_{j \in J_n} \coprod_{i \in I_{n,j}} S_{i,j}$ such that $\mathcal{E}_{n,j,H}|_{S_{i,j}}$ is finite locally free of rank i .

Definition 3.8 A linear S -scheme $(X, (X_n)_{n \geq 0})$ is called *strict* for all $n \geq 0$, we have that

$$\max \left\{ i \in \bigcup_{j \in J_{n-1}} I_{n-1,j} \right\} < \min \left\{ i \in \bigcup_{j \in J_n} I_{n,j} \right\},$$

where the notation is as in [Remark 3.7](#).

Example 3.9 Let us give the example, that motivates the definitions above. Let G be a split reductive S -group scheme with maximal split torus T contained in a Borel subgroup B . By [Proposition 3.4](#) the Schubert cells C_w of G/B are isomorphic to $\mathbb{A}_S^{l(w)}$. Let us set the closed subscheme X_n as the schematic image of $\bigsqcup_{l(w) \leq n} C_w$ inside G/B . This yields a linear structure on G/B by

$$X_0 \subseteq X_1 \subseteq \cdots \subseteq G/B,$$

where $X_n \setminus X_{n-1} \cong \bigsqcup_{l(w)=n} C_w$. By construction, each of the X_i are T -invariant and further, the linear structure $(G/B, (X_n)_{n \in \mathbb{N}_0})$ is strict. Thus, this construction yields a strict affinely linear T -equivariant structure on G/B .

From now on, we assume that H is a nice S -group scheme.

Theorem 3.10 Let $(X, (X_n)_{n \geq 0})$ be a H -equivariant linear S -scheme such that X is proper over S . Further, let us set $U_n := X_n \setminus X_{n-1} = \bigsqcup_{j \in J_n} V(\mathcal{E}_{n,j})$. Then the equivalences

$$M_{BH}([X/H]) \simeq \bigoplus_{n \geq 0} M_{BH}^c([U_n/H]) = \bigoplus_{n \geq 0} \bigoplus_{j \in J_n} \bigoplus_{i \in I_{n,j}} M_{BH}\langle i \rangle,$$

hold, where the notation is as in [Remark 3.7](#), if

- (i) M_{BH} admits an orientation and the linear structure above is strict, or
- (ii) the linear structure above is strict affinely linear.

Further, if M_{BH} satisfies [Assumption 2.15](#), then we can omit the strictness in (i) and (ii).

Proof We will prove the theorem under the assumption (i). The proof under the other assumptions will follow easily by the same arguments. By definition, X admits a filtration

$$X_0 \hookrightarrow X_1 \hookrightarrow \cdots \hookrightarrow X,$$

such that each $X_{i-1} \rightarrow X_i$ is a closed immersion with complement given by U_i . For simplicity, we will assume that every quotient in the following is taken with respect to the étale topology. By H -equivariance we may assume that each X_i is stabilized by H . We see that X/H admits a filtration

$$X_0/H \hookrightarrow X_1/H \hookrightarrow \cdots \hookrightarrow X/H,$$

where each of the $X_{n-1}/H \hookrightarrow X_n/H$ is a closed immersion with complement given by $U_n/H = \bigsqcup_{j \in J_n} V(\mathcal{E}_{n,j,H})$. As the X_n are proper, the X_n/H are also proper over BH and therefore

$$M_{c,BH}(X_n/H) \simeq M_{BH}(X_n/H)$$

by definition. In particular the localization sequence for motives with compact support (see [Lemma 2.12](#)) yields the fiber diagram

$$M_{BH}(X_{n-1}/H) \rightarrow M_{BH}(X_n/H) \rightarrow M_{BH}^c(U_n/H).$$

We claim that this sequence splits.

Indeed, as explained in [Remark 3.7](#), we have

$$M_{c,BH}([U_n/H]) \simeq \bigoplus_{j \in J_n} \bigoplus_{i_n \in I_{n,j}} M_{BH}\langle i_n \rangle.$$

By induction we may assume that

$$M_{BH}(X_{n-1}) \simeq \bigoplus_{k=0}^{n-1} \bigoplus_{j_k \in J_k} \bigoplus_{i_k \in I_{k,j_k}} M_{BH}\langle i_k \rangle.$$

In particular, any morphism $\delta: M_{c,BH}(U_n/H) \rightarrow M_{BH}(X_{n-1}/H)[1]$ corresponds to an element in

$$\prod_{k=0}^{n-1} \prod_{j_k \in J_k} \prod_{i_k \in I_{k,j_k}} \prod_{j \in J_n} \prod_{i_n \in I_{n,j}} \mathrm{Hom}_{\mathrm{SH}(BH)}(M_{BH}, M_{BH}\langle i_k - i_n \rangle[1]).$$

As $i_k - i_n < 0$ for any $i_k \in I_{k,j_k}$ with $0 \leq k \leq n-1$ and $i_n \in I_{n,j}$ by strictness, we see using [Assumption 2.13](#) that $\delta = 0$.

Now induction over n concludes the proof of the first assertion using that the motives only depend on the underlying reduced structure. $\square$

Corollary 3.11 *Let G be a split reductive S -group scheme with maximal split torus T that is contained in a Borel subgroup B of G . Then*

$$M_{BT}([T \backslash G/B]) \simeq \bigoplus_{w \in W} M_{BT}\langle l(w) \rangle.$$

Proof This immediately follows with [Theorem 3.10](#) and [Example 3.9](#). $\square$

3.3 Application to T -equivariant cohomology theories of flag varieties

In this subsection, we will prove [Theorem 3](#), [Corollaries 4](#) and [5](#) and [Proposition 6](#).

So, let S be a Noetherian regular affine scheme. In the following G is a split reductive S -group scheme with maximal split torus T contained in a Borel subgroup B of G .

3.3.1 Integral equivariant K -theory Let KGL_{BT} be the K -theory spectrum computing homotopy invariant K -theory for smooth representable stacks over BT (see [Theorem 2.6\(xii\)](#)). Note that KGL_{BT} satisfies [Assumption 2.13](#) by [Example 2.14](#). Then [Corollary 3.11](#) yields the following computation.

Corollary 3.12 *Let G be a split reductive S -group scheme with maximal split torus T that is contained in a Borel subgroup B of G . Then*

$$\mathrm{KH}([T \backslash G/B]) \simeq \bigoplus_{w \in W} \mathrm{KH}(BT).$$

Proof This follows from Bott periodicity⁴ of KGL , [Corollary 3.11](#) and [Theorem 2.6](#). $\square$

⁴The Bott periodicity yields $\mathrm{KGL}_X\langle \mathcal{E} \rangle \simeq \mathrm{KGL}_X$ for any finite locally free sheaf $\mathcal{E}$ over a scalloped stack X .

By construction $[T \backslash G/B]$ and BT are quotients of smooth Noetherian S -schemes by a nice S -group scheme. In particular, we see that $\mathrm{KH}([T \backslash G/B])$ and $\mathrm{KH}(BT)$ are connective and their homotopy groups are computed by genuine equivariant K -theory (see [23, Theorem 5.7]). Hence, for any $i \in \mathbb{Z}$ we have

$$(1) \quad K_i^T(G/B) = \bigoplus_{w \in W} K_i^T(S).$$

For rational K_0 this is nothing new over a field, as the Schubert classes yield an $R(T)_{\mathbb{Q}}$ -basis of $K_0^T(G/B)_{\mathbb{Q}}$ (see [31]).

Let $S = \mathrm{Spec}(k)$. With rational coefficients, we can use [32, Proposition A.5] to get isomorphisms of $R(T)$ -modules

$$(2) \quad K_i^T(G/B)_{\mathbb{Q}} \cong \bigoplus_{w \in W} K_i(k)_{\mathbb{Q}} \otimes_{\mathbb{Q}} R(T)_{\mathbb{Q}} \cong K_i(k)_{\mathbb{Q}} \otimes_{\mathbb{Q}} K_0^T(G/B)_{\mathbb{Q}}.$$

3.3.2 Completed equivariant K -theory In the following, we want to extend our results to other formalisms of motives. To look at these different definitions all at once, we use the formalism of a motivic ∞ -category $\mathcal{D}$ with full six-functor formalism on scalloped stacks (see [30, Proposition 5.13]). We do not want to give an explicit definition of a motivic ∞ -category, as it boils down to rewriting the axioms of the 6-functor formalism. But we want to give 2 examples that are of interest to us.

Example 3.13 The following 2 examples are motivic ∞ -categories with a full 6-functor formalism on scalloped stacks.

(1) Let X be an Artin stack over S . Let us now further assume that X is quasiseparated with quasiseparated representable diagonal and has a smooth cover that admits Nisnevich locally sections (we call such covers *smooth-Nisnevich*), eg any quasiseparated algebraic stacks with separated diagonal admits a smooth-Nisnevich cover (see [13, Theorem 1.2(1)]). The lisse-extended stable homotopy category $\mathcal{D} = \mathrm{SH}_{\triangleleft}$ is defined via

$$\mathrm{SH}_{\triangleleft}(X) := \mathrm{colim}_{(T,t)} \mathrm{SH}(T),$$

where limit is taken in the ∞ -category of pairs (T, t) , where $t: T \rightarrow X$ is a smooth morphism and T an algebraic space (see [30, Section 12]). The ∞ -category $\mathrm{SH}_{\triangleleft}(X)$ is equivalent to the right Kan extension of SH to smooth-Nisnevich stacks, evaluated at X (see [30, Corollary 12.28.]). This proves the extension of a full six-functor formalism, homotopy invariance and existence of a localization sequence to Artin stacks for this ∞ -category (this follows from the arguments of [28, Appendix A]). Note that for the existence of $\sharp$ -pushforward and $!$ -formalism there is no need for representability in this context. Further, they show that when S is the spectrum of a perfect field and X is the quotient of a smooth scheme by an algebraic group, the motivic cohomology spectrum $\mathrm{Hom}_{\mathrm{SH}_{\triangleleft}(X)}(1_X, M)$ for $M \in \mathrm{SH}_{\triangleleft}(X)$ can be computed by a Borel construction (see [30, Theorem 12.16]) — this is precisely the construction Edidin and Graham make to define equivariant Chow groups as seen in Example 3.20.

(2) Let us consider the étale localized rational stable homotopy category $\mathrm{SH}_{\mathbb{Q},\text{ét}}$. This ∞ -category can be right Kan extended to Artin stacks over S . As by definition $\mathrm{SH}_{\mathbb{Q},\text{ét}}$ satisfies étale descent, we can extend the 6-functor formalism, homotopy invariance and the localization sequence to $\mathrm{SH}_{\mathbb{Q},\text{ét}}$ on Artin stacks (see [28, Appendix A]). Again, there is no need for representability for the existence of a 6-functor formalism.

The universal property of the stable homotopy category yields a unique system of comparison maps $R_X: \mathrm{SH}(X) \rightarrow \mathbf{D}(X)$ (see [30, Proposition 5.13]) for any motivic ∞ -category $\mathbf{D}(X)$. The family of functors R is compatible with $\sharp$ -pushforward, $*$ -inverse image, Thom twists and tensor products.

As the family of functors R_X is monoidal, we can extend this to modules over any E_∞ -ring spectrum in $\mathrm{SH}(X)$, ie if $M_X \in \mathrm{SH}(X)$ is an E_∞ -ring spectrum, then there is a functor

$$R_M(X): \mathrm{SH}(X)_M \rightarrow R_X(M)\text{-Mod}(\mathbf{D}(X))$$

compatible with pullbacks of M and thus also $\sharp$ -pushforward, $*$ -inverse image, Thom twists and tensor products in module spectra.

Let us relate this construction to our examples above by applying it to motivic cohomology.

Example 3.14 Let X be an Artin stack over S and let $M\mathbb{Z} \in \mathrm{SH}(X)$ be the motivic cohomology ring spectrum.

(1) Let us now further assume that X admits a smooth-Nisnevich cover (see Example 3.13(1)). Let $M\mathbb{Z}^\triangleleft$ be the image of $M\mathbb{Z}$ under $R_X: \mathrm{SH}(X) \rightarrow \mathrm{SH}^\triangleleft(X)$.

The ∞ -category $\mathrm{SH}^\triangleleft(X)$ is equivalent to the right Kan extension of SH to smooth-Nisnevich stacks, evaluated at X (see Example 3.13(1)). Using this construction, the ∞ -category of $M_X^\triangleleft$ -modules in $\mathrm{SH}^\triangleleft(X)$ can be describe as follows.

We right Kan extends Spitzweck motives DM to prestacks along Nisnevich covers (see [9]). In this way, one can construct the exceptional pullback and pushforward for finite type morphisms. Further, for prestacks given by a quotient X/G of a scheme by a group scheme, the ∞ -category $\mathrm{DM}(X/G)$ is equivalent to $\lim_\Delta \mathrm{DM}(\mathrm{Bar}^\bullet(X, G))$, where $\mathrm{Bar}^\bullet(X, G)$ denotes the Bar resolution of X with respect to G . If G is special (eg $G = T$), then any étale G -torsor is Zariski locally trivial and in particular the étale sheafification of X/G , which we usually denoted by $[X/G]$, agrees with the Nisnevich sheafification. As seen in [loc. cit.] this allows one to compute $\mathrm{DM}([X/G]) = \mathrm{DM}(X/G) = \lim_\Delta \mathrm{DM}(\mathrm{Bar}^\bullet(X, G))$ (here $\mathrm{DM}(X/G)$ denotes the evaluation of DM at the presheaf quotient X/G).

(2) Let us assume that S is of finite type over an excellent Noetherian scheme of dimension ≤ 1 . Further, let us denote the image of $M\mathbb{Z}$ in $\mathrm{SH}(X)_{\mathbb{Q},\text{ét}}$ with $M\mathbb{Q}$. Then $M\mathbb{Q}$ can be glued from the Beilinson motivic cohomology spectrum along a smooth cover X and we can describe the $M\mathbb{Q}$ -modules in $\mathrm{SH}(X)_{\mathbb{Q},\text{ét}}$ as follows.

The right Kan extension of Beilinson motives $\mathrm{DM}_{\mathbb{Q}}$ to Artin stacks admits an extension of the full six-functor formalism (see [38]). This is achieved by gluing along smooth covers. There are different ways to see this, but we do not want to go into details and refer to the paragraph before [28, Section A.2]).

Remark 3.15 An important example in our context, that fits into the setting of [Example 3.13\(1\)](#), are quotient stacks of the form $[X/H]$, where X is a quasiseparated scheme and H is a nice group scheme (see [\[30, Remark 12.24\]](#)).

From now on let us assume that the six-functor formalism in $\mathbf{D}$ exists for not necessarily representable morphisms.

Notation 3.16 In the following we fix an E_∞ -ring spectrum $M_S \in \mathrm{SH}(S)$ we will denote its image in $\mathbf{D}(S)$ with $M_S^{\mathbf{D}}$. Again, we will denote the pullback of $M_S^{\mathbf{D}}$ under a map $X \rightarrow S$, where X is a scalloped stack, with $M_X^{\mathbf{D}}$. Further, we will denote the ∞ -category of $M_X^{\mathbf{D}}$ -modules in $\mathbf{D}(X)$ with $\mathbf{D}(X)_M$.

For any morphism scalloped stacks $f: X \rightarrow Y$ of finite type, we will denote the motives in $\mathbf{D}$ with $M_Y^{\mathbf{D}}(X) := f_! f^! M_X^{\mathbf{D}}$ and $M_Y^{c, \mathbf{D}} := f_* f^! M_X^{\mathbf{D}}$.

We also define *motivic cohomology of a scalloped stack X with coefficients in $M^{\mathbf{D}}$* as

$$H_{\mathbf{D}}^{n,m}(X, M) := \mathrm{Hom}_{\mathbf{D}(X)}(1_X, M_X^{\mathbf{D}}(n)[m]).$$

Remark 3.17 Assume $f: X \rightarrow S$ is a smooth scalloped stack over S . Then $f_! f^! M_S \simeq f_{\sharp} M_X$ and we see that

$$H_{\mathbf{D}}^{n,m}(X, M) \simeq \mathrm{Hom}_{\mathbf{D}(X)_M}(M_S^{\mathbf{D}}(X), M_S^{\mathbf{D}}(n)[m]).$$

If X is smooth over BH for some nice group scheme H , we therefore can transport all of the results of [Section 3.2](#) to $M_S(X)$ via $\sharp$ -pushforward along the structure map $BH \rightarrow S$.

Working over the base S , we can analyze the motive of strict linear schemes. This is classical, and the proof is achieved by the proof, mutatis mutandis, of [Theorem 3.10](#).

Proposition 3.18 *Let us assume that $H_{\mathbf{D}}^{i,1+2i}(S, M) = 0$ for all $i > 0$. Let $(X, (X_n)_{n \geq 0})$ be a linear S -scheme such that X is proper over S . Further, let us set $U_n := X_n \setminus X_{n-1} = \coprod_{j \in J_n} V(\mathcal{E}_{n,j})$. Then the equivalences*

$$M_S^{\mathbf{D}}(X) \simeq \bigoplus_{n \geq 0} M_S^{\mathbf{D},c}(U_n) = \bigoplus_{n \geq 0} \bigoplus_{j \in J_n} \bigoplus_{i \in I_{n,j}} M_S \langle i \rangle,$$

hold, where the notation is as in [Remark 3.7](#), if

- (i) $M_S^{\mathbf{D}}$ admits an orientation and the linear structure above is strict, or
- (ii) the linear structure above is strict affinely linear.

Further, if $H_{\mathbf{D}}^{i,1+2i}(S, M) = 0$ for all $i \in \mathbb{Z}$, then we can omit the strictness in (i) and (ii).

Proof This is completely analogous to the proof of [Theorem 3.10](#). □

The structure of the motive of linear stacks allows us to rewrite [Theorem 3.10](#).

Corollary 3.19 *Let H be a nice S -group scheme. Let $(X, (X_n)_{n \geq 0})$ be a H -equivariant strict linear S -scheme such that X is smooth and proper over S . Then*

$$M_S^D([X/H]) \simeq M_S^D(X) \otimes M_S^D(BH).$$

Proof This follows immediately from [Proposition 3.18](#) and [Theorem 3.10](#) via $\sharp$ -pushforward along the structure map $BH \rightarrow S$ (see [Remark 3.17](#)). $\square$

Example 3.20 Let us validate [Corollary 3.19](#) using a more direct computation over $S = \operatorname{Spec}(k)$, where k is a field and restricting ourselves to the case $\mathbf{D} = \operatorname{SH}_{\mathbb{Q}, \acute{e}t}$ and $M_S^D = M_{\mathbb{Q}, S} \in \mathbf{D}(S)$ the rational motivic cohomology ring spectrum.

Let $G = \operatorname{SL}_{2,S}$ and $T = \mathbb{G}_{m,S}$ the standard diagonal torus. Let B be the Borel subgroup of upper triangular matrices in G . Then $G/B \cong \mathbb{P}_S^1$ and the action of T on $\mathbb{P}_S^1$ induced by conjugation on G/B . In particular, the action of T on $\mathbb{P}_S^1$ is given by multiplication. The motive of $[T \backslash \mathbb{P}_k^1]$ can be computed in the following way — analogous to the computation of its intersection ring (see [\[15, Section 3.3\]](#)).

Let us fix an $i \in \mathbb{Z}$. Further, let V be a nontrivial representation of T over k . We denote by $\mathbb{V} := \operatorname{Spec}(\operatorname{Sym}(V))$ the associated vector bundle over S . Then we define the scheme U_i for each $h: Q \rightarrow S$ via

$$U_i(Q) := \{u \in \operatorname{Hom}(h^*\mathbb{V}, h^*\mathbb{V}^i) \mid \operatorname{Coker}(u) \text{ is finite free of rank } i\}$$

Then T acts freely on U_i and one can show that the codimension of U_i in the k -vector bundle $\mathbb{V}_i := \operatorname{Spec}(\operatorname{Sym}(V \otimes V^\vee)^i)$ is greater than $-i$. Further, we can see that

$$\mathbb{P}_k^1 \times_k^T U_{i-1} \cong \mathbb{P}(\mathcal{O}_{\mathbb{P}^{i-1}}(1) \oplus \mathcal{O}_{\mathbb{P}^{i-1}}(1)) \rightarrow \mathbb{P}_k^{i-1} \cong U_{i-1}/T$$

is a $\mathbb{P}_k^1$ -bundle. Thus, the projective bundle formula yields

$$M_S^D(\mathbb{P}_k^1 \times_k^T U_{i-1}) \cong M_S^D(\mathbb{P}_k^{i-1}) \oplus M_S^D(\mathbb{P}_k^{i-1})\langle 1 \rangle.$$

Now the motive of $M_S^D([T \backslash G/B])$ is isomorphic to the colimit over i of $M_S^D(\mathbb{P}_k^1 \times_k^T U_{i-1})$ (this can be followed from [\[21, Proposition A.7; 42, page 2107\]](#)). Therefore, we finally have

$$M_S^D([T \backslash G/B]) \simeq M_S^D(B \mathbb{G}_{m,S}) \oplus M_S^D(B \mathbb{G}_{m,S})\langle 1 \rangle \simeq M_S^D(\mathbb{P}_S^1) \otimes M_S^D(\mathbb{G}_{m,S}),$$

where the last equivalence follows again from the projective bundle formula.

Remark 3.21 The above example and computations also hold for $\operatorname{SH}_{\triangleleft}$ and the integral motivic cohomology ring spectrum if either k has characteristic 0 or after inverting the characteristic of k (see [\[30, Theorem 12.16\]](#)).

Let us specialize to the case of T acting on the flag variety G/B and the case where $S = \operatorname{Spec}(k)$ is the spectrum of a field. Assume for this paragraph that k has characteristic 0. As mentioned in [Example 3.13](#), the motivic cohomology spectrum can be computed using the Borel construction in the cases that are of interest to us. For Chow groups of stacks, this gives the right computation. For K -theory this is no longer true. In fact, one can show that the Borel construction yields the completion of equivariant K -theory along the augmentation ideal (see [\[33, Theorem 1.2\]](#)), ie

$$H_{\mathrm{SH}_{\triangleleft}}^{0,-i}([T \backslash G/B], \mathrm{KGL}^{\triangleleft}) \simeq K_i^T(G/B)^{\wedge_{I_T}},$$

where $I_T \subseteq R(T)$ is the ideal generated by virtual rank 0 representations and $K_i^T(G/B)^{\wedge_{I_T}}$ is the completion along $I_T K_i^T(G/B)$. Again the above stays true in characteristic $p > 0$, after inverting p .

Example 3.22 Let $G = \mathbb{G}_m$. Then $R(G) = \mathbb{Z}[T, T^{-1}]$ and the augmentation ideal I_G is generated by $1 - T$. Thus, $R(G)^{\wedge_{I_G}} = \mathbb{Z}[[T]]$ and we see that indeed $R(G)$ is not I_G -complete. Therefore, the lisse-extended K -theory spectrum does not recover K -theory.

A similar result appears when one wants to prove an equivariant form of the Riemann–Roch theorem (see [\[16\]](#)). The same holds, if we consider cohomology theories in $\mathrm{SH}_{\mathbb{Q}, \acute{\mathrm{e}}\mathrm{t}}$ as they satisfy étale descent. This descent property is the ambiguity here. One can show that even rational K -theory of stacks does not satisfy étale descent (for G -theory one can give precise conditions on quotient stacks, see [\[26, Section 3\]](#)). Nevertheless, we want to show the implications of our calculations for Chow groups and completed K -theory.

The upshot of [Corollary 3.19](#) is that it gives us a tensor description

$$(3) \quad M_S^D([T \backslash G/B]) = M_S^D(G/B) \otimes M_S^D(BT)$$

and we want to use this to get a tensor description of completed T -equivariant K -theory of G/B . In this case, we would need a Künneth formula for K -theory. For equivariant K_0 , in our special case, this is known and follows by the spectral sequence induced for example on G -theory (see [\[25, Theorem 4.1\]](#)). For higher equivariant K -groups there is no Künneth formula, as this fails even for nonequivariant K -theory.

Example 3.23 Let us consider $\mathbb{A}_k^1 \rightarrow \operatorname{Spec}(k)$ the projection of the affine line. Then

$$K_1(\mathbb{A}_k^2) = K_1(\mathbb{A}_k^1 \times_k \mathbb{A}_k^1) = k^\times$$

by homotopy invariance, whereas $K_1(\mathbb{A}_k^1) \otimes_{\mathbb{Z}} K_0(\mathbb{A}_k^1) \oplus K_0(\mathbb{A}_k^1) \otimes_{\mathbb{Z}} K_1(\mathbb{A}_k^1) = k^\times \oplus k^\times$.

Instead of a Künneth formula, one gets a spectral sequence for K -theory and higher Chow theory, at least when one of the factors comes from a linear scheme (see [\[25, Theorem 4.3\]](#)). In fact, Totaro shows that Chow groups commute with tensor products if and only if one of the associated motives of the factors is Tate (relative over a field, see [\[42, Theorem 7.2\]](#)). Thus, if X is a (smooth) strict linear S -scheme and Y an arbitrary (smooth) scheme, we have $A^*(X \times_S Y) = A^*(X) \otimes_{\mathbb{Z}} A^*(Y)$. In [\[25\]](#) this is also a

result of the associated spectral sequence of motivic cohomology of linear varieties. Dugger and Isaksen generalize this idea to arbitrary cellular motivic cohomology theories like motivic cohomology, algebraic K -theory and algebraic cobordism (see [14]).

Proposition 3.24 (Tor spectral sequence) *Let $\mathbf{D} = \mathrm{SH}_{\mathbb{Q}, \text{ét}}$. Let M_S be either $\mathrm{KGL}_{\mathbb{Q}, S}$ or $M_{\mathbb{Q}, S}$ inside $\mathbf{D}(S)$. Let $(X, (X_n)_{n \geq 0})$ be a linear S -scheme such that X is proper over S and let Y be a smooth algebraic stack. Then for each $n \in \mathbb{Z}$ there is a spectral sequence*

$$\mathrm{Tor}_p^{H_{\mathbf{D}}^{*,n}(S, M)}(H_{\mathbf{D}}^{*,n}(X, M), H_{\mathbf{D}}^{*,n}(Y, M))_q \Rightarrow H_{\mathbf{D}}^{p+q,n}(X \times_S Y, M).$$

Proof By Proposition 3.18 the motive $M_S^{\mathbf{D}}(X)$ is a direct sum of Tate-motives. Therefore, the result follows with [14, Theorem 6.2, 6.4, Proposition 7.7]. $\square$

We can use Proposition 3.24 to see that

$$(4) \quad K_0^T(G/B)_{\mathbb{Q}}^{\wedge I_T} \cong K_0(G/B)_{\mathbb{Q}} \otimes_{\mathbb{Q}} K_0^T(S)_{\mathbb{Q}}^{\wedge I_T}$$

noting that the K -theory of G/B and BT is connective (see Section 3.3.1).

The comparison of Beilinson motivic cohomology with higher Chow groups yield

$$H_{\mathrm{DM}}^{n, 2n}(X, \mathbb{Q}) = A^n(X)_{\mathbb{Q}},$$

as for $k < n$, we have $H_{\mathrm{DM}}^{k, n}(X, \mathbb{Q}) = A^n(X, 2k - 2n)_{\mathbb{Q}} = 0$. If $S = \mathrm{Spec}(k)$ is the spectrum of a field, we can use that equivariant Chow groups are given via the Borel construction, we see as before that

$$(5) \quad A_T^*(G/B)_{\mathbb{Q}} \cong A^*(G/B)_{\mathbb{Q}} \otimes_{\mathbb{Q}} A_T^*(S)_{\mathbb{Q}}.$$

Remark 3.25 It should not be difficult to generalize Proposition 3.24 to the case of Spitzweck motives and get an integral version of the above results for completed K -theory and Chow theory, at least after inverting the characteristic of the ground field (in the positive characteristic setting). But as this probably boils down to just rewriting the results of Dugger and Isaksen, we did not follow this further and leave it to the reader.

References

- [1] **J Alper**, *Good moduli spaces for Artin stacks*, PhD thesis, Stanford University (2008) Available at <https://www.proquest.com/docview/304468872>
- [2] **J Alper, J Hall, D Rydh**, *The étale local structure of algebraic stacks*, preprint (2019) [arXiv 1912.06162](https://arxiv.org/abs/1912.06162)
- [3] **D Anderson, S Griffeth, E Miller**, *Positivity and Kleiman transversality in equivariant K -theory of homogeneous spaces*, J. Eur. Math. Soc. 13 (2011) 57–84 [MR](#) [Zbl](#)
- [4] **J Ayoub**, *Les six opérations de Grothendieck et le formalisme des cycles évanescents dans le monde motivique, I*, Astérisque 314, Soc. Math. France, Paris (2007) [MR](#) [Zbl](#)

- [5] **N Bourbaki**, *Groupes et algèbres de Lie, Chapitres 4–6*, Actualités Scientifiques et Industrielles 1337, Hermann, Paris (1968) [MR](#) [Zbl](#)
- [6] **M Brion**, *Equivariant Chow groups for torus actions*, Transform. Groups 2 (1997) 225–267 [MR](#) [Zbl](#)
- [7] **M Brion**, *Lectures on the geometry of flag varieties*, from “Topics in cohomological studies of algebraic varieties”, Birkhäuser, Basel (2005) 33–85 [MR](#) [Zbl](#)
- [8] **G Carlsson, R Joshua**, *Equivariant algebraic K -theory, G -theory and derived completions*, Adv. Math. 430 (2023) art. id. 109194 [MR](#) [Zbl](#)
- [9] **R Cass, T van den Hove, J Scholbach**, *The geometric Satake equivalence for integral motives*, preprint (2022) [arXiv 2211.04832](#)
- [10] **C Chowdhury**, *Motivic homotopy theory of algebraic stacks*, Ann. K -Theory 9 (2024) 1–22 [MR](#) [Zbl](#)
- [11] **D-C Cisinski, F Déglise**, *Triangulated categories of mixed motives*, Springer (2019) [MR](#) [Zbl](#)
- [12] **M Demazure**, *Groupes réductifs: Déploiements, sous-groupes, groupes-quotients*, from “Schémas en groupes, Tome III: Structure des schémas en groupes réductifs, (SGA 3_{III})”, Lecture Notes in Math. 153, Springer (1970) exposé 22 [MR](#)
- [13] **N Deshmukh**, *On the motivic homotopy type of algebraic stacks*, preprint (2023) [arXiv 2304.10631](#)
- [14] **D Dugger, D C Isaksen**, *Motivic cell structures*, Algebr. Geom. Topol. 5 (2005) 615–652 [MR](#) [Zbl](#)
- [15] **D Edidin, W Graham**, *Equivariant intersection theory*, Invent. Math. 131 (1998) 595–634 [MR](#) [Zbl](#)
- [16] **D Edidin, W Graham**, *Riemann–Roch for equivariant Chow groups*, Duke Math. J. 102 (2000) 567–594 [MR](#) [Zbl](#)
- [17] **W Fulton**, *Intersection theory*, Ergebnisse der Math. 2, Springer (1984) [MR](#) [Zbl](#)
- [18] **W Graham**, *Positivity in equivariant Schubert calculus*, Duke Math. J. 109 (2001) 599–614 [MR](#) [Zbl](#)
- [19] **S Griffeth, A Ram**, *Affine Hecke algebras and the Schubert calculus*, European J. Combin. 25 (2004) 1263–1283 [MR](#) [Zbl](#)
- [20] **J Hornbostel, V Kiritchenko**, *Schubert calculus for algebraic cobordism*, J. Reine Angew. Math. 656 (2011) 59–85 [MR](#) [Zbl](#)
- [21] **V Hoskins, S Pepin Lehalleur**, *On the Voevodsky motive of the moduli stack of vector bundles on a curve*, Q. J. Math. 72 (2021) 71–114 [MR](#) [Zbl](#)
- [22] **M Hoyois**, *The six operations in equivariant motivic homotopy theory*, Adv. Math. 305 (2017) 197–279 [MR](#) [Zbl](#)
- [23] **M Hoyois, A Krishna**, *Vanishing theorems for the negative K -theory of stacks*, Ann. K -Theory 4 (2019) 439–472 [MR](#) [Zbl](#)
- [24] **J C Jantzen**, *Representations of algebraic groups*, 2nd edition, Math. Surv. Monogr. 107, Amer. Math. Soc., Providence, RI (2003) [MR](#) [Zbl](#)
- [25] **R Joshua**, *Algebraic K -theory and higher Chow groups of linear varieties*, Math. Proc. Cambridge Philos. Soc. 130 (2001) 37–60 [MR](#) [Zbl](#)
- [26] **R Joshua**, *Riemann–Roch for algebraic stacks, I*, Compos. Math. 136 (2003) 117–169 [MR](#) [Zbl](#)
- [27] **D Kazhdan, G Lusztig**, *Schubert varieties and Poincaré duality*, from “Geometry of the Laplace operator”, Proc. Sympos. Pure Math. 36, Amer. Math. Soc., Providence, RI (1980) 185–203 [MR](#) [Zbl](#)

- [28] **A A Khan**, *Virtual fundamental classes of derived stacks, I*, preprint (2019) [arXiv 1909.01332](#)
- [29] **A A Khan, C Ravi**, *Equivariant generalized cohomology via stacks*, preprint (2022) [arXiv 2209.07801](#)
- [30] **A A Khan, C Ravi**, *Generalized cohomology theories for algebraic stacks*, Adv. Math. 458 (2024) art. id. 109975 [MR](#) [Zbl](#)
- [31] **B Kostant, S Kumar**, *T-equivariant K-theory of generalized flag varieties*, J. Differential Geom. 32 (1990) 549–603 [MR](#) [Zbl](#)
- [32] **A Krishna**, *Riemann–Roch for equivariant K-theory*, Adv. Math. 262 (2014) 126–192 [MR](#) [Zbl](#)
- [33] **A Krishna**, *The completion problem for equivariant K-theory*, J. Reine Angew. Math. 740 (2018) 275–317 [MR](#) [Zbl](#)
- [34] **M Levine**, *K-theory and motivic cohomology of schemes*, preprint (1999)
- [35] **M Levine, S Pauli**, *Quadratic counts of twisted cubics*, preprint (2022) [arXiv 2206.05729](#)
- [36] **J Lurie**, *Higher topos theory*, Ann. of Math. Stud. 170, Princeton Univ. Press (2009) [MR](#) [Zbl](#)
- [37] **F Morel, V Voevodsky**, $\mathbb{A}^1$ -homotopy theory of schemes, Inst. Hautes Études Sci. Publ. Math. 90 (1999) 45–143 [MR](#) [Zbl](#)
- [38] **T Richarz, J Scholbach**, *The intersection motive of the moduli stack of shtukas*, Forum Math. Sigma 8 (2020) art. id. e8 [MR](#) [Zbl](#)
- [39] **H Schubert**, *Kalkül der abzählenden Geometrie*, Teubner, Leipzig (1879) [Zbl](#)
- [40] **M Spitzweck**, *A commutative $\mathbb{P}^1$ -spectrum representing motivic cohomology over Dedekind domains*, Mém. Soc. Math. France 157, Soc. Math. France, Paris (2018) [MR](#) [Zbl](#)
- [41] *The Stacks project*, electronic reference (2005–) Available at <http://stacks.math.columbia.edu>
- [42] **B Totaro**, *The motive of a classifying space*, Geom. Topol. 20 (2016) 2079–2133 [MR](#) [Zbl](#)
- [43] **V Uma**, *Equivariant K-theory of flag varieties revisited and related results*, Colloq. Math. 132 (2013) 151–175 [MR](#) [Zbl](#)

TU Darmstadt

Darmstadt, Germany

yaylali@mathematik.tu-darmstadt.de

Received: 31 August 2023

Revised: 1 February 2024

Small Heegaard genus and $SU(2)$

JOHN A BALDWIN

STEVEN SIVEK

Let Y be a closed, orientable 3-manifold with Heegaard genus 2. We prove that if $H_1(Y; \mathbb{Z})$ has order 1, 3, or 5, then there is a representation $\pi_1(Y) \rightarrow SU(2)$ with nonabelian image. Similarly, if $H_1(Y; \mathbb{Z})$ has order 2 then we find a nonabelian representation $\pi_1(Y) \rightarrow SO(3)$. We also prove that a knot K in S^3 is a trefoil if and only if there is a unique conjugacy class of irreducible representations $\pi_1(S^3 \setminus K) \rightarrow SU(2)$ sending a fixed meridian to $\text{diag}(i, -i)$.

57R58; 57K31, 57M12

1 Introduction

The instanton Floer homology of a homology 3-sphere Y is generated as a chain complex by conjugacy classes of irreducible representations $\pi_1(Y) \rightarrow SU(2)$, so it is natural to ask whether these exist. This question has been studied from many different perspectives in recent years, including surgeries on knots [25], branched double covers [46], Stein fillings [3], and incompressible tori [4; 32]; in every case so far the answer has been yes, unless $Y \cong S^3$.

In this article, we add small Heegaard genus to the list. If Y has a Heegaard splitting of genus at most 1 then $\pi_1(Y)$ is abelian, so the first interesting case is Heegaard genus 2. In this case we can apply work of Birman and Hilden [7] to identify such manifolds as branched covers of links, and thus reduce the problem to one about representations of link groups. Our main result is the following.

Theorem 1.1 *Let Y be a rational homology 3-sphere with Heegaard genus 2, and with first homology $H_1(Y; \mathbb{Z})$ of order 1, 3, or 5. Then there is a representation $\pi_1(Y) \rightarrow SU(2)$ with nonabelian image.*

In general, we will say that Y is $SU(2)$ -abelian if every homomorphism $\pi_1(Y) \rightarrow SU(2)$ has abelian image; then manifolds satisfying the hypotheses of Theorem 1.1 are *not* $SU(2)$ -abelian.

One could ask the same question about rational homology 3-spheres with other homology groups: for fixed $n \geq 1$, if Y has Heegaard genus 2 and $|H_1(Y)| = n$, must there be a nonabelian representation $\pi_1(Y) \rightarrow SU(2)$? The answer is negative for many other values of n . For example:

- If $n \geq 4$ is even, then the connected sum $\mathbb{R}P^3 \# L(n/2, 1)$ is $SU(2)$ -abelian.

- For $n = |24g + 11|$, where $g \neq 0, -1$, the toroidal manifolds $Y = Y(T_{2,3}, T_{2,2g+1})$ studied by Zentner [46] (and originally constructed by Motegi [35, Section 3]) are $SU(2)$ -abelian. These have Heegaard genus 2 because they are branched double covers of 3-bridge knots $L(T_{2,3}, T_{2,2g+1})$ [46, Theorem 4.14]; see [42, Figure 1]. For example, when $n = 35$ the branch locus is identified in [46, Section 5] as the knot 8_{16} .
- The smallest known hyperbolic example, pointed out to us by Nathan Dunfield, is $\text{Vol}3 = m007(3, 1)$, with $n = 18$. It has Heegaard genus 2 because it is the branched double cover of the 3-bridge link $L10n46$.
- The smallest known hyperbolic example with n odd is $m036(-3, 2)$, with $n = 45$; it is the branched double cover of the 3-bridge knot 8_{18} , or equivalently the 4-fold branched cyclic cover of the figure eight knot. (See [11, Section 6], or [42, Example 5.6] for more discussion.) It is also the unique double cover of $\text{Vol}3$.

Unfortunately, we do not expect that the strategy we use to prove [Theorem 1.1](#) will work for larger odd values of $|H_1(Y)|$, because the order-5 case already requires several special results that do not apply to higher orders. One is a theorem of Li and Liang [31, Theorem 1.4] asserting that if the instanton knot homology $KHI(K)$ of Kronheimer and Mrowka [28] has rank 1 in Alexander gradings $\pm g, \pm(g-1)$, and 0, and it vanishes everywhere else, then either K or its mirror is an *instanton L-space knot* of genus g . The analogous claim for $\widehat{HFK}(K)$ is certainly true but does not seem to exist in the Heegaard Floer literature. The second theorem, by Farber, Reinoso, and Wang [16], says that the $(2, 5)$ torus knot is the only instanton L-space knot of genus 2. We would like to generalize the first to larger values of $\dim KHI(K)$, and the second to genera $g \geq 3$, but both of these seem entirely out of reach.

[Theorem 1.1](#) makes use of some facts about $SU(2)$ -simple knots, whose definition we recall now. Here and in the sequel we view $SU(2)$ as the unit quaternions; this identifies the trace of a 2×2 matrix with twice the real part of the corresponding quaternion.

Definition 1.2 [46] Given a link $L \subset S^3$, a representation

$$\rho: \pi_1(S^3 \setminus L) \rightarrow SU(2)$$

is *meridian-traceless* if $\text{Re}(\rho(\mu)) = 0$ for every meridian μ of L . We then say that L is $SU(2)$ -*simple* if every meridian-traceless representation is conjugate to one with image in the binary dihedral group $\{e^{i\theta}\} \cup \{e^{i\theta}j\}$.

What we really show en route to proving [Theorem 1.1](#) is that the only $SU(2)$ -simple knots of determinant 1 or 3 are the unknot and the trefoils $T_{\pm 2,3}$, and that the only $SU(2)$ -simple knots of determinant 5 and bridge index at most 3 are the figure eight and the cinquefoils $T_{\pm 2,5}$.

Building on this, we can study the representation variety

$$\mathcal{R}(K, i) = \{\rho: \pi_1(S^3 \setminus K) \rightarrow SU(2) \mid \rho(\mu) = i\},$$

which consists of representations that send a fixed meridian μ of K to the quaternion i . Kronheimer and Mrowka [28, Theorem 7.17] proved that $\mathcal{R}(K, i) \cong \{*\}$ if and only if K is the unknot. For the trefoils,

$$\mathcal{R}(T_{\pm 2,3}, i) \cong \{*\} \sqcup S^1,$$

and [5, Conjecture 1.9] says that this should uniquely characterize the trefoils. The classification of $SU(2)$ -simple knots of determinant 3 allows us to prove that this is indeed the case.

Theorem 1.3 *Suppose for a knot $K \subset S^3$ that $\mathcal{R}(K, i) \cong \{*\} \sqcup S^1$. Then K is a trefoil.*

Partial results toward Theorem 1.3 were known under the assumption that the S^1 component consists of *nondegenerate* representations: Kronheimer and Mrowka [28, Corollary 7.20] first proved that K must be fibered, and then in [5, Theorem 1.10] we proved that K must be a trefoil. Theorem 1.3 removes the nondegeneracy assumption, establishing the conjecture in full generality.

Along the way to classifying $SU(2)$ -simple knots of small determinant, we deduce the following result which may be independently useful, as part of Theorem 3.3.

Theorem 1.4 *Let $K \subset S^3$ be an $SU(2)$ -simple knot whose determinant is prime, and write its Alexander polynomial as*

$$\Delta_K(t) = \sum_{i=-g}^g a_i t^i$$

where g is the Seifert genus of K . Then $(-1)^{i+\sigma(K)/2} a_i \geq 0$ for all i , where $\sigma(K)$ is the signature of K , and $|a_g| \geq 1$ with equality if and only if K is fibered.

In fact, in Theorem 3.3 we show that $(-1)^{i+\sigma(K)/2} a_i$ is nonnegative because it is equal to the dimension of the summand $KHI(K, i)$ of the instanton knot homology of K .

The reason we insist that $\det(K)$ be prime in Theorem 1.4 is that it allows us to verify a nondegeneracy condition for binary dihedral representations of $\pi_1(S^3 \setminus K)$, following work of Heusener and Klassen [19]; we do not know whether the same conclusions should hold when K is $SU(2)$ -simple but $\det(K)$ is composite. These nondegeneracy conditions allow us to show that $\dim KHI(K) = \det(K)$, and then in the determinant-3 case, we know from [5] that the only such knots are the trefoils. When $\det(K) = 5$ this is not quite enough, and we also need to understand Stein fillings of $SU(2)$ -abelian manifolds; see Proposition 4.3 for details.

Finally, the reader might wonder what happens when $|H_1(Y)|$ is even. In this case we have to consider $SU(2)$ -simplicity for links rather than knots, and less is known in this setting, but recent work of Xie and Zhang [45] provides a starting point. Even so, it is hard to guarantee the existence of $SU(2)$ representations, because our methods might lead to $SO(3)$ representations that do not lift to $SU(2)$. (When $|H_1(Y)|$ is odd, there is no obstruction to lifting.) But we can often at least find nonabelian $SO(3)$ representations, and in particular we prove the following.

Theorem 1.5 *If Y has Heegaard genus 2 and satisfies $|H_1(Y; \mathbb{Z})| = 2$, then there is a representation*

$$\pi_1(Y) \rightarrow \mathrm{SO}(3)$$

with nonabelian image.

As in the case of determinants 1 and 3, what we really prove is [Proposition 7.5](#), asserting that the Hopf link is the only $\mathrm{SU}(2)$ -simple link of determinant 2, and then [Theorem 1.5](#) follows from the special case of 3-bridge links.

Organization

In [Section 2](#) we prove [Theorem 1.1](#) for integral homology spheres. In [Section 3](#) we study the instanton knot homology of $\mathrm{SU}(2)$ -simple knots, and in [Section 4](#) we prove [Proposition 4.3](#), asserting that simply connected Stein fillings of many $\mathrm{SU}(2)$ -abelian manifolds must have negative definite intersection form. In [Section 5](#) we apply these to prove [Theorem 1.1](#) in the cases where $H_1(Y)$ has order 3 or 5. [Section 6](#) uses this work to prove [Theorem 1.3](#), that the variety $\mathcal{R}(K, i)$ detects the trefoils. Finally, [Section 7](#) is devoted to the proof of [Theorem 1.5](#).

Acknowledgements

We are grateful to Tye Lidman, Mike Miller Eismeier, Danny Ruberman, and Raphael Zentner for interesting conversations regarding this work. We also thank Nathan Dunfield for telling us about $\mathrm{Vol}3$ and its character variety, and the referee for their feedback. Baldwin was supported by NSF FRG grant DMS-1952707.

2 From $\mathrm{SU}(2)$ -simple knots to $\mathrm{SU}(2)$ -abelian 3-manifolds

Given a knot $K \subset S^3$, we will let $\Sigma_2(K)$ denote the branched double cover of K . The proof of [Theorem 1.1](#) then relies on the following proposition.

Proposition 2.1 *For each odd integer $d \geq 1$, the map $K \mapsto \Sigma_2(K)$ gives a well-defined surjection*

$$(2-1) \quad \left\{ \begin{array}{l} \text{SU}(2)\text{-simple 3-bridge knots} \\ K \subset S^3 \text{ of determinant } d \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \text{SU}(2)\text{-abelian 3-manifolds } Y \text{ with} \\ \text{Heegaard genus 2 and } |H_1(Y)| = d \end{array} \right\}.$$

Proof We first check that this map is well defined. If K is an $\mathrm{SU}(2)$ -simple 3-bridge knot, then its branched double cover is $\mathrm{SU}(2)$ -abelian by [\[42, Theorem 1.6\]](#). Moreover, $\Sigma_2(K)$ certainly has Heegaard genus at most 2, since we can split K in half using a bridge sphere, and the branched double cover of the 3-ball over either half will be a genus-2 handlebody. In fact, the Heegaard genus of $\Sigma_2(K)$ must be exactly 2, since otherwise $\Sigma_2(K)$ would be S^3 or a lens space, and then Hodgson and Rubinstein [\[20\]](#) proved that K would have bridge index at most 2. So this proves that $\Sigma_2(K)$ belongs to the codomain, as claimed.

We now wish to show that the map is surjective, so let Y be an element of the codomain. Since Y has Heegaard genus 2, Birman and Hilden [7, Theorem 1] proved that it is the branched double cover of a 3-bridge link $L \subset S^3$. Writing $\mathbb{F} = \mathbb{Z}/2\mathbb{Z}$, we have a short exact sequence of chain complexes

$$0 \rightarrow C_*(S^3, L; \mathbb{F}) \rightarrow C_*(\Sigma_2(L); \mathbb{F}) \rightarrow C_*(S^3; \mathbb{F}) \rightarrow 0,$$

where the first map is a transfer map and the second is induced by projection, and this leads to an isomorphism

$$(2-2) \quad H_1(\Sigma_2(L); \mathbb{F}) \cong H_1(S^3, L; \mathbb{F}) \cong \tilde{H}_0(L; \mathbb{F})$$

using the long exact sequence of the pair (S^3, L) . Thus if L has ℓ components then

$$0 = \dim_{\mathbb{F}} H_1(Y; \mathbb{F}) = \dim_{\mathbb{F}} H_1(\Sigma_2(L); \mathbb{F}) = \ell - 1,$$

so $\ell = 1$ and we can write $Y = \Sigma_2(K)$ where $K = L$ is a 3-bridge knot.

We know from [46, Lemma 3.2] and the fact that $H_1(\Sigma_2(K); \mathbb{F}) = 0$ that $\Sigma_2(K)$ is $SU(2)$ -abelian if and only if it is $SO(3)$ -abelian, and the latter implies that K is $SU(2)$ -simple by [46, Proposition 3.1], so if Y is $SU(2)$ -abelian then K must be $SU(2)$ -simple. Thus K belongs to the preimage of Y , proving the desired surjectivity. $\square$

Remark 2.2 We note from the proof of Proposition 2.1 (and in particular from [46, Section 3]) the fact that if a knot $K \subset S^3$ has an $SU(2)$ -abelian branched double cover, then K must be $SU(2)$ -simple. This holds regardless of the bridge index of K or the Heegaard genus of $\Sigma_2(K)$. By contrast, an $SU(2)$ -simple knot of bridge index greater than 3 need not have an $SU(2)$ -abelian branched double cover.

The following fact constrains the knots considered in Proposition 2.1, though we will not use it in this paper.

Proposition 2.3 *All $SU(2)$ -simple 3-bridge knots are hyperbolic.*

Proof Suppose that K is an $SU(2)$ -simple 3-bridge knot, but that it is not hyperbolic. Then K is either a torus knot or a satellite [43, Corollary 2.5], so we address these cases separately below.

Suppose first that K is a torus knot. Schubert [40, Satz 10] showed that the only 3-bridge torus knots are $T(3, q)$ where $|q| > 3$, so $K = T(3, q)$ for some such q . Then $\det(K)$ is either 1 or 3 depending on whether q is odd or even, so $\Sigma_2(K)$ is an $SU(2)$ -abelian Seifert fibered space, with first homology of order 1 or 3. The only such manifolds are S^3 and lens spaces [42, Theorem 1.2], contradicting the fact that $\Sigma_2(K)$ has Heegaard genus 2.

Now suppose that K is a satellite knot, say with pattern $P \subset S^1 \times D^2$ and companion C . Let $\alpha > 0$ be the wrapping number of P , meaning the minimal number of times that some meridional disk of $S^1 \times D^2$ intersects P . Then Schubert [40, Satz 3] proved that the bridge index satisfies

$$b(K) \geq \alpha \cdot b(C).$$

Since $b(K) = 3$ and the knottedness of C implies that $b(C) \geq 2$, we must have $\alpha = 1$; hence K is a connected sum $C \# C'$. But then $b(C \# C') = b(C) + b(C') - 1$ by [40, Satz 7], so C and C' are both 2-bridge knots, and thus the $\mathrm{SU}(2)$ -abelian manifold

$$\Sigma_2(K) \cong \Sigma_2(C) \# \Sigma_2(C')$$

is a connected sum of two lens spaces of odd order. Now $\pi_1(\Sigma_2(K)) \cong (\mathbb{Z}/p\mathbb{Z}) * (\mathbb{Z}/q\mathbb{Z})$ for some odd integers $p, q \geq 3$, and there is an irreducible representation $\pi_1(\Sigma_2(K)) \rightarrow \mathrm{SU}(2)$ defined by sending the generators of the $\mathbb{Z}/p\mathbb{Z}$ and $\mathbb{Z}/q\mathbb{Z}$ factors to $e^{2\pi i/p}$ and $e^{2\pi j/q}$ respectively, so $\Sigma_2(K)$ is not $\mathrm{SU}(2)$ -abelian after all and we have a contradiction. $\square$

Proposition 2.1, together with some gauge-theoretic input from Kronheimer and Mrowka [28], suffices to establish the homology sphere case of **Theorem 1.1**.

Proof of Theorem 1.1 in the case $H_1(Y; \mathbb{Z}) = 0$ We wish to show that there are no $\mathrm{SU}(2)$ -abelian integer homology 3-spheres of Heegaard genus 2. By **Proposition 2.1**, it suffices to show that there are no nontrivial $\mathrm{SU}(2)$ -simple knots of determinant 1, so let us suppose that such a knot K exists. Then Klassen [24, Theorem 10] proved that there are exactly $\frac{1}{2}(\det(K) - 1) = 0$ conjugacy classes of nonabelian representations $\pi_1(S^3 \setminus K) \rightarrow \mathrm{SU}(2)$ with binary dihedral image. By assumption, every representation

$$\rho: \pi_1(S^3 \setminus K) \rightarrow \mathrm{SU}(2)$$

such that $\mathrm{Re} \rho(\mu) = 0$ is binary dihedral, so now every such ρ has abelian image. But Kronheimer and Mrowka [28, Corollary 7.17] proved that since K is a nontrivial knot, there is at least one nonabelian ρ with $\mathrm{Re} \rho(\mu) = 0$, a contradiction. $\square$

Remark 2.4 Zentner [46, Proposition 9.1] gave a slightly different proof of the fact that a homology sphere which is the branched double cover of a knot cannot be $\mathrm{SU}(2)$ -abelian. We argue differently here because the idea of counting binary dihedral representations will turn out to be useful for manifolds with nontrivial first homology.

3 Instanton knot homology and $\mathrm{SU}(2)$ -simple knots

When Y is not an integer homology sphere, we can still occasionally say something about representations $\pi_1(Y) \rightarrow \mathrm{SU}(2)$, though it will require substantially more machinery. In preparation, we study the instanton knot homology [28] of $\mathrm{SU}(2)$ -simple knots with prime determinant.

Proposition 3.1 *Let $K \subset S^3$ be an $\mathrm{SU}(2)$ -simple knot, with determinant an odd prime p . Then its signature satisfies $|\sigma(K)| \leq p - 1$ and $\sigma(K) \equiv p - 1 \pmod{4}$, and its instanton knot homology satisfies $\dim_{\mathbb{Q}} KHI(K) = p$.*

Proof Since K is $SU(2)$ -simple, all of the irreducible $SU(2)$ representations $\rho: \pi_1(S^3 \setminus K) \rightarrow SU(2)$ with $\operatorname{Re}(\rho(\mu)) = 0$ are conjugate to binary dihedral ones, and again by [24, Theorem 10] there are $n = \frac{1}{2}(\det(K) - 1)$ of these up to conjugacy. Thus by [46, Proposition 7.3], we will have

$$\dim_{\mathbb{Q}} I^{\natural}(K) = 2n + 1 = \det(K),$$

where $I^{\natural}(K)$ is the reduced singular instanton knot homology of [29], if each such representation ρ satisfies the following two conditions:

- (a) $\dim_{\mathbb{R}} H^1(S^3 \setminus K; \mathfrak{su}(2)_{\rho}) = 1$, and
- (b) the restriction map $H^1(S^3 \setminus K; \mathfrak{su}(2)_{\rho}) \rightarrow H^1(\mu; \mathfrak{su}(2)_{\rho})$ is onto.

If ρ is one of these irreducible, binary dihedral representations, with image in $\{e^{i\theta}\} \cup \{e^{i\theta}j\}$, then one can extract from the proof of [24, Theorem 10] that it has finite image, of order dividing $4\det(K) = 4p$. (For completeness, we will work out the details explicitly in Lemma 7.3 and Remark 7.4.) Since $\rho(\mu)$ is traceless it has order 4, so then the order of $\operatorname{Im}(\rho)$ is also a multiple of 4, and it is not equal to 4 or else $\operatorname{Im}(\rho)$ would be abelian. Then $\operatorname{Im}(\rho)$ has order exactly $4p$ since p is prime. Now Proposition 3.2 below asserts that ρ satisfies both of the desired conditions.

We deduce from this that $\dim_{\mathbb{Q}} I^{\natural}(K) = \det(K) = p$ after all, and then the isomorphism

$$I^{\natural}(K; \mathbb{Q}) \cong KHI(K; \mathbb{Q})$$

of [29, Proposition 1.4] tells us that $\dim_{\mathbb{Q}} KHI(K) = p$ as claimed. These computations of the twisted cohomology of $S^3 \setminus K$ also suffice to ensure that the Casson–Lin invariant $h(K)$, defined in [34] and reinterpreted using gauge theory by Herald [18], is in fact a signed count of the conjugacy classes of irreducible representations with $\operatorname{Re}(\rho(\mu)) = 0$. It follows that

$$|h(K)| \leq \frac{1}{2}(\det(K) - 1),$$

but Lin [34, Corollary 2.10] proved that $h(K) = \frac{1}{2}\sigma(K)$, so this establishes the inequality

$$|\sigma(K)| \leq \det(K) - 1.$$

Finally, the congruence $\sigma(K) \equiv \det(K) - 1 \pmod{4}$ actually holds for arbitrary knots $K \subset S^3$, as an immediate consequence of the evenness of $\sigma(K)$ and the relation

$$(3-1) \quad \det(K) \equiv (-1)^{\sigma(K)/2} \pmod{4}$$

of [36, Theorem 5.6], so the proof is complete. □

The following proposition completes the missing part of Proposition 3.1. The hard technical details of its proof are essentially contained in [19], though we have to do some work to extract them.

Proposition 3.2 *Let $K \subset S^3$ be a knot, and let $\rho: \pi_1(S^3 \setminus K) \rightarrow SU(2)$ be a representation whose image is binary dihedral of order $4p$ for some odd prime p . Then*

- (a) $\dim_{\mathbb{R}} H^1(S^3 \setminus K; \mathfrak{su}(2)_\rho) = 1$, and
 (b) the restriction map $H^1(S^3 \setminus K; \mathfrak{su}(2)_\rho) \rightarrow H^1(\mu; \mathfrak{su}(2)_\rho)$ is onto.

Proof Since $\rho: \pi_1(S^3 \setminus K) \rightarrow \mathrm{SU}(2)$ has binary dihedral image of order 4 times an odd prime, Heusener and Klassen [19, Section 3] proved that $H^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$ is 1-dimensional, verifying (a). In fact, their proof also implies that the restriction map

$$(3-2) \quad H^1(S^3 \setminus K; \mathfrak{su}(2)_\rho) \rightarrow H^1(\mu; \mathfrak{su}(2)_\rho) \cong \mathbb{R}$$

is onto, as we will now explain. To set the stage, we assume without loss of generality that $\rho(\mu) = j$. Then $H^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$ is the group $Z^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$ of cocycles, which consists of functions

$$\xi: \pi_1(S^3 \setminus K) \rightarrow \mathfrak{su}(2)$$

such that

$$\xi(gh) = \xi(g) + \mathrm{Ad}_{\rho(g)} \xi(h) \quad \text{for all } g, h,$$

modulo the group $B^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$ of coboundaries, which have the form

$$\zeta^\#: \pi_1(S^3 \setminus K) \rightarrow \mathfrak{su}(2), \quad g \mapsto \zeta - \mathrm{Ad}_{\rho(g)} \zeta$$

for some $\zeta \in \mathfrak{su}(2)$. (Here we view $\mathfrak{su}(2)$ as the vector space of purely imaginary quaternions.) If $\zeta = ai + bj + ck$, then since $\rho(\mu) = j$ we have

$$\zeta^\#(\mu) = (ai + bj + ck) - \mathrm{Ad}_j(ai + bj + ck) = (ai + bj + ck) - (-ai + bj - ck) = 2(ai + ck),$$

so the j -component of $\zeta^\#(\mu)$ is zero. It follows that if $\xi \in Z^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$ is a 1-cocycle for which $\xi(\mu)$ has nonzero j -component, then its restriction to an element of $Z^1(\mu; \mathfrak{su}(2)_\rho)$ is not a coboundary, so this is how we will prove the surjectivity of (3-2).

In order to determine $Z^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$, Heusener and Klassen take a Wirtinger presentation [19, Equation (5)] of the form

$$\pi_1(S^3 \setminus K) = \langle S_1, \dots, S_n \mid S_{j_\ell}^{\epsilon_\ell} S_\ell S_{j_\ell}^{-\epsilon_\ell} = S_{\ell+1}, \ell = 1, \dots, n-1 \rangle$$

with each $\epsilon_\ell \in \{\pm 1\}$, and turn this into a system of $n-1$ linear equations in $n+1$ real variables $s_1, \dots, s_n, t$, given in [19, Equation (14)] as

$$(3-3) \quad \alpha(j_\ell, \ell) s_{j_\ell} - s_\ell - s_{\ell+1} - \epsilon_{j_\ell} \beta(j_\ell, \ell) t = 0, \quad 1 \leq \ell \leq n-1.$$

Writing $\rho(S_\ell) = \zeta_\ell j$ for a complex unit ζ_ℓ as in [19, Equation (6)], they show that a collection of values $\xi(S_\ell) = x_\ell i + y_\ell j + z_\ell k$ determines a cocycle if and only if

- we have $x_1 = x_2 = \dots = x_n$ (coming from [19, Equation (9)]), and
- if we write $y_\ell + z_\ell i = \zeta_\ell(t_\ell + s_\ell i)$, then $t_1 = t_2 = \dots = t_\ell$ have some common value t , and $(s_1, \dots, s_n, t)$ is a solution to (3-3).

Now $\dim H^1(S^3 \setminus K; \mathfrak{su}(2)_\rho) = 1$ implies that $\dim Z^1(S^3 \setminus K; \mathfrak{su}(2)_\rho) = 4$, as explained after [19, Remark 5], so the space of solutions to (3-3) is 3-dimensional.

Suppose we fix $\mu = S_1$ and $\rho(\mu) = j$ as above, and let $\xi \in Z^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$ be a cocycle corresponding to a solution $(s_1, \dots, s_n, t)$ of (3-3). Following the above notation, $\zeta_1 = 1$ and so $y_1 + z_1 i = t_1 + s_1 i$; this means that the j -coordinate y_1 of $\xi(\mu)$ is nonzero if and only if $t_1 = t$ is. Thus it will suffice to find a solution of (3-3) with $t \neq 0$, since this will give a cocycle $\xi \in Z^1(S^3 \setminus K; \mathfrak{su}(2)_\rho)$ whose restriction to $Z^1(\mu; \mathfrak{su}(2)_\rho)$ is not a coboundary.

The proof of [19, Theorem 1] concludes by determining that the simpler system

$$(3-4) \quad \alpha(j_\ell, \ell) s_{j_\ell} - s_\ell - s_{\ell+1} = 0, \quad 1 \leq \ell \leq n-1$$

of $n-1$ equations in n variables has at most a 2-dimensional solution space (its rank is at least $n-2$). If every solution to (3-3) has $t = 0$, then it follows that the space of solutions to (3-3) is also at most 2-dimensional, since (by forgetting the t coordinate) it injects into the space of solutions to (3-4). But we know that (3-3) actually has a 3-dimensional solution space, so there must be a solution with $t \neq 0$. Thus the map (3-2) must be onto after all, verifying (b). $\square$

We can use Proposition 3.1 to say a bit more about the instanton knot homology of knots which satisfy its hypotheses. We first recall that if K has genus g , then its instanton knot homology comes with a symmetric Alexander decomposition

$$KHI(K) = \bigoplus_{i=-g}^g KHI(K, i),$$

each of whose summands has a natural $\mathbb{Z}/2\mathbb{Z}$ -grading. Kronheimer and Mrowka [27] and Lim [33] proved that this recovers the Alexander polynomial $\Delta_K(t)$, by the formula

$$(3-5) \quad \pm \Delta_K(t) = \sum_{i=-g}^g \chi(KHI(K, i)) \cdot t^i,$$

where the sign ambiguity comes from using different conventions for the grading.

Theorem 3.3 *Let $K \subset S^3$ be an $SU(2)$ -simple knot for which $\det(K)$ is prime. Then each Alexander summand $KHI(K, i)$ is supported in a single $\mathbb{Z}/2\mathbb{Z}$ grading. Moreover, if K has genus g and Alexander polynomial*

$$\Delta_K(t) = \sum_{i=-g}^g a_i t^i,$$

then the following must be true:

- $a_i = (-1)^{i+\sigma(K)/2} \dim_{\mathbb{Q}} KHI(K, i)$ for all i , and
- $|a_g| \geq 1$, with equality if and only if K is fibered.

Proof Proposition 3.1 says that $\dim_{\mathbb{Q}} KHI(K) = \det(K)$. Now substituting $t = -1$ into (3-5), taking absolute values, and applying the triangle inequality gives

$$\det(K) = \left| \sum_{i=-g}^g \chi(KHI(K, i)) \cdot (-1)^i \right| \leq \sum_{i=-g}^g |(-1)^i \chi(KHI(K, i))| \leq \sum_{i=-g}^g \dim_{\mathbb{Q}} KHI(K, i) = \det(K).$$

In particular, equality must hold in both of the inequalities above. We deduce from the second one that $|\chi(KHI(K, i))| = \dim_{\mathbb{Q}} KHI(K, i)$ for each i , so each $KHI(K, i)$ is supported in a single mod 2 grading. The first one says that all of the nonzero terms $(-1)^i \chi(KHI(K, i))$ must have the same sign, say $(-1)^{\delta_0}$, and then the mod 2 grading of each $KHI(K, i)$ is $i + \delta_0$. Then

$$\chi(KHI(K, i)) = (-1)^{i+\delta_0} \dim_{\mathbb{Q}} KHI(K, i),$$

so we substitute this information back into (3-5) to get

$$\pm \Delta_K(t) = \sum_{i=-g}^g ((-1)^{i+\delta_0} \dim_{\mathbb{Q}} KHI(K, i)) t^i.$$

Let δ be either δ_0 or $\delta_0 + 1$, depending on whether the left side is $+\Delta_K(t)$ or $-\Delta_K(t)$. Then

$$a_i = (-1)^{i+\delta} \dim_{\mathbb{Q}} KHI(K, i)$$

for all i , and so we have

$$\Delta_K(-1) = \sum_{i=-g}^g (-1)^i a_i = (-1)^{\delta} \sum_{i=-g}^g \dim_{\mathbb{Q}} KHI(K, i) = (-1)^{\delta} \cdot \det(K).$$

But we also know in general that $\Delta_K(-1) = (-1)^{\sigma(K)/2} \det(K)$: this follows from the congruence (3-1) and the fact that

$$\Delta_K(-1) = a_0 + \sum_{i=1}^g ((-1)^i + (-1)^{-i}) a_i = a_0 + 2 \sum_{i=1}^g (-1)^i a_i$$

is congruent modulo 4 to $a_0 + 2 \sum_{i=1}^g a_i = \Delta_K(1) = 1$. Thus we have $(-1)^{\delta} = (-1)^{\sigma(K)/2}$, completing the determination of a_i in terms of $KHI(K, i)$.

The last detail is the claim about $|a_g|$, which is equal to $\dim_{\mathbb{Q}} KHI(K, g)$. We know that this dimension is strictly positive [28, Proposition 7.6], and that equality holds if and only if K is fibered [27, Proposition 4.1], so the claim follows immediately. $\square$

4 Symplectic fillings of $SU(2)$ -abelian manifolds

An *instanton L-space* is a rational homology 3-sphere Y whose framed instanton homology has the smallest possible rank, namely $\dim_{\mathbb{Q}} I^{\#}(Y) = |H_1(Y)|$. By analogy with Heegaard Floer homology [38, Theorem 1.4], we expect that any symplectic filling of an instanton L-space should have negative definite

intersection form. We do not prove this here, but in this section we will prove this claim for a restricted class of instanton L-spaces which will suffice for our purposes. We begin with the following definition.

Definition 4.1 [10] An $SU(2)$ -abelian rational homology 3-sphere Y has *cyclically finite* fundamental group if for every representation

$$\rho: \pi_1(Y) \rightarrow SU(2),$$

the finite cover $\tilde{Y}$ whose fundamental group $\pi_1(\tilde{Y})$ is the kernel of

$$\text{ad } \rho: \pi_1(Y) \rightarrow SO(3)$$

is a rational homology sphere.

For example, in Definition 4.1 it suffices to know that the universal abelian cover of Y is a rational homology sphere. The following are Proposition 4.9 and Corollary 4.10 of [3].

Lemma 4.2 [3] Suppose that Y is an $SU(2)$ -abelian rational homology sphere, and that $|H_1(Y)| \leq 5$ or that $H_1(Y)$ is cyclic of order a prime power. Then $\pi_1(Y)$ is cyclically finite.

In particular, [3, Theorem 4.6] says that if Y is an $SU(2)$ -abelian rational homology sphere and $\pi_1(Y)$ is cyclically finite, then Y is an instanton L-space. We now prove an analogue of [38, Theorem 1.4] for such manifolds.

Proposition 4.3 Let Y be an $SU(2)$ -abelian rational homology sphere, and suppose that $\pi_1(Y)$ is cyclically finite. Then any simply connected, weak symplectic filling (X, ω) of a contact structure on Y must satisfy $b_2^+(X) = 0$.

Proof Let (X, ω) be a weak symplectic filling of (Y, ξ) , with $\pi_1(X) = 1$, and suppose that $b_2^+(X) \geq 1$. We attach a Weinstein 2-handle H along a Legendrian torus knot $T_{2,5}$ with Thurston–Bennequin invariant 3, inside a Darboux ball $B^3 \subset Y$, to get a symplectic cobordism to a new (Y', ξ') . Then following Eliashberg [14], we can construct a concave symplectic filling (Z_0, ω_0) of (Y', ξ') from an open book decomposition, by first performing 0-surgery on the binding B and then extending the symplectic form across a Lefschetz fibration with concave boundary $Y'_0(B)$. Kronheimer and Mrowka [26, Lemma 11] observed that one can take $H_1(Z_0) = 0$ by choosing a collection of vanishing cycles which generate the homology of the base, but the same argument with π_1 instead of H_1 allows us to take Z_0 to be simply connected.

With this in mind, we write $Z = H \cup_{Y'} Z_0$, and note that Z contains a surface of positive self-intersection: this is the core of the handle H , glued to a Seifert surface for the attaching curve $T_{2,5}$ in Y . Thus we have a closed symplectic manifold $W = X \cup_Y Z$ such that

- both X and Z are simply connected, and
- both $b_2^+(X)$ and $b_2^+(Z)$ are strictly positive.

We choose a class $w \in H^2(W; \mathbb{Z})$ and a symplectic form on W with integral cohomology class, and if the latter class has Poincaré dual $h \in H_2(W; \mathbb{Z})$ then we know from [28; 41] that the Donaldson invariants

$$D_W^w(h^k)$$

are nonzero for all large enough k in some residue class mod 4.

On the other hand, the assumption that Y is $\mathrm{SU}(2)$ -abelian means that all critical points of the Chern–Simons functional on Y are reducible, and in fact the Chern–Simons functional on Y is Morse–Bott since $\pi_1(Y)$ is cyclically finite [3, Corollary 4.5]. Under these circumstances, Austin and Braam [1, Proposition 6.3] proved that any Donaldson invariant of the form $D_W^w(x^k)$ must vanish, where $x \in H_2(W; \mathbb{Z})$ is a class whose restriction to $H_1(Y)$ is zero. Letting $n = |H_1(Y)|$, the restriction of $x = nh$ to $H_1(Y)$ is certainly zero, so the Donaldson invariant

$$D_W^w((nh)^k) = n^k \cdot D_W^w(h^k)$$

must be zero as well, but this contradicts the nonvanishing result mentioned above. It follows that we must have had $b_2^+(X) = 0$ after all. $\square$

Remark 4.4 In the proof of Proposition 4.3, we can choose the class $w \in H^2(W; \mathbb{Z})$ to be nonzero mod 2 on each of X and Z , and then the desired vanishing for the invariants $D_W^w(h^k)$ follows equally well from the “dimension counting argument” outlined by Donaldson following the statement of [12, Theorem 4.9]. The key facts we need to carry out this argument are that every flat $\mathrm{SU}(2)$ connection A on Y is reducible and nondegenerate (meaning that $H_A^1(Y) = 0$), or in other words that Y is $\mathrm{SU}(2)$ -abelian and $\pi_1(Y)$ is cyclically finite. This allows us to glue the relevant moduli spaces of instantons over X and Z , each with the same asymptotic limit at Y , as described for example in [13, Section 4.4.1].

We remark that Proposition 4.3 should apply equally well to fillings that are not simply connected, but we do not need this more general statement here. We also observe the following corollary; cf [30, Theorem 2.1] or [38, Theorem 1.4].

Corollary 4.5 *Let Y be an $\mathrm{SU}(2)$ -abelian rational homology sphere. If $\pi_1(Y)$ is cyclically finite, then Y does not admit a coorientable taut foliation.*

Proof If Y has a coorientable taut foliation then we can perturb it to a weakly fillable contact structure [9; 15; 23], and use [14] to construct a simply connected weak symplectic filling of (Y, ξ) with $b^+ > 0$, exactly as in the proof of [26, Proposition 15]. This contradicts Proposition 4.3. $\square$

5 Homology of orders 3 and 5

We can now prove the second case of Theorem 1.1, which asserts that if Y has Heegaard genus 2 and first homology of order 3, then it is not $\mathrm{SU}(2)$ -abelian. Here we will need the fact, proved in [5], that the dimension of KHI is enough to detect the trefoils.

Lemma 5.1 *If $K \subset S^3$ is an $SU(2)$ -simple knot of determinant 3, then K is a trefoil.*

Proof Proposition 3.1 says that $\dim_{\mathbb{Q}} KHI(K) = 3$, and by [5, Theorem 1.6] the only knots with three-dimensional instanton knot homology are the trefoils. $\square$

Proof of Theorem 1.1 in the case $|H_1(Y)| = 3$ Suppose that Y has Heegaard genus 2 and is $SU(2)$ -abelian, and that $H_1(Y)$ has order 3. By Proposition 2.1, we can write $Y \cong \Sigma_2(K)$ for an $SU(2)$ -simple 3-bridge knot K of determinant 3. But Lemma 5.1 says that there are no such knots, since the trefoils have bridge index 2, so such Y cannot exist. $\square$

We cannot quite prove the analogue of Lemma 5.1 for knots of determinant 5, but we can achieve a partial characterization. We recall a bit of terminology first: an *instanton L-space knot* is a knot in S^3 on which some positive Dehn surgery produces an instanton L-space. Such knots are known to be fibered and strongly quasipositive [6, Theorem 1.15].

Lemma 5.2 *Let K be an $SU(2)$ -simple knot of determinant 5 other than the figure eight or $T_{\pm 2,5}$. Then, after possibly replacing it with its mirror, K is an instanton L-space knot, hence fibered and strongly quasipositive, with Seifert genus $g = g(K)$ at least 3 and signature $\sigma(K) \in \{-4, 0, 4\}$. Moreover, its branched double cover $Y = \Sigma_2(K)$ bounds a simply connected Stein domain X with*

$$b_2^+(X) = g + \frac{1}{2}\sigma(K) \geq 1.$$

Proof Proposition 3.1 says that $|\sigma(K)| \leq 4$ and $\sigma(K) \equiv \det(K) - 1 \equiv 0 \pmod{4}$, so $\sigma(K)$ must be either 0 or ± 4 ; and that $\dim_{\mathbb{Q}} KHI(K) = 5$. We will use the Alexander decomposition of $KHI(K)$ to understand more about K .

First, we suppose that $\dim_{\mathbb{Q}} KHI(K, g) > 1$. Then since $\dim_{\mathbb{Q}} KHI(K) = 5$ we must have

$$KHI(K; \mathbb{Q}) \cong \mathbb{Q}_g^2 \oplus \mathbb{Q}_0 \oplus \mathbb{Q}_{-g}^2,$$

where the subscripts denote the Alexander grading. But then Theorem 3.3 and the fact that $\frac{1}{2}\sigma(K)$ is even would tell us that

$$\Delta_K(t) = 2(-t)^g + 1 + 2(-t)^{-g},$$

which is impossible since $\Delta_K(1) = 1$. It follows that $\dim_{\mathbb{Q}} KHI(K, g) = 1$ and that K is fibered. Moreover, if K were fibered of genus 1 then it would be the figure eight (the trefoils having determinant 3), but we have excluded this possibility. Therefore K must have genus $g \geq 2$.

Now since K is fibered we can apply [5, Theorem 1.7] to show that $KHI(K, g-1) \neq 0$, and since $KHI(K)$ has total dimension 5, we have

$$KHI(K; \mathbb{Q}) \cong \mathbb{Q}_g \oplus \mathbb{Q}_{g-1} \oplus \mathbb{Q}_0 \oplus \mathbb{Q}_{1-g} \oplus \mathbb{Q}_{-g}.$$

According to work of Li and Liang [31, Theorem 1.4], this tells us that K is an instanton L-space knot, up to mirroring; we replace K with its mirror as needed to guarantee that it has a positive instanton L-space

surgery. Then K is also strongly quasipositive [6, Theorem 1.15], as claimed. Moreover, Farber, Reinoso, and Wang [16, Corollary 1.4] proved (building on a partial characterization in [2, Section 2]) that $T_{2,5}$ is the only genus-2 instanton L-space knot, and we have assumed that K is not $T_{2,5}$, so in fact $g \geq 3$.

Boileau, Boyer, and Gordon [8, Proposition 6.1] proved that since K is strongly quasipositive, if its branched double cover is a Heegaard Floer L-space, then $|\sigma(K)| = 2g(K)$. We adapt part of their argument to the instanton setting: they push a genus- g Seifert surface F for $K \subset S^3$ into the 4-ball so that F is the intersection of a smooth complex curve with that ball, and then observe that $X = \Sigma_2(F)$ is a Stein manifold [17; 39] with boundary $\Sigma_2(K)$. Then X is simply connected [8, Section 3.1.1], and has $b_3(X) = 0$ because it is Stein, and this implies that $b_2(X) = 2g$ by [8, Remark 3.7]. Its signature $\sigma(X)$ is equal to $\sigma(K)$ [22, Lemma 1.1], so we determine that X gives a Stein filling of $\Sigma_2(K)$ satisfying

$$b_2^+(X) = g + \frac{1}{2}\sigma(K), \quad b_2^-(X) = g - \frac{1}{2}\sigma(K),$$

as claimed. Moreover, since $g \geq 3$ and $\sigma(K) \geq -4$ we conclude that $b_2^+(X)$ is positive. $\square$

We can now apply Proposition 2.1 to complete the proof of Theorem 1.1.

Proof of Theorem 1.1 We have already proved the cases where $H_1(Y; \mathbb{Z})$ has order 1 (in Section 2) or 3, so now we suppose that Y has Heegaard genus 2, that $|H_1(Y; \mathbb{Z})| = 5$, and that Y is $\mathrm{SU}(2)$ -abelian. Then Proposition 2.1 says that Y is the branched double cover of an $\mathrm{SU}(2)$ -simple 3-bridge knot K , with $\det(K) = 5$. Since K is not a 2-bridge knot, it is neither the figure eight nor the torus knot $T_{\pm 2,5}$.

We now apply Lemma 5.2, replacing K with its mirror (and hence Y with $-Y$) if necessary to ensure that K is an instanton L-space knot. The conclusion of this lemma says that Y has a simply connected Stein filling (X, J) satisfying $b_2^+(X) \geq 1$. But Lemma 4.2 guarantees that $\pi_1(Y)$ is cyclically finite, so this contradicts Proposition 4.3, and thus such Y cannot exist after all. $\square$

Remark 5.3 Unlike in the cases $\det(K) = 1$ and $\det(K) = 3$, we have not shown that there are no $\mathrm{SU}(2)$ -simple knots of determinant 5 and bridge index greater than 2. What we have shown is that if such a knot K exists, then its branched double cover $Y = \Sigma_2(K)$ cannot be $\mathrm{SU}(2)$ -abelian. We saw in the proof of Proposition 2.1 that this cannot happen if K is a 3-bridge knot, but K could still have bridge index 4 or greater, as per Remark 2.2.

6 Meridian-traceless representation varieties and the trefoils

Given a knot $K \subset S^3$, we define the representation variety

$$\mathcal{R}(K, i) = \{\rho: \pi_1(S^3 \setminus K) \rightarrow \mathrm{SU}(2) \mid \rho(\mu) = i\},$$

where μ is a fixed meridian. This carries an action of

$$U(1)/(\mathbb{Z}/2\mathbb{Z}) = \{e^{i\theta}\}/\{\pm 1\}$$

by conjugation, which fixes the unique reducible representation in $\mathcal{R}(K, i)$ and partitions the irreducible representations into S^1 orbits. In this section, we prove [Theorem 1.3](#), which asserts that there is a single S^1 orbit of irreducible representations if and only if K is a trefoil.

Theorem 6.1 *If $\mathcal{R}(K, i) \cong \{*\} \sqcup S^1$, then K is a trefoil.*

Proof Suppose for now that K is $SU(2)$ -simple, so that every irreducible, meridian-traceless representation $\rho: \pi_1(S^3 \setminus K) \rightarrow SU(2)$ has binary dihedral image. We recall that according to Klassen [\[24, Theorem 10\]](#), there are exactly $\frac{1}{2}(\det(K) - 1)$ such ρ up to conjugacy. But by hypothesis there is a unique conjugacy class of such ρ , so now K is an $SU(2)$ -simple knot with $\det(K) = 3$, and [Lemma 5.1](#) says that K must therefore be a trefoil.

We now show that K is indeed $SU(2)$ -simple. Fixing the nontrivial central character

$$\chi: \pi_1(S^3 \setminus K) \twoheadrightarrow H_1(S^3 \setminus K) \cong \mathbb{Z} \rightarrow \{\pm 1\},$$

we can take any representation in $\mathcal{R}(K, i)$, say

$$\rho: \pi_1(S^3 \setminus K) \rightarrow SU(2),$$

and consider the representation

$$\rho'(\gamma) = j \cdot \chi(\gamma) \rho(\gamma) \cdot j^{-1}.$$

We note that $\rho' \in \mathcal{R}(K, i)$, since $\chi(\mu) = -1$ and $\rho(\mu) = i$ imply that

$$\rho'(\mu) = j \cdot (-i) \cdot j^{-1} = i.$$

The operation $\rho \mapsto \rho'$ defines an involution not just on $\mathcal{R}(K, i)$ but on the entire meridian-traceless $SL_2(\mathbb{C})$ character variety of K , and Nagasato and Yamaguchi [\[37, Proposition 3\]](#) proved that the irreducible characters fixed by this involution are precisely the metabelian ones, which are the same as the binary dihedral characters [\[37, Proposition 2\]](#).

If ρ is conjugate to ρ' then they have the same characters, so the above says that $\text{tr } \rho$ is equal to the character of an irreducible binary dihedral representation ρ_0 , which we can take to have image in $SU(2)$. Facts (1) and (2) in the proof of [\[24, Proposition 15\]](#) say that ρ and ρ_0 are conjugate in $SL_2(\mathbb{C})$, and then that they are conjugate by an element of $SU(2)$. Thus ρ can only be conjugate to ρ' if ρ is itself a binary dihedral representation. But if K is not $SU(2)$ -simple then we can choose some irreducible $\rho \in \mathcal{R}(K, i)$ which is not binary dihedral, and the corresponding $\rho' \in \mathcal{R}(K, i)$ will not be conjugate to ρ . In this case $\mathcal{R}(K, i)$ has at least two distinct S^1 orbits, which contradicts the hypothesis that $\mathcal{R}(K, i) \cong \{*\} \sqcup S^1$. We conclude that K must be $SU(2)$ -simple after all, and hence K is a trefoil as argued above. $\square$

7 Homology of order 2

In this section we will prove [Proposition 7.5](#), which classifies $SU(2)$ -simple links of determinant 2. We will then use this to prove [Theorem 1.5](#), constructing nonabelian $SO(3)$ -representations of arbitrary 3-manifolds Y with Heegaard genus 2 and first homology of order 2.

Lemma 7.1 *If $L \subset S^3$ is a link of $\ell \geq 1$ components, then $\det(L)$ is a multiple of $2^{\ell-1}$. If $\ell = 1$ then $\det(L)$ is odd, while if $\ell = 2$ and we write $L = L_1 \cup L_2$, then*

$$\det(L) \equiv 2\text{lk}(L_1, L_2) \pmod{4}.$$

Proof Let $Y = \Sigma_2(L)$. We will assume $\det(L) \neq 0$ for now, since otherwise $2^{\ell-1}$ divides it anyway. Then $H_1(Y; \mathbb{Z})$ is finite of order $\det(L)$, and according to the universal coefficient theorem and (2-2),

$$H_1(Y; \mathbb{Z}) \otimes (\mathbb{Z}/2\mathbb{Z}) \cong H_1(Y; \mathbb{Z}/2\mathbb{Z}) \cong (\mathbb{Z}/2\mathbb{Z})^{\ell-1}.$$

Writing $H_1(Y; \mathbb{Z}) \cong \bigoplus_i \mathbb{Z}/d_i\mathbb{Z}$ as a sum of cyclic groups, so that $\det(L) = \prod_i d_i$, each $\mathbb{Z}/d_i\mathbb{Z}$ summand contributes a $\mathbb{Z}/2\mathbb{Z}$ summand to the tensor product on the left if d_i is even, and the trivial group otherwise. Thus there must be exactly $\ell - 1$ even values of d_i , and it follows that $2^{\ell-1}$ divides $\prod_i d_i = \det(L)$. If $\ell = 1$ then all of the d_i are odd, so their product $\det(L)$ is odd as well.

Now suppose that $\ell = 2$ and that $\det(L)$ is arbitrary. In this case it equals

$$|H_1(Y; \mathbb{Z})| = 2\Delta_L(-1, -1),$$

where Δ_L is the multivariable Alexander polynomial of $L = L_1 \cup L_2$, as shown by Hosokawa and Kinoshita [21]. On the other hand, Torres [44] proved that the linking number $\text{lk}(L_1, L_2)$ satisfies

$$\Delta_L(1, 1) = \pm \text{lk}(L_1, L_2),$$

so in particular

$$\Delta_L(-1, -1) \equiv \Delta_L(1, 1) \equiv \text{lk}(L_1, L_2) \pmod{2}.$$

Thus $\det(L) \equiv 2\text{lk}(L_1, L_2) \pmod{4}$ as claimed. □

In what follows we will let

$$D = \{e^{i\theta}\} \cup \{e^{i\theta}j\} \subset \text{SU}(2)$$

denote the binary dihedral subgroup of $\text{SU}(2)$.

Lemma 7.2 *Let $L \subset S^3$ be a link with $\det(L) \equiv 2 \pmod{4}$, and fix a meridian-traceless representation*

$$\rho: \pi_1(S^3 \setminus L) \rightarrow \text{SU}(2)$$

with nonabelian image in the binary dihedral group D . Then ρ sends every meridian of L to an element of the form $e^{i\theta}j$.

Proof Lemma 7.1 tells us that L must have exactly two components, since $\det(L)$ is neither odd nor a multiple of 4, and that if we write $L = L_1 \cup L_2$ then $\text{lk}(L_1, L_2)$ is odd. We will consider the Wirtinger presentation associated to some diagram of L , which is generated by the meridians around each strand. From this presentation we know that all of the meridians of L_1 are conjugate to each other, and likewise the meridians of L_2 are all mutually conjugate.

Both the normal subgroup $N = \{e^{i\theta}\}$ of D and its coset $Nj = \{e^{i\theta}j\}$ are fixed by conjugation in D , so either ρ sends all the meridians of L_1 to elements of N , or it sends them all to elements of Nj . The same is true for L_2 . If ρ sends both sets of meridians to Nj then we are done; if instead both sets are sent to N then all of $\text{Im}(\rho)$ lies in N , contradicting our assumption that $\text{Im}(\rho)$ is nonabelian. Thus we will assume without loss of generality that the meridians of L_1 are sent to $N = \{e^{i\theta}\}$, while the meridians of L_2 are sent to $Nj = \{e^{i\theta}j\}$.

We now walk along L_1 in a circle and observe the sequence of values of ρ at each strand. Each of these values is either i or $-i$, since these are the only purely imaginary elements of N . If two such strands μ and μ' of L_1 meet at a crossing c , where the overcrossing strand has meridian μ_c , then the Wirtinger presentation says that

$$\rho(\mu') = \rho(\mu_c^{\pm 1}) \cdot \rho(\mu) \cdot \rho(\mu_c^{\mp 1}).$$

If μ_c belongs to L_1 then its image lies in N , hence commutes with $\rho(\mu) = \pm i$, and so $\rho(\mu') = \rho(\mu)$. On the other hand, if μ_c belongs to L_2 then we can write $\rho(\mu_c^{\pm 1}) = e^{i\theta}j$ for some θ , and we have

$$e^{i\theta}j \cdot (\pm i) \cdot (e^{i\theta}j)^{-1} = e^{i\theta}j \cdot (\pm i) \cdot (-je^{-i\theta}) = \mp i,$$

so then $\rho(\mu') = -\rho(\mu)$. In particular, as we travel along L_1 , the value of ρ at each meridian changes sign every time we cross under L_2 , and it stays the same otherwise.

In order to assign $\pm i$ to each strand of L_1 in a consistent way, it follows that L_1 must cross under L_2 an even number of times. Changing one of these undercrossings to an overcrossing changes $\text{lk}(L_1, L_2)$ by ± 1 ; if instead we change all of them, then the total change must be even. But doing this causes L_1 and L_2 to be unlinked from each other, since now L_1 always crosses over L_2 . Therefore $\text{lk}(L_1, L_2)$ differs from 0 by an even number. In other words, the linking number $\text{lk}(L_1, L_2)$ is even, but we saw at the beginning of the proof that it must also be odd and therefore we have a contradiction. $\square$

Lemma 7.3 *Let $L \subset S^3$ be a link of at least two components, with $\det(L) \neq 0$, and let*

$$\rho: \pi_1(S^3 \setminus L) \rightarrow D \subset SU(2)$$

be a meridian-traceless representation with nonabelian image. Suppose that ρ sends every meridian of L to an element of the form $e^{i\theta}j$. Then $\text{Im}(\rho)$ is a finite group of order dividing $2\det(L)$.

Proof We argue exactly as in the proof of [24, Theorem 10]. Suppose that we fix a diagram of L , and let $\mu_1, \dots, \mu_s$ denote the meridional generators of the corresponding Wirtinger presentation. If we have a meridian-traceless, nonabelian representation

$$\rho: \pi_1(S^3 \setminus L) \rightarrow D \subset SU(2)$$

such that $\rho(\mu_r) = e^{i\theta_r}j$ for each $r = 1, \dots, s$, then the relations in the Wirtinger presentation are equivalent to a system of linear equations, one for each of $r - 1$ crossings of the diagram, which at a given crossing take the form

$$\theta_a - 2\theta_b + \theta_c \equiv 0 \pmod{2\pi}$$

if the strand corresponding to μ_b crosses over the strands for μ_a and μ_c . We can conjugate ρ by $e^{-i\theta_s/2} \in D$ to replace a solution $(\theta_1, \dots, \theta_{s-1}, \theta_s)$ with $(\theta_1 - \theta_s, \dots, \theta_{s-1} - \theta_s, 0)$, using the relation

$$e^{-i\theta_s/2} \cdot e^{i\alpha} j \cdot (e^{-i\theta_s/2})^{-1} = e^{i(\alpha - \theta_s)} j.$$

Thus up to conjugacy we are free to take $\theta_s = 0$, and then nonabelian representations will correspond to solutions $(\theta_1, \dots, \theta_{s-1})$ that are not identically zero modulo π .

The $(s-1) \times (s-1)$ integer matrix A describing this system is obtained from an Alexander matrix by setting $t = -1$, so it has determinant

$$\det(A) = \pm \Delta_L(-1) = \pm \det(L).$$

In particular, if we write $\Delta = \det(L)$ for readability then any nonzero solution satisfies

$$(\theta_1, \dots, \theta_{s-1}) \in \frac{1}{\det(A)} (2\pi\mathbb{Z})^{s-1} = \frac{2\pi}{\Delta} \mathbb{Z}^{s-1},$$

and in this case the image $\rho(\mu_r)$ of each meridian lies in the set

$$(7-1) \quad \{e^{2\pi im/\Delta} \mid 0 \leq m < \Delta\} \cup \{e^{2\pi im/\Delta} \cdot j \mid 0 \leq m < \Delta\}.$$

Since $\Delta = \det(L)$ is even, this set is closed under multiplication; hence it is a binary dihedral subgroup of D , with order $2\det(L)$. Then $\text{Im}(\rho)$ is a subgroup of this, so its order divides $2\det(L)$, as promised. $\square$

Remark 7.4 If L were a knot, then the set (7-1) would not be closed under multiplication, because it does not contain -1 if Δ is odd. But adjoining -1 turns it into a subgroup, so in this case we would conclude that $\text{Im}(\rho)$ has order dividing $4\det(L)$ instead.

Proposition 7.5 *Let $L \subset S^3$ be an $\text{SU}(2)$ -simple link of determinant 2. Then L is a Hopf link.*

Proof Suppose that L is not a Hopf link. Xie and Zhang [45, Theorem 1.5] proved that the following are equivalent:

- L is not an unknot, a Hopf link, or a connected sum of Hopf links.
- There is an irreducible representation

$$\rho: \pi_1(S^3 \setminus L) \rightarrow \text{SU}(2)$$

sending every meridian of L to a traceless element.

Now we know by Lemma 7.1 that L must be a two-component link, and we have assumed that it is not a Hopf link, so such a representation ρ must exist, with nonabelian image.

Suppose in addition that after possibly replacing ρ with a conjugate, the image $\text{Im}(\rho)$ lies in the binary dihedral group D . Lemma 7.2 says that ρ sends every meridian of L to an element of the form $e^{i\theta} j$, so by Lemma 7.3 we know that $\text{Im}(\rho)$ is a finite group of order dividing $2\det(L) = 4$. But any such group must be abelian, so we have a contradiction. Thus ρ cannot be conjugate to a binary dihedral representation, and this proves that L is not $\text{SU}(2)$ -simple. $\square$

We are now ready to prove [Theorem 1.5](#), asserting that if Y has Heegaard genus 2 and $H_1(Y)$ has order 2 then there is a nonabelian representation $\pi_1(Y) \rightarrow SO(3)$. To do so, we recall how to construct a map

$$(7-2) \quad B: \left\{ \begin{array}{l} \text{meridian-traceless representations} \\ \rho: \pi_1(S^3 \setminus L) \rightarrow SU(2) \end{array} \right\} \rightarrow \text{Hom}(\pi_1(\Sigma_2(L)), SO(3)).$$

We have a short exact sequence

$$1 \rightarrow \pi_1(X_2) \rightarrow \pi_1(S^3 \setminus L) \rightarrow \mathbb{Z}/2\mathbb{Z} \rightarrow 1,$$

where the map to $\mathbb{Z}/2\mathbb{Z}$ sends the meridian μ_i of each component L_i (say $i = 1, \dots, r$) to 1, and where X_2 is the corresponding double cover of the complement of L . Then we obtain $Y \cong \Sigma_2(L)$ by Dehn filling X_2 along each lift of μ_i^2 , so that

$$\pi_1(Y) \cong \frac{\pi_1(X_2)}{\langle\langle \mu_1^2, \mu_2^2, \dots, \mu_r^2 \rangle\rangle}.$$

Now given a meridian-traceless ρ , we know that $\rho(\mu_i^2) = -1$ for each i , so

$$\text{ad } \rho: \pi_1(S^3 \setminus L) \rightarrow SO(3)$$

sends each μ_i^2 to the identity; thus its restriction to $\pi_1(X_2)$ induces a well-defined

$$\tilde{\rho}: \pi_1(Y) \rightarrow SO(3).$$

The map (7-2) is defined by $B(\rho) = \tilde{\rho}$.

The following is part of the proof of [\[46, Proposition 3.1\]](#).

Lemma 7.6 [\[46\]](#) Suppose that $\det(L) \neq 0$, and let $\rho: \pi_1(S^3 \setminus L) \rightarrow SU(2)$ be a meridian-traceless representation with nonabelian image. If $B(\rho)$ has abelian image, then ρ is conjugate to a representation with image in the binary dihedral subgroup

$$D = \{e^{i\theta}\} \cup \{e^{i\theta}j\} \subset SU(2).$$

The condition $\det(L) \neq 0$ is needed to ensure that $\text{Im}(B(\rho))$ is a *finite* subgroup of $SO(3)$.

Proof of Theorem 1.5 Just as in the proof of [Proposition 2.1](#), we can appeal to Birman and Hilden [\[7, Theorem 1\]](#) to write

$$Y \cong \Sigma_2(L)$$

for some 3-bridge link L in S^3 , with determinant $\det(L) = |H_1(Y; \mathbb{Z})| = 2$ and hence (by [Lemma 7.1](#)) with two components. If L were the Hopf link then $Y \cong \Sigma_2(L)$ would be $\mathbb{RP}^3$, whose Heegaard genus is only 1, so now [Proposition 7.5](#) says that L is not $SU(2)$ -simple. In particular, there is a meridian-traceless, nonabelian representation

$$\rho: \pi_1(S^3 \setminus L) \rightarrow SU(2)$$

whose image is not conjugate to a subgroup of the binary dihedral group $D \subset SU(2)$.

We now use ρ to construct a representation $\tilde{\rho} = B(\rho): \pi_1(Y) \rightarrow \mathrm{SO}(3)$ as in (7-2). Since ρ does not have binary dihedral image, Lemma 7.6 says that $\tilde{\rho}$ cannot have abelian image and so this proves that Y is not $\mathrm{SO}(3)$ -abelian. $\square$

References

- [1] **D M Austin, P J Braam**, *Equivariant Floer theory and gluing Donaldson polynomials*, Topology 35 (1996) 167–200 [MR](#) [Zbl](#)
- [2] **J A Baldwin, Z Li, S Sivek, F Ye**, *Small Dehn surgery and $\mathrm{SU}(2)$* , Geom. Topol. 28 (2024) 1891–1922 [MR](#) [Zbl](#)
- [3] **J A Baldwin, S Sivek**, *Stein fillings and $\mathrm{SU}(2)$ representations*, Geom. Topol. 22 (2018) 4307–4380 [MR](#) [Zbl](#)
- [4] **J A Baldwin, S Sivek**, *Instanton L -spaces and splicing*, Ann. H. Lebesgue 5 (2022) 1213–1233 [MR](#) [Zbl](#)
- [5] **J A Baldwin, S Sivek**, *Khovanov homology detects the trefoils*, Duke Math. J. 171 (2022) 885–956 [MR](#) [Zbl](#)
- [6] **J A Baldwin, S Sivek**, *Instantons and L -space surgeries*, J. Eur. Math. Soc. 25 (2023) 4033–4122 [MR](#) [Zbl](#)
- [7] **J S Birman, H M Hilden**, *Heegaard splittings of branched coverings of S^3* , Trans. Amer. Math. Soc. 213 (1975) 315–352 [MR](#) [Zbl](#)
- [8] **M Boileau, S Boyer, C M Gordon**, *Branched covers of quasi-positive links and L -spaces*, J. Topol. 12 (2019) 536–576 [MR](#) [Zbl](#)
- [9] **J Bowden**, *Approximating C^0 -foliations by contact structures*, Geom. Funct. Anal. 26 (2016) 1255–1296 [MR](#) [Zbl](#)
- [10] **S Boyer, A Nicas**, *Varieties of group representations and Casson’s invariant for rational homology 3-spheres*, Trans. Amer. Math. Soc. 322 (1990) 507–522 [MR](#) [Zbl](#)
- [11] **C R Cornwell**, *Character varieties of knot complements and branched double-covers via the cord ring*, preprint (2015) [arXiv 1509.04962](#)
- [12] **S K Donaldson**, *Polynomial invariants for smooth four-manifolds*, Topology 29 (1990) 257–315 [MR](#) [Zbl](#)
- [13] **S K Donaldson**, *Floer homology groups in Yang–Mills theory*, Cambridge Tracts in Math. 147, Cambridge Univ. Press (2002) [MR](#) [Zbl](#)
- [14] **Y Eliashberg**, *A few remarks about symplectic filling*, Geom. Topol. 8 (2004) 277–293 [MR](#) [Zbl](#)
- [15] **Y M Eliashberg, W P Thurston**, *Confoliations*, Univ. Lect. Ser. 13, Amer. Math. Soc., Providence, RI (1998) [MR](#) [Zbl](#)
- [16] **E Farber, B Reinoso, L Wang**, *Fixed-point-free pseudo-Anosov homeomorphisms, knot Floer homology and the cinquefoil*, Geom. Topol. 28 (2024) 4337–4381 [MR](#) [Zbl](#)
- [17] **S Harvey, K Kawamuro, O Plamenevskaya**, *On transverse knots and branched covers*, Int. Math. Res. Not. 2009 (2009) 512–546 [MR](#) [Zbl](#)
- [18] **C M Herald**, *Flat connections, the Alexander invariant, and Casson’s invariant*, Comm. Anal. Geom. 5 (1997) 93–120 [MR](#) [Zbl](#)

- [19] **M Heusener, E Klassen**, *Deformations of dihedral representations*, Proc. Amer. Math. Soc. 125 (1997) 3039–3047 [MR](#) [Zbl](#)
- [20] **C Hodgson, J H Rubinstein**, *Involutions and isotopies of lens spaces*, from “Knot theory and manifolds”, Lecture Notes in Math. 1144, Springer (1985) 60–96 [MR](#) [Zbl](#)
- [21] **F Hosokawa, S Kinoshita**, *On the homology group of branched cyclic covering spaces of links*, Osaka Math. J. 12 (1960) 331–355 [MR](#) [Zbl](#)
- [22] **L H Kauffman, L R Taylor**, *Signature of links*, Trans. Amer. Math. Soc. 216 (1976) 351–365 [MR](#) [Zbl](#)
- [23] **W H Kazez, R Roberts**, *C^0 approximations of foliations*, Geom. Topol. 21 (2017) 3601–3657 [MR](#) [Zbl](#)
- [24] **E P Klassen**, *Representations of knot groups in $SU(2)$* , Trans. Amer. Math. Soc. 326 (1991) 795–828 [MR](#) [Zbl](#)
- [25] **P B Kronheimer, T S Mrowka**, *Dehn surgery, the fundamental group and $SU(2)$* , Math. Res. Lett. 11 (2004) 741–754 [MR](#) [Zbl](#)
- [26] **P B Kronheimer, T S Mrowka**, *Witten’s conjecture and property P*, Geom. Topol. 8 (2004) 295–310 [MR](#) [Zbl](#)
- [27] **P Kronheimer, T Mrowka**, *Instanton Floer homology and the Alexander polynomial*, Algebr. Geom. Topol. 10 (2010) 1715–1738 [MR](#) [Zbl](#)
- [28] **P Kronheimer, T Mrowka**, *Knots, sutures, and excision*, J. Differential Geom. 84 (2010) 301–364 [MR](#) [Zbl](#)
- [29] **P B Kronheimer, T S Mrowka**, *Khovanov homology is an unknot-detector*, Publ. Math. Inst. Hautes Études Sci. 113 (2011) 97–208 [MR](#) [Zbl](#)
- [30] **P Kronheimer, T Mrowka, P Ozsváth, Z Szabó**, *Monopoles and lens space surgeries*, Ann. of Math. 165 (2007) 457–546 [MR](#) [Zbl](#)
- [31] **Z Li, Y Liang**, *Some computations on instanton knot homology*, J. Knot Theory Ramifications 32 (2023) art. id. 2350007 [MR](#) [Zbl](#)
- [32] **T Lidman, J Pinzón-Caicedo, R Zentner**, *Toroidal integer homology three-spheres have irreducible $SU(2)$ -representations*, J. Topol. 16 (2023) 344–367 [MR](#)
- [33] **Y Lim**, *Instanton homology and the Alexander polynomial*, Proc. Amer. Math. Soc. 138 (2010) 3759–3768 [MR](#) [Zbl](#)
- [34] **X-S Lin**, *A knot invariant via representation spaces*, J. Differential Geom. 35 (1992) 337–357 [MR](#) [Zbl](#)
- [35] **K Motegi**, *Haken manifolds and representations of their fundamental groups in $SL(2, \mathbb{C})$* , Topology Appl. 29 (1988) 207–212 [MR](#) [Zbl](#)
- [36] **K Murasugi**, *On a certain numerical invariant of link types*, Trans. Amer. Math. Soc. 117 (1965) 387–422 [MR](#) [Zbl](#)
- [37] **F Nagasato, Y Yamaguchi**, *On the geometry of the slice of trace-free $SL_2(\mathbb{C})$ -characters of a knot group*, Math. Ann. 354 (2012) 967–1002 [MR](#) [Zbl](#)
- [38] **P Ozsváth, Z Szabó**, *Holomorphic disks and genus bounds*, Geom. Topol. 8 (2004) 311–334 [MR](#) [Zbl](#)
- [39] **L Rudolph**, *Some 3-dimensional transverse $\mathbb{C}$ -links (constructions of higher-dimensional $\mathbb{C}$ -links, I)*, from “Interactions between low-dimensional topology and mapping class groups”, Geom. Topol. Monogr. 19, Geom. Topol. Publ., Coventry (2015) 367–413 [MR](#) [Zbl](#)

- [40] **H Schubert**, *Über eine numerische Knoteninvariante*, Math. Z. 61 (1954) 245–288 [MR](#) [Zbl](#)
- [41] **S Sivek**, *Donaldson invariants of symplectic manifolds*, Int. Math. Res. Not. 2015 (2015) 1688–1716 [MR](#) [Zbl](#)
- [42] **S Sivek, R Zentner**, *A menagerie of $SU(2)$ -cyclic 3-manifolds*, Int. Math. Res. Not. 2022 (2022) 8038–8085 [MR](#) [Zbl](#)
- [43] **W P Thurston**, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. 6 (1982) 357–381 [MR](#) [Zbl](#)
- [44] **G Torres**, *On the Alexander polynomial*, Ann. of Math. 57 (1953) 57–89 [MR](#) [Zbl](#)
- [45] **Y Xie, B Zhang**, *On meridian-traceless $SU(2)$ -representations of link groups*, Adv. Math. 418 (2023) art. id. 108947 [MR](#) [Zbl](#)
- [46] **R Zentner**, *A class of knots with simple $SU(2)$ -representations*, Selecta Math. 23 (2017) 2219–2242 [MR](#) [Zbl](#)

Department of Mathematics, Boston College
Chestnut Hill, MA, United States

Department of Mathematics, Imperial College London
London, United Kingdom

john.baldwin@bc.edu, s.sivek@imperial.ac.uk

<https://sites.google.com/bc.edu/john-baldwin/>, <https://www.ma.imperial.ac.uk/~ssivek/>

Received: 20 September 2023 Revised: 29 February 2024

Harmonic measures and rigidity for surface group actions on the circle

MASANORI ADACHI
YOSHIFUMI MATSUDA
HIRAKU NOZAWA

We study rigidity properties of actions of a torsion-free lattice of $\mathrm{PSU}(1, 1)$ on the circle S^1 . We follow the approaches of Frankel and Thurston proposed in preprints via foliated harmonic measures on the suspension bundles. Our main results are a curvature estimate and a Gauss–Bonnet formula for the S^1 connection obtained by taking the average of the flat connection with respect to a harmonic measure. As consequences, we give a precise description of the harmonic measure on suspension foliations with maximal Euler number and an alternative proof of rigidity theorems of Matsumoto and Burger, Iozzi and Wienhard.

22D50, 53C12; 31A15, 57M60, 57R20, 58D19

1 Introduction

Let Σ be an orientable hyperbolic surface of finite type. The $\pi_1(\Sigma)$ -actions on the circle S^1 have rigidity properties governed by the Euler class. For closed Σ , the Euler class is bounded by the Milnor–Wood inequality [20; 25] and Matsumoto [17] proved that if the equality holds for a $\pi_1(\Sigma)$ -action in the Milnor–Wood inequality, then the action is semiconjugate to a Fuchsian action. Burger, Iozzi and Wienhard [7, Section 4.5] generalized the Milnor–Wood inequality and Matsumoto’s theorem to the case where Σ has finite volume and cusps. They defined and used the following generalization of the Euler number: Let $\mathbb{D}$ be the Poincaré disk. For a torsion-free lattice Γ in $\mathrm{PSU}(1, 1) = \mathrm{Isom}_+(\mathbb{D})$ and a homomorphism $\rho: \Gamma \rightarrow \mathrm{Homeo}_+(S^1)$, the Euler number $e(\rho) \in \mathbb{R}$ is defined based on the bounded cohomology. In the case where Γ is not uniform, they showed that it is expressed in terms of the translation number of any homomorphism lifting $\tilde{\rho}: \Gamma \rightarrow \widetilde{\mathrm{Homeo}}_+(S^1)$ of ρ to the universal cover group $\widetilde{\mathrm{Homeo}}_+(S^1)$ of $\mathrm{Homeo}_+(S^1)$: we have

$$e(\rho) = - \sum_{i=1}^m \tau(\tilde{\rho}(c_i)),$$

where $\tau: \widetilde{\mathrm{Homeo}}_+(S^1) \rightarrow \mathbb{R}$ is the translation number and $c_1, \dots, c_m$ are curves which go around each of cusps of the surface $\Sigma := \Gamma \backslash \mathbb{D}$.

In this article, we study the rigidity of surface group actions on S^1 via foliated harmonic measures. We are going to follow an approach proposed in the unpublished article by Frankel [12] and the unfinished paper by Thurston [23] (cf Calegari [8, Example 4.6]). During this course, we will address some

missing details in these previous works concerning regularity issues (see Lemmas 3.4 and 3.8). We will consider a specifically chosen S^1 -connection obtained by taking the average of the flat connection on the suspension bundle with respect to a harmonic measure. Our main results concern this connection: a curvature estimate, which we prove following the master thesis of the first author [1], and a Gauss–Bonnet formula (Theorem 1.1). These results led us to give a precise description of the harmonic measure with maximal Euler number (Theorem 1.3) in addition to an alternative proof of the Matsumoto and Burger–Iozzi–Wienhard results (Corollary 1.2).

To state our main result, we briefly explain a construction due to Thurston [23] of the specifically chosen S^1 -connection on the suspension bundle of a given action $\rho: \Gamma \rightarrow \text{Homeo}_+(S^1)$. Consider the suspension bundle $\pi: \Sigma \times_\rho S^1 \rightarrow \Sigma$ of ρ equipped with the horizontal foliation $\mathcal{F}$. Let μ be a harmonic measure on $\mathcal{F}$, which exists from the result of Alvarez [4]. We normalize and disintegrate μ along the fibers of π to obtain a family of measures $\{\mu_z\}_{z \in \Sigma}$ on the fibers with total measure 2π . When $\rho(\Gamma)$ has no finite orbit in S^1 , measures μ_z are nonatomic. Moreover, we may assume that the μ_z have full support on the fibers by collapsing the complement of the support of μ by a semiconjugacy. We may regard $\Sigma \times_\rho S^1$ as a principal S^1 -bundle whose principal S^1 -action $\{\phi_t\}$ preserves μ_z on each fiber. Let ω be the flat connection form defining $\mathcal{F}$. By taking the average of ω under the principal S^1 -action, we obtain an S^1 -connection of $\Sigma \times_\rho S^1$,

$$\bar{\omega} := \frac{1}{2\pi} \int_0^{2\pi} \phi_t^* \omega \, dt,$$

which we call the S^1 -connection form associated with μ . It should be noted that, after collapsing the complement of the support of μ , the foliation $\mathcal{F}$ is only transversely Lipschitz and ω is continuous (see Step 2 of Section 3). Then $\bar{\omega}$ is only continuous in general and its curvature is considered as in the sense of Definition 2.4.

Now we state our main result.

Theorem 1.1 *Let Γ be a torsion-free lattice in $\text{PSU}(1, 1)$ and $\rho: \Gamma \rightarrow \text{Homeo}_+(S^1)$ a homomorphism. Assume that $\rho(\Gamma)$ has no finite orbit in S^1 . For a harmonic measure μ on the suspension foliation of ρ , let $\bar{\omega}$ be the S^1 -connection form associated with μ . Then, $\bar{\omega}$ has curvature of the form $K \, \text{vol}$ such that $|K(z)| \leq 1$ a.e. z , and we have*

$$e(\rho) = \frac{1}{2\pi} \int_\Sigma K(z) \, \text{vol}(dz),$$

where vol denotes the hyperbolic volume form on $\Sigma := \Gamma \backslash \mathbb{D}$.

In the case where Γ is a uniform lattice, the Gauss–Bonnet formula is classical and is not new. The main point of this result is the estimate of the curvature, already stated in [23], and the behavior of the connection $\bar{\omega}$ near the cusps of Σ , which is proved based on Harnack’s inequality and the isoperimetric inequality. We remark that there is a similar curvature estimate in [5, §6] by Besson, Courtois and Gallot. As a corollary, we obtain an alternative proof of the generalization of the Milnor–Wood inequality and Matsumoto’s rigidity theorem to torsion-free lattices due to Burger, Iozzi and Wienhard [7].

Corollary 1.2 For a torsion-free lattice Γ in $\mathrm{PSU}(1, 1)$ and a homomorphism $\rho: \Gamma \rightarrow \mathrm{Homeo}_+(S^1)$, we have

$$(1-1) \quad |e(\rho)| \leq -e(\Gamma \setminus \mathbb{D}),$$

where $e(\Gamma \setminus \mathbb{D})$ denotes the Euler characteristic of the hyperbolic surface $\Gamma \setminus \mathbb{D}$. Furthermore, we have $e(\rho) = e(\Gamma \setminus \mathbb{D})$ (resp. $e(\rho) = -e(\Gamma \setminus \mathbb{D})$) if and only if ρ is semiconjugate to ρ_0 (resp. $\bar{\rho}_0$), where ρ_0 is the Fuchsian action of Γ induced from the inclusion $\Gamma \rightarrow \mathrm{PSU}(1, 1) \rightarrow \mathrm{Homeo}_+(S^1)$, and $\bar{\rho}_0$ is its complex conjugate.

In the case where Γ is uniform, $\Gamma \setminus \mathbb{D}$ is a closed surface. In that case, Corollary 1.2 is nothing but the Milnor–Wood inequality [20; 25] and the rigidity theorem of Matsumoto [17]. Note that Matsumoto [19] recently gave an alternative concise proof of the theorem of Burger, Iozzi and Wienhard [7] by using his rigidity theorem for closed surfaces.

The key of the proof of the rigidity part of Corollary 1.2 is to show a rigidity of harmonic measures in the equality case in the Milnor–Wood inequality (1-1). It is well known that the pull back of a harmonic measure on the suspension bundle to $\mathbb{D} \times \mathbb{R}$ is of the form

$$(1-2) \quad h(z, t) \mathrm{vol}(z) \nu(t) \quad (z \in \mathbb{D}, t \in \mathbb{R}),$$

where vol is the leafwise volume form, ν is a Borel measure on $\mathbb{R}$ and h is a positive leafwise harmonic function. Note that h is determined only up to the multiplication of leafwise constant functions. We will show that, if we have the equality in (1-1), then we can take h closely related to the Poisson kernel as stated in the following result:

Theorem 1.3 Let Γ be a torsion-free lattice in $\mathrm{PSU}(1, 1)$ and $\rho: \Gamma \rightarrow \mathrm{Homeo}_+(S^1)$ a homomorphism. If we have $e(\rho) = e(\Sigma)$, then the pull back of every harmonic measure on the suspension bundle of ρ to the universal cover $\mathbb{D} \times \mathbb{R}$ is of the form

$$\frac{1 - |z|^2}{|\mathfrak{m}(e^{it}) - z|^2} \mathrm{vol}(z) \nu(t) \quad (z \in \mathbb{D}, t \in \mathbb{R}),$$

where vol is the leafwise hyperbolic volume form, for a Borel measure ν on $\mathbb{R}$ and a continuous monotone mapping $\mathfrak{m}: S^1 \rightarrow S^1$ of degree one which is (ρ, ρ_0) -equivariant, namely, $\mathfrak{m} \circ \rho(\gamma) = \rho_0(\gamma) \circ \mathfrak{m}$ holds for all $\gamma \in \Gamma$.

The map $\mathfrak{m}$ gives a semiconjugacy from the $\rho(\Gamma)$ -action to the Fuchsian Γ -action. Matsumoto considered this map in [18]. Consider a Γ -action τ on a compact metric space F . Fix an ergodic harmonic measure on the suspension bundle. Then the lift to $\mathbb{D} \times F$ is of the form as in (1-2), where $\mathbb{R}$ is replaced with F . It is well known that, for each $t \in F$, the function $h(\cdot, t)$ is the integral of the Poisson kernel with respect to a Borel measure σ_t on the ideal boundary of the leaf $\mathbb{D} \times \{t\}$. Matsumoto proved the following dichotomy: We have either the support of σ_t is a point or the entire ideal boundary of the leaf for generic $t \in F$ simultaneously. In the former case, we have a (τ, ρ_0) -equivariant measurable map $\mathfrak{m}: F \rightarrow S^1$; $t \mapsto \mathrm{supp}(\sigma_t)$, which we call the *Matsumoto map*. The map $\mathfrak{m}$ in Theorem 1.3 is essentially this Matsumoto map.

Convention

Integrations and “almost everywhere” refer to the Lebesgue measure unless otherwise stated. The Dirac measure at a point a is denoted by δ_a . We use the identification $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ throughout this paper.

Acknowledgements

We are grateful to Masayuki Asaoka for his comments, in particular, pointing out a serious gap in the proof of (3-12) in a previous version of this article. We also thank Ken’ichi Yoshida for letting us know about the publication of Thurston’s article [23]. We thank the referee for careful reading and suggestions that highly improved the readability of this article.

Adachi is supported by JSPS KAKENHI grant numbers JP18K13422, JP19KK0347, JP21H00980, JP21K18579 and JP24K06776. Matsuda is supported by JSPS KAKENHI grant number JP17K05260. Nozawa is supported by JSPS KAKENHI grant numbers JP20K03620 and JP24K06723.

2 Preliminaries

2.1 Harmonic measures on the suspension foliation

A harmonic measure on a foliated manifold is a measure on M invariant under the leafwise heat flow, which was introduced by Garnett [13]. The advantage is that a nontrivial harmonic probability measure always exists for foliations on compact manifolds by Garnett’s theorem, while transverse invariant measures may not exist. There are many applications to foliations, group actions and related topics, for instance, [2; 10; 26].

Let Γ be a torsion-free lattice of $\mathrm{PSU}(1, 1)$, and $\Sigma := \Gamma \backslash \mathbb{D}$. For a homomorphism $\rho: \Gamma \rightarrow \mathrm{Homeo}_+(S^1)$, we can construct its suspension $M := \Sigma \times_\rho S^1 := \Gamma \backslash (\mathbb{D} \times S^1) \rightarrow \Sigma$, where $\gamma \cdot (z, t) = (\gamma z, \rho(\gamma)t)$ for $\gamma \in \Gamma$, $z \in \mathbb{D}$, $t \in S^1$. The suspension foliation $\mathcal{F}$ on M is induced from the product foliation $\mathbb{D} \times S^1 = \bigsqcup_{w \in S^1} \mathbb{D} \times \{w\}$. Since Γ preserves the Poincaré metric,

$$(2-1) \quad (g_{\mathrm{hyp}})_z = \frac{4dz d\bar{z}}{(1 - |z|^2)^2} \quad (z \in \mathbb{D}),$$

the leaves of $\mathcal{F}$ admit the natural Riemannian metric of constant curvature -1 . We call it the leafwise hyperbolic metric. Let us denote the leafwise Laplacian associated with the leafwise hyperbolic metric by Δ .

Definition 2.1 A Borel measure μ on M is called *harmonic* if

$$\int_M \Delta f(x) \mu(dx) = 0$$

for every compactly supported leafwise C^2 function f on M such that Δf is continuous on M .

If M is compact, then a harmonic probability measure on M exists by Garnett's theorem (see [9; 13]). Now M is not compact unless Γ is uniform. In this case, we need the following result of Alvarez [4]:

Theorem 2.2 *There exists a harmonic probability measure μ on $(M, \mathcal{F})$.*

Alvarez [4] constructed a harmonic probability measure by using a stationary measure for the Γ -action on S^1 . By the disintegration formula [13], the pull back of a harmonic measure μ to $\mathbb{D} \times \mathbb{R}$ is of the form

$$(2-2) \quad \tilde{\mu} = h(z, t) \operatorname{vol}(z) \nu(t),$$

where $h(z, t)$ is a Borel measurable function whose restriction to $\mathbb{D} \times \{t\}$ is a positive harmonic function for ν -a.e. t , vol is the leafwise volume form and ν is a Borel measure on $\mathbb{R}$.

2.2 Harnack's inequality

We equip the unit disk $\mathbb{D} = \{z = x_1 + i x_2 \in \mathbb{C} \mid |z| < 1\}$ with the Poincaré metric and denote by d_{hyp} the induced hyperbolic distance, and by $|\cdot|_{\text{hyp}}$ the induced norm on differential forms. The following is a fundamental inequality for positive harmonic functions on $\mathbb{D}$, which we apply to leafwise harmonic functions associated to harmonic measures.

Lemma 2.3 (Harnack's inequality) *Let h be a positive harmonic function on $\mathbb{D}$.*

- (1) *For any $z, w \in \mathbb{D}$, $e^{-d_{\text{hyp}}(z, w)} h(w) \leq h(z) \leq e^{d_{\text{hyp}}(z, w)} h(w)$.*
- (2) *$|d \log h|_{\text{hyp}} \leq 1$ on $\mathbb{D}$.*
- (3) *If $|d \log h|_{\text{hyp}} = 1$ at $\alpha \in \mathbb{D}$, then we have*

$$h(z) = h(0) \frac{1 - |z|^2}{|m - z|^2},$$

where

$$c := \frac{h_{x_1}(\alpha) + i h_{x_2}(\alpha)}{|h_{x_1}(\alpha) + i h_{x_2}(\alpha)|} \in S^1, \quad m := \frac{c + \alpha}{1 + \bar{\alpha}c} \in S^1.$$

Proof The first point is a standard fact (see, for instance, [3]). By taking the logarithm, we can deduce from the first point that

$$|\log h(z) - \log h(w)| \leq d_{\text{hyp}}(z, w),$$

hence, $|d \log h|_{\text{hyp}} \leq 1$ holds everywhere on $\mathbb{D}$.

Let us show the third point. We use the Poisson formula for positive harmonic functions (see, for instance, [24]),

$$(2-3) \quad h(z) = h(0) \int_{[0, 2\pi)} \frac{1 - |z|^2}{|e^{it} - z|^2} \mu(dt),$$

where μ is a Borel measure on $\mathbb{R}$ with $\mu([0, 2\pi)) = 1$ and invariant under $2\pi\mathbb{Z}$ -translation. Consider first the case when $|d \log h|_{\text{hyp}} = 1$ holds at $\alpha = 0 \in \mathbb{D}$. By a rotation, we may assume $(\log h)_{x_1}(0) = 2$, $(\log h)_{x_2}(0) = 0$, $m = 1$. Differentiating (2-3) at $z = 0$, we have

$$2 = (\log h)_{x_1}(0) = \int_{[0, 2\pi)} \left(\frac{\partial}{\partial x_1} \frac{1 - |z|^2}{|e^{it} - z|^2} \right) \Big|_{z=0} \mu(dt) = \int_{[0, 2\pi)} 2 \cos t \mu(dt).$$

Hence $1 \leq \cos t$ holds μ -a.e. and $\text{supp } \mu \subset 2\pi\mathbb{Z}$. Therefore, $\mu = \sum_{n \in \mathbb{Z}} \delta_{2\pi n}$ and

$$h(z) = h(0) \int_{[0, 2\pi)} \frac{1 - |z|^2}{|e^{it} - z|^2} \delta_0(dt) = h(0) \frac{1 - |z|^2}{|1 - z|^2}.$$

Now let $\alpha \in \mathbb{D}$ be arbitrary. Define $f_\alpha(z) = (z + \alpha)/(1 + \bar{\alpha}z) \in \text{PSU}(1, 1)$ and $g := h \circ f_\alpha$. Then

$$g_{x_1}(0) = (1 - |\alpha|^2)h_{x_1}(\alpha), \quad g_{x_2}(0) = (1 - |\alpha|^2)h_{x_2}(\alpha),$$

hence, $|d \log g|_{\text{hyp}} = 1$ holds at $0 \in \mathbb{D}$. Applying the argument in previous paragraph to g , we obtain

$$g(z) = g(0) \frac{1 - |z|^2}{|c - z|^2}, \quad c = \frac{h_{x_1}(\alpha) + i h_{x_2}(\alpha)}{|h_{x_1}(\alpha) + i h_{x_2}(\alpha)|},$$

and it follows that

$$h(z) = h(\alpha) \frac{1 - |f_\alpha^{-1}(z)|^2}{|c - f_\alpha^{-1}(z)|^2} = h(\alpha) \frac{1 - |\alpha|^2}{|c + \alpha|^2} \cdot \frac{1 - |z|^2}{|\frac{c+\alpha}{1+\bar{\alpha}c} - z|^2}.$$

By letting $z = 0$, we see that $h(\alpha)(1 - |\alpha|^2)/|c + \alpha|^2 = h(0)$, completing the proof. $\square$

2.3 The Euler number for surface group actions on S^1

Let Γ be a torsion-free lattice of $\text{PSU}(1, 1)$, and $\rho: \Gamma \rightarrow \text{Homeo}_+(S^1)$ a homomorphism. Let us recall the definition of the Euler number in [7]. If Γ is not uniform, clearly we cannot adopt the classical definition by the pairing of the Euler class and the fundamental class of $\Sigma = \Gamma \backslash \mathbb{D}$. Ghys [14] introduced the Euler class in the bounded cohomology. Burger, Iozzi and Wienhard [7] defined the Euler number based on Ghys' idea by using the relative fundamental class in the bounded cohomology as follows: Let Σ' be the compact surface with boundary obtained by cutting off all cusps of Σ . We can pull back the universal bounded Euler class $\text{eu} \in H_b^2(\text{Homeo}_+(S^1))$ by ρ to have $\rho^* \text{eu} \in H_b^2(\Gamma)$, which we regard as an element of $H_b^2(\Sigma')$ via the so-called Gromov isomorphism $H_b^2(\Gamma) \cong H_b^2(\Sigma')$. Since $H_b^1(\partial\Sigma') \cong H_b^2(\partial\Sigma') = 0$, by the relative exact sequence, we have an isomorphism

$$f: H_b^2(\Sigma', \partial\Sigma') \rightarrow H_b^2(\Sigma').$$

The Euler number $e(\rho) \in \mathbb{R}$ of ρ is defined by

$$e(\rho) = \langle f^{-1} \rho^* \text{eu}, [\Sigma', \partial\Sigma'] \rangle,$$

where $[\Sigma', \partial\Sigma']$ is the relative fundamental class of $(\Sigma', \partial\Sigma')$.

As with the classical Euler number, by a result of [7], the Euler number $e(\rho)$ is expressed in terms of the suspension bundle $\pi: M = \Sigma' \times_{\rho} S^1 \rightarrow \Sigma'$ and the translation numbers of the action of the boundary loops of Σ' . Let us consider the case where $\partial\Sigma' \neq \emptyset$. Let $c_1, \dots, c_m$ be the boundary loops of Σ' whose orientations are induced from Σ' . Take a homomorphism lift $\tilde{\rho}: \Gamma \rightarrow \widetilde{\text{Homeo}}_+(S^1)$ of ρ , where

$$\widetilde{\text{Homeo}}_+(S^1) = \{f \in \text{Homeo}_+(\mathbb{R}) \mid f(x + 2\pi) = f(x) + 2\pi, \forall x \in \mathbb{R}\}$$

is the universal cover group of $\text{Homeo}_+(S^1)$. Since $\pi: M \rightarrow \Sigma'$ is trivial as an S^1 -bundle, such lift exists. Let $\tau: \widetilde{\text{Homeo}}_+(S^1) \rightarrow \mathbb{R}$ be the translation number, which is defined by

$$\tau(f) = \frac{1}{2\pi} \lim_{n \rightarrow \infty} \frac{f^n(x) - x}{n},$$

where x is any point in $\mathbb{R}$ (see e.g. [21]). By [7], in the case where $\partial\Sigma' \neq \emptyset$, we have

$$e(\rho) = - \sum_{i=1}^m \tau(\tilde{\rho}(c_i)).$$

Note that the lift $\tilde{\rho}$ determines a trivialization of the S^1 -bundle $M \rightarrow \Sigma'$ up to homotopy, which is the image of the 0-section of the suspension $\mathbb{R}$ -bundle of $\tilde{\rho}$ under the projection $\Sigma' \times_{\tilde{\rho}} \mathbb{R} \rightarrow M$.

Our Theorem 1.1 states a Gauss–Bonnet formula that express this Euler number $e(\rho)$ in terms of a curvature integral with respect to a specifically chosen continuous connection. To this end, let us recall the usual Gauss–Bonnet formula for surfaces with boundaries. Let Σ' be a smooth compact surface with smooth boundary, and $\pi: M \rightarrow \Sigma'$ a smooth principal S^1 -bundle. We will call a continuous 1-form η on M a *continuous S^1 -connection form* on M if η is an S^1 -invariant continuous section of T^*M such that $\eta(\frac{\partial}{\partial\theta}) = 1$ on M , where $e^{i\theta} \in S^1$ and $\frac{\partial}{\partial\theta}$ denotes the generator of the principal S^1 -action. For each piecewise smooth closed curve γ in Σ' , $\text{hol}_{\eta}(\gamma) \in S^1$ denotes the holonomy along γ with respect to the parallel transport defined by the connection η .

Definition 2.4 Let Ω be an integrable 2-form on Σ' . For a (nonsmooth) continuous S^1 -connection form η on M , we call Ω the *curvature* of η if, for every simply connected subset U of Σ' with piecewise smooth boundary ∂U , we have $\text{hol}_{\eta}(\partial U) = \exp i \int_U \Omega$, where we regard $\text{hol}_{\eta}(\partial U) \in S^1$.

When η is smooth, it is well known that its curvature exists as a smooth 2-form Ω on Σ' such that $\pi^*\Omega = d\eta$. In the step 4 of the proof of the main result (Theorem 1.1), we will use the following Gauss–Bonnet formula for continuous connections whose curvature is well-defined in the sense of Definition 2.4.

Proposition 2.5 Let Σ' be a smooth compact surface with nonempty smooth boundary and $M \rightarrow \Sigma'$ a principal S^1 -bundle with a global smooth section σ . Let η be a continuous S^1 -connection form on M . If η has curvature Ω in the sense of Definition 2.4, then we have

$$(2-4) \quad \frac{1}{2\pi} \int_{\Sigma'} \Omega = \sum_{i=1}^m \tau(\widetilde{\text{hol}}_{\eta}(c_i)),$$

where $c_1, \dots, c_m$ are the boundary curves of Σ' whose orientations are induced from Σ' , and $\widetilde{\text{hol}}_{\eta}(c_i): \mathbb{R} \rightarrow \mathbb{R}$ is the lift of the holonomy map of η along c_i with respect to σ .

Proof Trivialize $M \cong \Sigma' \times S^1$ so that the image of σ is mapped to $\Sigma' \times \{1\}$. Consider the covering map $\varphi = \text{id} \times p: \Sigma' \times \mathbb{R} \rightarrow \Sigma' \times S^1$, where $p: \mathbb{R} \rightarrow S^1$ is the standard map $p(t) = e^{it}$. Equip $\Sigma' \times \mathbb{R}$ with a connection 1-form $\hat{\eta} = \varphi^* \eta$. Under this setting, the translation number of the holonomy of arbitrary paths with respect to σ makes sense: For a path c on Σ' , define $\widehat{\text{hol}}_\eta(c) = (\text{pr}_2(\hat{c}(1)) - \text{pr}_2(\hat{c}(0)))/2\pi$, where $\text{pr}_2: \Sigma' \times \mathbb{R} \rightarrow \mathbb{R}$ is the second projection and $\hat{c}$ is a horizontal lift of c with respect to $\hat{\eta}$. Note that, since the holonomy along every path is a translation on $\mathbb{R}$, this translation number is independent of the choice of $\hat{c}$. Moreover, if c is a piecewise smooth closed path, then we have $\tau(\widetilde{\text{hol}}_\eta(c)) = \widehat{\text{hol}}_\eta(c)$.

Triangulate Σ' piecewise smoothly. In order to prove (2-4), it suffices to show

$$(2-5) \quad \frac{1}{2\pi} \int_s \Omega = \tau(\widetilde{\text{hol}}_\eta(\partial s))$$

for each 2-simplex s . Indeed we have $\tau(\widetilde{\text{hol}}_\eta(\partial s)) = \sum_{k=0}^2 \widehat{\text{hol}}_\eta(t_k)$, where t_0, t_1, t_2 are the edges of s , and (2-4) is obtained as the sum of (2-5) for all 2-simplices. Finally let us explain that (2-5) follows from Definition 2.4. By definition, we have $\frac{1}{2\pi} \int_s \Omega - \tau(\widetilde{\text{hol}}_\eta(\partial s)) \in \mathbb{Z}$. Since Ω is integrable, this integer depends continuously on s , and hence it is constant. When we contract s to a point, the difference $\frac{1}{2\pi} \int_s \Omega - \tau(\widetilde{\text{hol}}_\eta(\partial s))$ goes to zero, which implies (2-5). $\square$

3 The Gauss–Bonnet formula for the connection associated with a harmonic measure

We shall prove a Gauss–Bonnet formula for the S^1 -connection $\bar{\omega}$ associated with a harmonic measure (Theorem 1.1). Let Γ be a torsion-free lattice of $\text{PSU}(1, 1)$, ρ a homomorphism $\Gamma \rightarrow \text{Homeo}_+(S^1)$ and $\Sigma = \Gamma \backslash \mathbb{D}$. We assume that $\rho(\Gamma)$ has no finite orbit in S^1 . Consider the suspension foliation on $M = \Sigma \times_\rho S^1$, which is induced from $\{\mathbb{D} \times \{e^{it}\}\}_{e^{it} \in S^1}$, equipped with the standard leafwise hyperbolic metric. The proof of Theorem 1.1 is divided into the following four steps.

Step 1 Collapsing the complement of the support of a harmonic measure

In this first step, we construct a semiconjugacy of the given representation ρ to a new representation ρ' by integrating fiberwise measures to collapse the complement of the support of the given harmonic measure.

As mentioned above, the suspension foliation $\mathcal{F}$ has a leafwise hyperbolic metric. Take a harmonic probability measure μ on $\mathcal{F}$, which exists from Theorem 2.2. By multiplying a scalar, we normalize μ so that $\mu(M) = 4\pi^2 |e(\Sigma)|$.

Since the natural projection $\mathbb{D} \times \mathbb{R} \rightarrow \Sigma \times_\rho S^1$ is a covering map, we can pull back μ to $\mathbb{D} \times \mathbb{R}$ and denote it by $\tilde{\mu}$, which is a harmonic measure on the product foliation $\{\mathbb{D} \times \{t\}\}_{t \in \mathbb{R}}$. By (2-2), we may write

$$\tilde{\mu} = q(z, t) \text{vol}(z) \nu(t),$$

where vol is the leafwise volume measure on $\mathbb{D}$, ν is a Borel measure on $\mathbb{R}$, q is a locally integrable function on $\mathbb{D} \times \mathbb{R}$ with respect to $\text{vol}(z)\nu(t)$ and, for ν -a.e. t , $q(\cdot, t)$ is a positive harmonic function on $\mathbb{D}$.

For each $z \in \mathbb{D}$, consider a measure μ_z on S^1 induced by $\tilde{\mu}_z := q(z, t)\nu(t)$ on $\mathbb{R}$. Since $\tilde{\mu}$ is Γ -invariant by construction, it follows that, for each $\gamma \in \Gamma$, we have

$$(3-1) \quad \mu_{\gamma z} = \rho(\gamma)_* \mu_z.$$

Claim 3.1 μ_z is a nonatomic measure for every $z \in \mathbb{D}$.

Proof Assume on the contrary that $\mu_{z_0}(\{a\}) > 0$ for some $z_0 \in \mathbb{D}$ and some $a \in S^1$. Then, $\nu(\{a\}) > 0$. Let L be the leaf of $\mathcal{F}$ induced from $\mathbb{D} \times \{a\}$. The measure $\nu(\{a\})q(z, a)\text{vol}(z)\delta_a(t)$ supported on $\mathbb{D} \times \{a\}$ induces a measure μ_L supported on L . Then,

$$0 < \mu_L(L) \leq \mu(M) < \infty$$

and $q(z, a)$ induces an $L^1(\text{vol})$ positive harmonic function on L . Since $\rho(\Gamma)$ is assumed to have no finite orbit in S^1 , any leaf of $\mathcal{F}$, being an infinite covering of Σ , is a complete hyperbolic surface of infinite volume. This leads to a contradiction by [16, Theorem 2.4] or [15, Theorem 1]. $\square$

Claim 3.2 $\mu_z(S^1) = 2\pi$ for every $z \in \mathbb{D}$.

Proof From the invariance (3-1), $\mu_z(S^1)$ defines an $L^1(\text{vol})$ positive harmonic function on Σ . This function must be constant from [16, Theorem 2.4] or [15, Theorem 1]. Since μ is normalized so that $\mu(M) = 4\pi^2|e(\Sigma)|$ and the volume of Σ is $2\pi|e(\Sigma)|$, this constant is 2π . $\square$

Define a monotone map $\tilde{\psi}: \mathbb{R} \rightarrow \mathbb{R}$ by

$$\tilde{\psi}(t) := \int_0^t \tilde{\mu}_0(ds) = \int_0^t q(0, s)\nu(ds).$$

By construction, $\tilde{\psi}$ induces a monotone continuous map $\psi: S^1 \rightarrow S^1$ of mapping degree one.

Claim 3.3 There exists a group homomorphism $\rho': \Gamma \rightarrow \text{Homeo}_+(S^1)$ such that ψ is (ρ, ρ') -equivariant, namely, we have

$$(3-2) \quad \psi \circ \rho(\gamma) = \rho'(\gamma) \circ \psi \quad (\text{for all } \gamma \in \Gamma).$$

Proof Fix $\gamma \in \Gamma$. In order to define $\rho'(\gamma): S^1 \rightarrow S^1$, take $e^{it'} \in S^1$. Since ψ is surjective, there exists $e^{it} \in S^1$ such that $\psi(e^{it}) = e^{it'}$. Then we define $\rho'(\gamma)e^{it'} := \psi(\rho(\gamma)e^{it})$. It is easy to see that this ρ' is well-defined, and that ρ' is a group homomorphism $\Gamma \rightarrow \text{Homeo}_+(S^1)$. Equation (3-2) follows from the construction. $\square$

Actually, ρ' is a representation of Γ in the bi-Lipschitz homeomorphism group of S^1 , as it will follow from the arguments in the next step.

Step 2 Construction of a smooth structure

Consider the suspension bundle $\pi': \Sigma \times_{\rho'} S^1 \rightarrow \Sigma$ for ρ' and let $\mathcal{F}'$ denote the suspension foliation on $M' := \Sigma \times_{\rho'} S^1$. In this step, we construct a smooth structure on our S^1 -bundle M' with respect to which $\mathcal{F}'$ is a transversely Lipschitz foliation.

Using ψ constructed in the first step, we define

$$\tilde{\Psi}: \mathbb{D} \times \mathbb{R} \rightarrow \mathbb{D} \times \mathbb{R}, \quad \tilde{\Psi}(z, t) := (z, \tilde{\psi}(t))$$

that induces a surjective continuous map $\Psi: M \rightarrow M'$. Let us consider the push-forward measures $\mu' := \Psi_*\mu$ and $\tilde{\mu}' := \tilde{\Psi}_*\tilde{\mu}$, which are harmonic measures on $\mathcal{F}'$ and the product foliation $\{\mathbb{D} \times \{t\}\}_{t \in \mathbb{R}}$, respectively. By (2-2), like as $\tilde{\mu}$, we have

$$(3-3) \quad \tilde{\mu}' = h(z, t) \operatorname{vol}(z) \lambda(t),$$

where λ is the Lebesgue measure on $\mathbb{R}$ and h is a locally integrable function on $\mathbb{D} \times \mathbb{R}$ such that, for a.e. t , $h(\cdot, t)$ is a positive harmonic function on $\mathbb{D}$. In particular, $h(0, t) = 1$ holds for a.e. t . Note that $\tilde{\mu}'_z := h(z, t) \lambda(t)$ induces a nonatomic measure μ'_z on S^1 with $\mu'_z(S^1) = 2\pi$ for each $z \in \mathbb{D}$.

To construct a desired smooth structure on M' , we integrate the harmonic measure μ' on each fiber of $\pi': M' \rightarrow \Sigma$. The next claim, a direct consequence of Harnack's inequality, is a key ingredient of our proof and will justify analytic arguments in next steps.

Lemma 3.4 Define a map $\tilde{\Phi}: \mathbb{D} \times \mathbb{R} \rightarrow \mathbb{D} \times \mathbb{R}$ by $\tilde{\Phi}(z, t) := (z, \varphi(z, t))$, where

$$\varphi(z, t) := \int_0^t h(z, s) ds$$

and h is the function that appeared in (3-3). Then $\tilde{\Phi}$ is a locally bi-Lipschitz homeomorphism. In particular, for every $z \in \mathbb{D}$, the map $\varphi(z, \cdot)$ is a locally bi-Lipschitz homeomorphism of $\mathbb{R}$, whose inverse is denoted by $\tau(z, \cdot)$.

Proof It is clear that $\tilde{\Phi}$ is bijective. To show that $\tilde{\Phi}$ is locally Lipschitz, it is enough to see that $\varphi: \mathbb{D} \times \mathbb{R} \rightarrow \mathbb{R}$ is Lipschitz in each variable on

$$D_N := \{(z, t) \in \mathbb{D} \times \mathbb{R} \mid d_{\text{hyp}}(0, z) \leq N, |t| \leq 2\pi N\}$$

for arbitrary $N \in \mathbb{N}$. Since Lemma 2.3 implies that, for a.e. t ,

$$0 < h(z, t) \leq e^{d_{\text{hyp}}(0, z)} h(0, t) = e^{d_{\text{hyp}}(0, z)} \leq e^N,$$

we have, for $(z, t_1), (z, t_2) \in D_N$,

$$|\varphi(z, t_1) - \varphi(z, t_2)| = \left| \int_{t_2}^{t_1} h(z, t) dt \right| \leq e^N |t_1 - t_2|.$$

Next, let $(z_1, t), (z_2, t) \in D_N$. [Lemma 2.3](#) also yields that, for a.e. t ,

$$|h(z_1, t) - h(z_2, t)| \leq (e^{d_{\text{hyp}}(z_1, z_2)} - 1) \max\{h(z_1, t), h(z_2, t)\}.$$

Since the exponential function is convex, we have

$$e^{d_{\text{hyp}}(z_1, z_2)} - 1 \leq \frac{e^{2N}}{2N} d_{\text{hyp}}(z_1, z_2).$$

Then, we have

$$(3-4) \quad |\varphi(z_1, t) - \varphi(z_2, t)| \leq \left| \int_0^t |h(z_1, s) - h(z_2, s)| ds \right| \leq \pi e^{3N} d_{\text{hyp}}(z_1, z_2).$$

Note that this inequality holds also for $t < 0$. Therefore, φ is locally Lipschitz, and hence so is $\tilde{\Phi}$.

It remains to show that $\tilde{\Phi}^{-1}$ is locally Lipschitz. We denote the inverse map of $\varphi(z, \cdot): \mathbb{R} \rightarrow \mathbb{R}$ by $\tau(z, \cdot)$ for each $z \in \mathbb{D}$. Again, it is enough to see that τ is Lipschitz in each variable on D_N for arbitrary $N \in \mathbb{N}$. Let $(z, \theta_1), (z, \theta_2) \in D_N$ and let $t_j := \tau(z, \theta_j)$ for $j = 1, 2$. From [Lemma 2.3](#), we have

$$h(z, t) \geq e^{-d_{\text{hyp}}(0, z)} h(0, t) = e^{-d_{\text{hyp}}(0, z)} \geq e^{-N}$$

for a.e. t . Hence, it follows that

$$|\varphi(z, t_1) - \varphi(z, t_2)| = \left| \int_{t_2}^{t_1} h(z_1, t) dt \right| \geq e^{-N} |t_1 - t_2|,$$

which is equivalent to

$$(3-5) \quad |\tau(z, \theta_1) - \tau(z, \theta_2)| \leq e^N |\theta_1 - \theta_2|.$$

Next, for each $(z_1, \theta), (z_2, \theta) \in D_N$, we deduce from [\(3-5\)](#) and [\(3-4\)](#) that

$$\begin{aligned} |\tau(z_1, \theta) - \tau(z_2, \theta)| &= |\tau(z_2, \varphi(z_2, \tau(z_1, \theta))) - \tau(z_2, \theta)| \leq e^N |\varphi(z_2, \tau(z_1, \theta)) - \theta| \\ &= e^N |\varphi(z_2, \tau(z_1, \theta)) - \varphi(z_1, \tau(z_1, \theta))| \\ &\leq \pi e^{4N} d_{\text{hyp}}(z_1, z_2). \end{aligned}$$

This completes the proof. □

Using $\tilde{\Phi}$, we define a new Γ -action on $\mathbb{D} \times S^1$ by

$$(3-6) \quad \gamma \cdot (z, e^{i\theta}) := (\gamma z, e^{i\varphi(\gamma z, \tilde{\rho}'(\gamma)(\tau(z, \theta)))}),$$

where $\gamma \in \Gamma$, $(z, e^{i\theta}) \in \mathbb{D} \times S^1$ and $\tilde{\rho}'(\gamma)$ denotes an arbitrary lift of $\rho'(\gamma) \in \text{Homeo}_+(S^1)$ to $\text{Homeo}_+(\mathbb{R})$. We see from the following claim that this Γ -action is smooth.

Claim 3.5 $\varphi(\gamma z, \tilde{\rho}'(\gamma)(\tau(z, \theta)))$ is a smooth function on $\mathbb{D} \times \mathbb{R}$.

Proof Using the invariance $\mu'_{\gamma z} = \rho'(\gamma)_* \mu'_z$, we have

$$\begin{aligned} \varphi(\gamma z, \tilde{\rho}'(\gamma)(\tau(z, \theta))) &= \int_0^{\tilde{\rho}'(\gamma)(\tau(z, \theta))} \mu'_{\gamma z}(dt) \\ &= \int_{\tilde{\rho}'(\gamma)(0)}^{\tilde{\rho}'(\gamma)(\tau(z, \theta))} \mu'_{\gamma z}(dt) + \int_0^{\tilde{\rho}'(\gamma)(0)} \mu'_{\gamma z}(dt) \\ &= \int_0^{\tau(z, \theta)} \mu'_z(dt) + \int_0^{\tilde{\rho}'(\gamma)(0)} \mu'_{\gamma z}(dt) \\ &= \varphi(z, \tau(z, \theta)) + \int_0^{\tilde{\rho}'(\gamma)(0)} \mu'_{\gamma z}(dt) \\ &= \theta + \int_0^{\tilde{\rho}'(\gamma)(0)} h(\gamma z, t) dt. \end{aligned}$$

The last term $\int_0^{\tilde{\rho}'(\gamma)(0)} h(\gamma z, t) dt$ is a positive harmonic function in z , hence, smooth in z . $\square$

Using the Γ -action given in (3-6), we define a smooth S^1 -bundle $P := \Gamma \backslash (\mathbb{D} \times S^1)$. The bi-Lipschitz homeomorphism $\tilde{\Phi}$ induces a homeomorphism $\Phi: M' \rightarrow P$.

Step 3 The S^1 -connection associated with a harmonic measure

In this step, we construct a continuous connection form on P whose curvature is bounded by one everywhere in its modulus.

Let $\mathcal{F}_P$ denote the transversely Lipschitz foliation on P obtained by mapping the suspension foliation $\mathcal{F}'$ on M' by Φ . Let $\mathcal{F}_{\tilde{P}}$ denote its lift to $\tilde{P} := \mathbb{D} \times \mathbb{R}$. Note that the leaves of $\mathcal{F}_{\tilde{P}}$ are the graphs of positive harmonic functions $\varphi(\cdot, t): \mathbb{D} \rightarrow \mathbb{R}$. Let $z = x_1 + ix_2 \in \mathbb{D}$. Let us compute the slope of this graph. Recall that, for each $z \in \mathbb{D}$, the map $\tau(z, \cdot): \mathbb{R} \rightarrow \mathbb{R}$ is the inverse of a locally bi-Lipschitz homeomorphism $\varphi(z, \cdot)$, which was defined by $\varphi(z, t) = \int_0^t h(z, s) ds$ in Lemma 3.4. For $j = 1, 2$, let

$$\omega_j(z, \theta) := \frac{\partial \varphi}{\partial x_j}(z, t) \Big|_{t=\tau(z, \theta)}.$$

Lemma 3.6 *The function ω_j is continuous on $\tilde{P}$. For each $z \in \mathbb{D}$, $\omega_j(z, \cdot)$ is a periodic Lipschitz function on $\mathbb{R}$ of period 2π .*

Proof By Lemma 2.3, we have

$$\left| \frac{\partial h}{\partial x_j}(z, t) \right| \leq \frac{2}{1-|z|^2} h(z, t) \leq \frac{2}{1-|z|^2} e^{d_{\text{hyp}}(0, z)},$$

which implies that, for any compact subset $K \subset \mathbb{D}$, it is bounded from above on $K \times \mathbb{R}$. Therefore we can change the order of differential and integration to have

$$\frac{\partial \varphi}{\partial x_j}(z, t) = \frac{\partial}{\partial x_j} \int_0^t h(z, s) ds = \int_0^t \frac{\partial h}{\partial x_j}(z, s) ds.$$

From this expression and [Lemma 3.4](#) it follows that ω_j is continuous.

We also have

$$\left| \frac{\partial \varphi}{\partial x_j}(z, t_1) - \frac{\partial \varphi}{\partial x_j}(z, t_2) \right| \leq \int_{t_2}^{t_1} \left| \frac{\partial h}{\partial x_j}(z, s) \right| ds \leq \frac{2}{1-|z|^2} \int_{t_2}^{t_1} h(z, s) ds = \frac{2}{1-|z|^2} |\varphi(z, t_1) - \varphi(z, t_2)|$$

for any $t_1 \geq t_2$. Therefore, we conclude that $\omega_j(z, \cdot)$ is a Lipschitz function on $\mathbb{R}$:

$$|\omega_j(z, \theta_1) - \omega_j(z, \theta_2)| \leq \frac{2}{1-|z|^2} |\theta_1 - \theta_2|.$$

The periodicity of $\omega_j(z, \cdot)$ is clear. □

Remark 3.7 We do not know if $\omega_j(\cdot, \theta)$ is locally Lipschitz or not, while it is continuous.

We will construct a connection on P by taking average of the slope $\omega_j(z, \cdot)$; consider an S^1 -invariant continuous 1-form

$$\bar{\omega} := d\theta - \sum_{j=1,2} \left(\int_0^{2\pi} \omega_j(z, \theta) \frac{d\theta}{2\pi} \right) dx_j$$

on $\mathbb{D} \times S^1$. Since $\bar{\omega}$ is Γ -invariant, we can identify it with a continuous 1-form on P . This gives a continuous connection on P .

Let us show that $\bar{\omega}$ has a curvature form. Define

$$K(z) := -\frac{(1-|z|^2)^2}{4} \int_0^{2\pi} \left(-\frac{\partial \omega_1}{\partial \theta}(z, \theta) \omega_2(z, \theta) + \frac{\partial \omega_2}{\partial \theta}(z, \theta) \omega_1(z, \theta) \right) \frac{d\theta}{2\pi},$$

and regard the Γ -invariant form $K(z) \text{vol}(dz)$ as a measurable 2-form on Σ . Recall that the hyperbolic volume form is of the form $4dz \wedge d\bar{z}/(1-|z|^2)^2$ ($z \in \mathbb{D}$).

Lemma 3.8 $\bar{\omega}$ has the curvature $K(z) \text{vol}(dz)$ in the sense of [Definition 2.4](#).

Proof Take a piecewise smooth simple closed curve $\gamma: [0, 1] \rightarrow \mathbb{D}$. Let U be the domain bounded by γ . It is enough to show

$$\text{hol}_{\bar{\omega}}(\gamma) = \exp i \int_U K(z) \text{vol}(dz)$$

assuming that γ is smooth. Consider the horizontal lift $\tilde{\gamma}: [0, 1] \rightarrow \mathbb{D} \times \mathbb{R}$ of γ to $\tilde{P}$ with respect to $\bar{\omega}$ whose initial value is $\tilde{\gamma}(0) = 0$. For $\exp i \tilde{\gamma}(1) = \text{hol}_{\bar{\omega}}(\gamma)$, we need to show $\tilde{\gamma}(1) = \int_U K(z) \text{vol}(dz)$.

Since $\tilde{\gamma}$ is a solution of an ordinary differential equation,

$$\frac{d\tilde{\gamma}}{du}(u) = \sum_{j=1,2} \left(\int_0^{2\pi} \omega_j(\gamma(u), \theta) \frac{d\theta}{2\pi} \right) \frac{d\gamma_j}{du}(u),$$

and the right hand side does not contain $\tilde{\gamma}$, we can integrate it to compute the solution:

$$(3-7) \quad \tilde{\gamma}(1) = \int_0^1 du \sum_{j=1,2} \left(\int_0^{2\pi} \omega_j(\gamma(u), \theta) \frac{d\theta}{2\pi} \right) \frac{d\gamma_j}{du}(u) = \int_0^{2\pi} \frac{d\theta}{2\pi} \int_{\gamma} \omega_1(z, \theta) dx_1 + \omega_2(z, \theta) dx_2.$$

We would like to apply Stokes theorem to compute the path integral, but the difficulty here is that the $\omega_j(z, \theta)$ are merely continuous in z . We need to approximate $\varphi(z, t)$ by a sequence of smooth functions. By using mollifiers, we can construct smooth functions $\{f_n(z, t)\}$ on $\bar{U} \times [0, 2\pi]$ such that

$$\sup_{\bar{U} \times [0, 2\pi]} |f_n - \varphi| + \sum_{j=1,2} \sup_{\bar{U} \times [0, 2\pi]} \left| \frac{\partial f_n}{\partial x_j} - \frac{\partial \varphi}{\partial x_j} \right| \rightarrow 0, \quad \sum_{j=1,2} \int_{\bar{U} \times [0, 2\pi]} \left| \frac{\partial^2 f_n}{\partial t \partial x_j} - \frac{\partial^2 \varphi}{\partial t \partial x_j} \right| dx_1 dx_2 dt \rightarrow 0$$

as $n \rightarrow \infty$. Note that $\varphi(z, t)$ and $(\partial\varphi/\partial x_j)(z, t)$ are continuous, and

$$\frac{\partial \varphi}{\partial t}(z, t) = h(z, t) \quad \text{and} \quad \frac{\partial^2 \varphi}{\partial t \partial x_j}(z, t) = \frac{\partial h}{\partial x_j}(z, t)$$

are bounded measurable and integrable from Lemmas 2.3 and 3.6. Moreover, from Lemma 3.4,

$$\int_{\bar{U} \times [0, 2\pi]} \left| \frac{\partial^2 f_n}{\partial t \partial x_j}(z, \tau(z, \theta)) - \frac{\partial^2 \varphi}{\partial t \partial x_j}(z, \tau(z, \theta)) \right| dx_1 dx_2 d\theta \rightarrow 0$$

follows. Hence, Fubini's theorem yields that, for a.e. θ ,

$$(3-8) \quad \int_{\bar{U}} \left| \frac{\partial^2 f_n}{\partial t \partial x_j}(z, \tau(z, \theta)) - \frac{\partial^2 \varphi}{\partial t \partial x_j}(z, \tau(z, \theta)) \right| dx_1 dx_2 \rightarrow 0.$$

Using this sequence, we can approximate

$$\int_{\gamma} \omega_1(z, \theta) dx_1 + \omega_2(z, \theta) dx_2 = \lim_{n \rightarrow \infty} \int_{\gamma} \frac{\partial f_n}{\partial x_1}(z, \tau(z, \theta)) dx_1 + \frac{\partial f_n}{\partial x_2}(z, \tau(z, \theta)) dx_2.$$

From Lemma 3.4, the integrand in the approximation sequence is a Lipschitz 1-form and we can apply Stokes theorem for it. Note that we can show Stokes theorem for Lipschitz 1-forms by using that fact that the second fundamental theorem of calculus holds for Lipschitz functions (see e.g., [22, Section 1.6.4]). Note also that Stokes theorem holds in more general situations (see [11, Theorem 4.5.6]).

It follows that

$$\begin{aligned} \int_{\gamma} \frac{\partial f_n}{\partial x_1}(z, \tau(z, \theta)) dx_1 + \frac{\partial f_n}{\partial x_2}(z, \tau(z, \theta)) dx_2 \\ = \int_U \left(-\frac{\partial}{\partial x_2} \frac{\partial f_n}{\partial x_1}(z, \tau(z, \theta)) + \frac{\partial}{\partial x_1} \frac{\partial f_n}{\partial x_2}(z, \tau(z, \theta)) \right) dx_1 dx_2 \\ = \int_U \left(-\frac{\partial^2 f_n}{\partial t \partial x_1}(z, \tau(z, \theta)) \frac{\partial \tau}{\partial x_2}(z, \theta) + \frac{\partial^2 f_n}{\partial t \partial x_2}(z, \tau(z, \theta)) \frac{\partial \tau}{\partial x_1}(z, \theta) \right) dx_1 dx_2. \end{aligned}$$

Notice that $\partial^2 f_n / (\partial x_1 \partial x_2)$ is canceled out. By letting $n \rightarrow \infty$, (3-8) implies

$$\begin{aligned} \int_{\gamma} \omega_1(z, \theta) dx_1 + \omega_2(z, \theta) dx_2 \\ = \int_U \left(-\frac{\partial^2 \varphi}{\partial t \partial x_1}(z, \tau(z, \theta)) \frac{\partial \tau}{\partial x_2}(z, \theta) + \frac{\partial^2 \varphi}{\partial t \partial x_2}(z, \tau(z, \theta)) \frac{\partial \tau}{\partial x_1}(z, \theta) \right) dx_1 dx_2 \end{aligned}$$

for a.e. θ . On the other hand, thanks to Lemma 3.4, we can differentiate $\varphi(z, \tau(z, \theta)) = \theta$ by the chain rule and see that

$$\frac{\partial \varphi}{\partial x_j}(z, \tau(z, \theta)) + \frac{\partial \varphi}{\partial t}(z, \tau(z, \theta)) \frac{\partial \tau}{\partial x_j}(z, \theta) = 0$$

holds for a.e. z and all θ . Also,

$$\frac{\partial \varphi}{\partial t}(z, \tau(z, \theta)) \frac{\partial \tau}{\partial \theta}(z, \theta) = 1$$

is true for a.e. θ and all z . Hence, it follows that, for a.e. θ and a.e. z ,

$$\frac{\partial \tau}{\partial x_j}(z, \theta) = -\omega_j(z, \theta) \frac{\partial \tau}{\partial \theta}(z, \theta)$$

holds, and we have

$$\int_{\gamma} \omega_1(z, \theta) dx_1 + \omega_2(z, \theta) dx_2 = \int_U \left(\frac{\partial \omega_1}{\partial \theta}(z, \theta) \omega_2(z, \theta) - \frac{\partial \omega_2}{\partial \theta}(z, \theta) \omega_1(z, \theta) \right) dx_1 dx_2$$

for a.e. θ . Therefore, by changing the order of integrals in (3-7), we have $\tilde{\gamma}(1) = \int_U K(z) \text{vol}(dz)$. $\square$

Applying the isoperimetric inequality and Harnack's inequality, we show that $|K(z)|$ is bounded by one everywhere.

Claim 3.9 For every $z \in \mathbb{D}$, we have $|K(z)| \leq 1$.

Proof Among the definition of $K(z)$,

$$\int_0^{2\pi} \frac{1}{2} \left(-\frac{\partial \omega_1}{\partial \theta}(z, \theta) \omega_2(z, \theta) + \frac{\partial \omega_2}{\partial \theta}(z, \theta) \omega_1(z, \theta) \right) d\theta$$

is the signed area of the domain bounded by a Lipschitz closed curve $(\omega_1(z, \theta), \omega_2(z, \theta))$ ($0 \leq \theta \leq 2\pi$) on $\mathbb{R}^2$. The tangent vector of this curve is defined for a.e. θ and it is

$$\begin{aligned} \frac{\partial}{\partial \theta} \omega_j(z, \theta) &= \frac{\partial}{\partial \theta} \int_0^{\tau(z, \theta)} \frac{\partial h}{\partial x_j}(z, t) dt = \frac{\partial \tau}{\partial \theta}(z, \theta) \frac{\partial h}{\partial x_j}(z, \tau(z, \theta)) \\ &= \left(\frac{\partial \varphi}{\partial t}(z, \tau(z, \theta)) \right)^{-1} \frac{\partial h}{\partial x_j}(z, \tau(z, \theta)) = \frac{\partial \log h}{\partial x_j}(z, \tau(z, \theta)). \end{aligned}$$

Then, by the isoperimetric inequality, we have

$$|K| \leq \frac{(1 - |z|^2)^2}{4\pi} \cdot \frac{1}{4\pi} \left(\int_0^{2\pi} \sqrt{\left(\frac{\partial \log h}{\partial x_1} \right)^2 + \left(\frac{\partial \log h}{\partial x_2} \right)^2} d\theta \right)^2 = \frac{1}{4\pi^2} \left(\int_0^{2\pi} |d \log h|_{\text{hyp}} d\theta \right)^2,$$

which implies $|K| \leq (2\pi)^2/4\pi^2 = 1$ by Harnack's inequality. Note that the factor $\frac{1}{4}(1 - |z|^2)^2$ comes from the Poincaré metric on $\mathbb{D}$ (see (2-1)). $\square$

This completes the proof for the first part of [Theorem 1.1](#).

Step 4 The translation number of the holonomy along horocircles

We will show the remaining part of [Theorem 1.1](#), namely, the Gauss–Bonnet formula

$$e(\rho) = \frac{1}{2\pi} \int_{\Sigma} K(z) \operatorname{vol}(dz).$$

Let ρ' be the Γ -action on S^1 obtained in [Step 1](#) by collapsing the complement of the harmonic measure μ . The suspension foliation of ρ' has the harmonic measure induced from μ , which was denoted by μ' . Let us consider the S^1 -connection $\bar{\omega}$ on the suspension bundle of ρ' constructed in [Step 3](#), which is the fiberwise μ' -average of the flat connection. Since the Euler class in the bounded cohomology is invariant under semiconjugacy due to an observation of Ghys [\[14\]](#) (see also [\[6, Theorem 4.3\]](#)), we have $e(\rho) = e(\rho')$. Therefore, in order to prove [Theorem 1.1](#), it suffices to show

$$(3-9) \quad e(\rho') = \frac{1}{2\pi} \int_{\Sigma} K(z) \operatorname{vol}(dz).$$

This is well-known when Σ is a closed surface. We discuss the case where Σ has cusps. Take a homomorphism lift $\tilde{\rho}': \Gamma \rightarrow \widetilde{\operatorname{Homeo}}_+(S^1)$ of ρ' . Then, by [\[7\]](#), we have

$$(3-10) \quad e(\rho') = - \sum_{i=1}^m \tau(\tilde{\rho}'(c_i)),$$

where $\tau: \widetilde{\operatorname{Homeo}}_+(S^1) \rightarrow \mathbb{R}$ is the translation number and $c_i \in \Gamma$ corresponds to a loop that goes around the i^{th} cusp of Σ for $i = 1, \dots, m$. A neighborhood of each cusp is foliated by closed horocircles. For each i , let $\{c_i^s\}_{s \gg 0}$ be the family of horocircles in a neighborhood of the i^{th} cusp so that the hyperbolic diameter δ_i^s of c_i^s tends to 0 and c_i^s approaches to the i^{th} cusp as $s \rightarrow \infty$. For $s \gg 0$, let Σ^s be the compact subsurface of Σ bounded by $c_1^s, \dots, c_m^s$. We denote by σ the global section of $P \rightarrow \Sigma$ that corresponds to the choice of the lift $\tilde{\rho}'$ of ρ' . For $s \gg 0$, by [Proposition 2.5](#), we have

$$(3-11) \quad \frac{1}{2\pi} \int_{\Sigma^s} K(z) \operatorname{vol}(dz) = \sum_{i=1}^m \tau(\widetilde{\operatorname{hol}}_{\bar{\omega}}(c_i^s)),$$

where $\widetilde{\operatorname{hol}}_{\bar{\omega}}(c_i^s): \mathbb{R} \rightarrow \mathbb{R}$ is the lift of the holonomy map of $\bar{\omega}$ along c_i^s with respect to σ . Since $\Gamma \backslash \mathbb{D}$ has finite volume and $|K(z)| \leq 1$ a.e. z from [Claim 3.9](#), we have $\int_{\Sigma^s} K(z) \operatorname{vol}(dz) \rightarrow \int_{\Sigma} K(z) \operatorname{vol}(dz)$ as $s \rightarrow \infty$. Therefore, by (3-10) and (3-11), for the proof of the Gauss–Bonnet formula (3-9), it remains to find a sequence $\{s_n\}$ such that

$$(3-12) \quad \tau(\widetilde{\operatorname{hol}}_{\bar{\omega}}(c_i^{s_n})) \rightarrow -\tau(\tilde{\rho}'(c_i))$$

as $n \rightarrow \infty$ for each fixed i .

Take a lift of c_i^s on $\mathbb{D}$,

$$\gamma^s = \gamma_1^s + i\gamma_2^s: [0, 1] \rightarrow \mathbb{D},$$

so that it is continuous in s . We use the coordinate (z, t) and (z, θ) of $\mathbb{D} \times \mathbb{R}$ used in [Step 3](#) for the pull back bundle of $\Sigma \times_{\tilde{\rho}'} \mathbb{R}$ over $\mathbb{D}$, and write $\tilde{P}$ for $\mathbb{D} \times \mathbb{R}$ equipped with (z, θ) -coordinate. Denote by $\tilde{\sigma}$ a lift of σ to a global section of $\mathbb{D} \times \mathbb{R} \rightarrow \mathbb{D}$, and express $\tilde{\sigma}$ as the graph of a smooth function $\tilde{\sigma}: \mathbb{D} \rightarrow \mathbb{R}$ in (z, θ) -coordinate.

Since $(\gamma^s(0), \tilde{\sigma}(\gamma^s(0))), (\gamma^s(1), \tilde{\sigma}(\gamma^s(1))) \in \tilde{P}$ are identified in $\Sigma \times_{\tilde{\rho}'} \mathbb{R}$, it follows from [\(3-7\)](#) that

$$2\pi \tau(\widetilde{\text{hol}}_{\tilde{\omega}}(c_i^s)) = \int_0^{2\pi} \frac{d\theta}{2\pi} \int_{\gamma^s} (\omega_1(z, \theta) dx_1 + \omega_2(z, \theta) dx_2) + \tilde{\sigma}(\gamma^s(0)) - \tilde{\sigma}(\gamma^s(1)).$$

We set $H(s)$ to be the first term in the right-hand side. We use the locally bi-Lipschitz homeomorphism $\mathbb{D} \times \mathbb{R} \rightarrow \tilde{P}$,

$$(z, t) \mapsto (z, \theta), \quad \theta = \varphi(z, t) = \int_0^t h(z, s) ds$$

considered in [Lemma 3.4](#) to change variables of the integration, which is justified thanks to Rademacher's theorem (see e.g., [\[11, Theorem 3.2.3\]](#)). It follows that

$$\begin{aligned} H(s) &= \int_{[0,1] \times [0,2\pi]} \left(\omega_1(\gamma^s(u), \theta) \frac{d\gamma_1^s}{du}(u) + \omega_2(\gamma^s(u), \theta) \frac{d\gamma_2^s}{du}(u) \right) \frac{du d\theta}{2\pi} \\ &= \int_{[0,1] \times [0,2\pi]} \left(\frac{\partial \varphi}{\partial x_1}(\gamma^s(u), t) \frac{d\gamma_1^s}{du}(u) + \frac{\partial \varphi}{\partial x_2}(\gamma^s(u), t) \frac{d\gamma_2^s}{du}(u) \right) h(\gamma^s(u), t) \frac{du dt}{2\pi}. \end{aligned}$$

Harnack's inequality implies that

$$(3-13) \quad e^{-\delta_i^s} h(\gamma^s(0), t) \leq h(\gamma^s(u), t) \leq e^{\delta_i^s} h(\gamma^s(0), t)$$

for any $u \in [0, 1]$. So, we may approximate $H(s)$ by

$$H_0(s) := \int_{[0,1] \times [0,2\pi]} \left(\frac{\partial \varphi}{\partial x_1}(\gamma^s(u), t) \frac{d\gamma_1^s}{du}(u) + \frac{\partial \varphi}{\partial x_2}(\gamma^s(u), t) \frac{d\gamma_2^s}{du}(u) \right) h(\gamma^s(0), t) \frac{du dt}{2\pi}.$$

We can rewrite

$$\begin{aligned} H_0(s) &= \int_0^{2\pi} \frac{h(\gamma^s(0), t) dt}{2\pi} \int_0^1 \left(\frac{\partial \varphi}{\partial x_1}(\gamma^s(u), t) \frac{d\gamma_1^s}{du}(u) + \frac{\partial \varphi}{\partial x_2}(\gamma^s(u), t) \frac{d\gamma_2^s}{du}(u) \right) du \\ &= \int_0^{2\pi} (\varphi(\gamma^s(1), t) - \varphi(\gamma^s(0), t)) \frac{\mu'_{\gamma^s(0)}(dt)}{2\pi}, \end{aligned}$$

where we used the notation $\mu'_z = h(z, t)\lambda(t)$. Since $(\gamma^s(0), t), (\gamma^s(1), \tilde{\rho}'(c_i)(t)) \in \mathbb{D} \times \mathbb{R}$ are identified in $\Sigma \times_{\tilde{\rho}'} \mathbb{R}$,

$$\varphi(\gamma^s(1), t) - \tilde{\sigma}(\gamma^s(1)) = \varphi(\gamma^s(0), \tilde{\rho}'(c_i)^{-1}(t)) - \tilde{\sigma}(\gamma^s(0)).$$

Thus,

$$\begin{aligned}
 (3-14) \quad H_0(s) + \tilde{\sigma}(\gamma^s(0)) - \tilde{\sigma}(\gamma^s(1)) &= \int_0^{2\pi} (\varphi(\gamma^s(0), \tilde{\rho}'(c_i)^{-1}(t)) - \varphi(\gamma^s(0), t)) \frac{\mu'_{\gamma^s(0)}(dt)}{2\pi} \\
 &= \int_0^{2\pi} \frac{\mu'_{\gamma^s(0)}(dt)}{2\pi} \int_t^{\tilde{\rho}'(c_i)^{-1}(t)} h(\gamma^s(0), u) du \\
 &= \frac{1}{2\pi} (\mu'_{\gamma^s(0)} \otimes \mu'_{\gamma^s(0)}(E_i^+) - \mu'_{\gamma^s(0)} \otimes \mu'_{\gamma^s(0)}(E_i^-)),
 \end{aligned}$$

where

$$E_i^+ := \{(t, u) \in [0, 2\pi] \times \mathbb{R} \mid t \leq u \leq \tilde{\rho}'(c_i)^{-1}(t)\}, \quad E_i^- := \{(t, u) \in [0, 2\pi] \times \mathbb{R} \mid \tilde{\rho}'(c_i)^{-1}(t) \leq u \leq t\}.$$

Since $\{\mu'_{\gamma^s(0)}/2\pi\}_{s \gg 0}$ can be seen as a family of probability measures on S^1 , we can extract a weakly convergent subsequence, say, $\{\mu'_{\gamma^{s_n}(0)}/2\pi\}_{n \in \mathbb{N}}$. We denote its weak limit by ν . We show that ν is $\tilde{\rho}'(c_i)$ -invariant: From the Γ -invariance of the harmonic measure as in (3-1),

$$\mu'_{\gamma^{s_n}(1)} = \tilde{\rho}'(c_i) * \mu'_{\gamma^{s_n}(0)}.$$

For any Borel subset B in $\mathbb{R}$, (3-13) implies that

$$|\mu'_{\gamma^{s_n}(1)}(B) - \mu'_{\gamma^{s_n}(0)}(B)| \leq \max\{(e^{\delta_i^{s_n}} - 1), (1 - e^{-\delta_i^{s_n}})\} \mu'_{\gamma^{s_n}(0)}(B).$$

By letting $n \rightarrow \infty$, we conclude that $\nu(B) = \tilde{\rho}'(c_i)_* \nu(B)$.

Now we apply (3-13) and (3-14) for $s = s_n$. Taking $n \rightarrow \infty$ yields

$$\lim_{n \rightarrow \infty} \tau(\widetilde{\text{hol}}_{\bar{\omega}}(c_i^{s_n})) = \frac{1}{2\pi} \lim_{n \rightarrow \infty} (H_0(s_n) + \tilde{\sigma}(\gamma^{s_n}(0)) - \tilde{\sigma}(\gamma^{s_n}(1))) = \nu \otimes \nu(E_i^+) - \nu \otimes \nu(E_i^-).$$

Since ν is $\tilde{\rho}'(c_i)$ -invariant,

$$\tau_\nu(\tilde{\rho}'(c_i)^{-1}) := \begin{cases} \nu([t, \tilde{\rho}'(c_i)^{-1}(t)]) & \text{if } t \leq \tilde{\rho}'(c_i)^{-1}(t), \\ -\nu(\tilde{\rho}'(c_i)^{-1}(t), t]) & \text{if } \tilde{\rho}'(c_i)^{-1}(t) \leq t, \end{cases}$$

is independent of the choice of $t \in \mathbb{R}$. Hence,

$$\lim_{n \rightarrow \infty} \tau(\widetilde{\text{hol}}_{\bar{\omega}}(c_i^{s_n})) = \int_0^{2\pi} \tau_\nu(\tilde{\rho}'(c_i)^{-1}) \nu(dt) = \tau_\nu(\tilde{\rho}'(c_i)^{-1}).$$

This coincides with the translation number $-\tau(\tilde{\rho}'(c_i))$, which implies (3-12). The proof is completed.

4 An application to rigidity theory

Let us prove Corollary 1.2 and Theorem 1.3 in this section. We first show the inequality (1-1). If $e(\rho) = 0$, then it is clear that the inequality (1-1) strictly holds. Thus let us assume that $e(\rho) \neq 0$. In that case, it is well known that $\rho(\Gamma)$ has no finite orbit in S^1 . We can apply the construction in Step 1 and obtain a representation $\rho' : \Gamma \rightarrow \text{Homeo}_+(S^1)$ which is semiconjugate to ρ . We have $e(\rho) = e(\rho')$ as in Step 4.

We employ the continuous connection form constructed in [Step 3](#). Then, from [Theorem 1.1](#), we have

$$(4-1) \quad |e(\rho')| = \left| \frac{1}{2\pi} \int_{\Sigma} K(z) \operatorname{vol}(dz) \right| \leq \frac{1}{2\pi} \int_{\Sigma} |K(z)| \operatorname{vol}(dz) \leq \frac{1}{2\pi} \operatorname{vol}(\Sigma) = |e(\Sigma)|,$$

which shows the Milnor–Wood type inequality (1-1).

Next, we show the rigidity statement of [Corollary 1.2](#), and [Theorem 1.3](#). For simplicity, we consider the case where $e(\rho) = e(\Sigma)$. The proof is similar in the other case where $e(\rho) = -e(\Sigma)$. Since the equality holds in (4-1), we have $K(z) = -1$ for a.e. z . We fix $z_0 \in \Sigma$ where $K(z_0) = -1$.

We use the notations in the proof of [Claim 3.9](#). Recall that two inequalities, the isoperimetric inequality and Harnack's inequality, are involved in this proof:

$$\begin{aligned} K(z) &= -\frac{(1-|z|^2)^2}{4\pi} \int_0^{2\pi} \frac{1}{2} \left(-\frac{\partial \omega_1}{\partial \theta}(z, \theta) \omega_2(z, \theta) + \frac{\partial \omega_2}{\partial \theta}(z, \theta) \omega_1(z, \theta) \right) d\theta \\ &\geq -\frac{(1-|z|^2)^2}{4\pi} \cdot \frac{1}{4\pi} \left(\int_0^{2\pi} \sqrt{\left(\frac{\partial \log h}{\partial x_1} \right)^2 + \left(\frac{\partial \log h}{\partial x_2} \right)^2} d\theta \right)^2 \\ &= -\frac{1}{4\pi^2} \left(\int_0^{2\pi} |d \log h|_{\text{hyp}} d\theta \right)^2 \\ &\geq -1. \end{aligned}$$

From $K(z_0) = -1$, the equality holds in the isoperimetric inequality for the Lipschitz curve given by $(\omega_1(z_0, \theta), \omega_2(z_0, \theta))$ and this curve must be a round circle of radius $2/(1-|z_0|^2)$ that is positively oriented. Also, the equality holds in Harnack's inequality for a.e. θ and the curve must have a constant speed. Therefore, for some $\alpha(z_0) \in \mathbb{R}$, we have

$$(h_{x_1}, h_{x_2})(z_0, \tau(z_0, \theta)) = \frac{2h(z_0, \tau(z_0, \theta))}{1-|z_0|^2} (\cos(\theta + \alpha(z_0)), \sin(\theta + \alpha(z_0))),$$

and hence

$$\frac{h_{x_1}(z_0, t) + i h_{x_2}(z_0, t)}{|h_{x_1}(z_0, t) + i h_{x_2}(z_0, t)|} = e^{i(\varphi(z_0, t) + \alpha(z_0))}.$$

From this and [Lemma 2.3](#), we have shown that the harmonic measure is of the form $h(z, t) \operatorname{vol}(z) \lambda(t)$, where

$$(4-2) \quad h(z, t) = \frac{1-|z|^2}{|m(e^{it}) - z|^2}$$

for a.e. t with

$$m(e^{it}) := \frac{e^{i(\varphi(z_0, t) + \alpha(z_0))} + z_0}{1 + \bar{z}_0 e^{i(\varphi(z_0, t) + \alpha(z_0))}} \in S^1.$$

Note that, by (4-2), for a.e. t , $m(e^{it})$ is the unique point in S^1 such that $h(z, t) \rightarrow +\infty$ as $z \rightarrow m(e^{it})$, which is independent of the choice of z_0 .

The following claim finishes the proof of [Corollary 1.2](#) since ρ is semiconjugate to ρ' by ψ .

Claim 4.1 $m: S^1 \rightarrow S^1$ is a (ρ', ρ_0) -equivariant orientation-preserving homeomorphism, which conjugates ρ' to ρ_0 .

Proof By Lemma 3.4, φ is an orientation-preserving homeomorphism, and hence so is m .

Let us prove the equivariance of m , which follows from the invariance of $\tilde{\mu}'$ under the Γ -action on $\mathbb{D} \times \mathbb{R}$ as in the sequel: Since the harmonic measure $\tilde{\mu}' = h(z, t) \operatorname{vol}(z) \lambda(t)$ on $\mathbb{D} \times \mathbb{R}$ is the pull-back of a measure on M' , we have, for every $\gamma \in \Gamma$, $\gamma_* \tilde{\mu}' = \tilde{\mu}'$. It follows that

$$(4-3) \quad h(\rho_0(\gamma)z, \tilde{\rho}'(\gamma)(t)) \operatorname{vol}(z) \rho'(\gamma)_* \lambda(t) = h(z, t) \operatorname{vol}(z) \lambda(t).$$

Then, for a.e. t , $h(\rho_0(\gamma)z, \tilde{\rho}'(\gamma)(t))$ and $h(z, t)$ differ only up to a constant which may depend on t . As $z \rightarrow m(e^{it})$, the right hand side of (4-3) diverges to $+\infty$, and hence so does the left hand side. Since $m(e^{it})$ is the unique point in S^1 such that $h(z, t) \rightarrow +\infty$ as $z \rightarrow m(e^{it})$ for a.e. t by (4-2), it follows that $\rho_0(\gamma) \circ m = m \circ \rho'(\gamma)$. $\square$

Now consider $\mathfrak{m} := m \circ \psi: S^1 \rightarrow S^1$. The map $\mathfrak{m}$ is continuous, monotone, of degree one and (ρ, ρ_0) -equivariant from Claims 3.3 and 4.1 giving a semiconjugacy from ρ to ρ_0 . Equation (4-2) completely describes the leafwise harmonic function associated with μ' on $M \times_{\rho'} S^1$ and Ψ . From $\mu'_z = \psi_* \mu_z$, it follows that

$$\begin{aligned} \mu_z(\psi^{-1}(B)) &= \int_B \frac{1 - |z|^2}{|m(e^{it}) - z|^2} \lambda(dt) = \int_B \frac{1 - |z|^2}{|m(e^{it}) - z|^2} (\psi_* \mu_0)(dt) \\ &= \int_{\psi^{-1}(B)} \frac{1 - |z|^2}{|\mathfrak{m}(e^{it}) - z|^2} q(0, t) \nu(dt) \end{aligned}$$

for any Borel set B on S^1 . Since, for any interval $I \subset \mathbb{R}$, $\tilde{\psi}^{-1}(\tilde{\psi}(I)) \setminus I$ is ν -null set, we find

$$\mu_z(B) = \int_B \frac{1 - |z|^2}{|\mathfrak{m}(e^{it}) - z|^2} q(0, t) \nu(dt)$$

for any open set, hence, any Borel set B on S^1 . We have therefore obtained the description of the leafwise harmonic function associated with μ on $M \times_{\rho} S^1$,

$$q(z, t) = q(0, t) \frac{1 - |z|^2}{|\mathfrak{m}(e^{it}) - z|^2}$$

for ν -a.e. t . Replacing ν with $q(0, t)\nu$, we complete the proof of Theorem 1.3.

References

- [1] **M Adachi**, *A differential-geometric approach to the Milnor–Wood inequality*, master’s thesis, Nagoya University (2010) In Japanese.
- [2] **M Adachi**, *A local expression of the Diederich–Fornaess exponent and the exponent of conformal harmonic measures*, Bull. Braz. Math. Soc. 46 (2015) 65–79 [MR](#) [Zbl](#)

- [3] **L V Ahlfors**, *Complex analysis: an introduction to the theory of analytic functions of one complex variable*, 3rd edition, McGraw-Hill, New York (1978) [MR](#) [Zbl](#)
- [4] **S Alvarez**, *Discretization of harmonic measures for foliated bundles*, C. R. Math. Acad. Sci. Paris 350 (2012) 621–626 [MR](#) [Zbl](#)
- [5] **G Besson, G Courtois, S Gallot**, *Entropies et rigidités des espaces localement symétriques de courbure strictement négative*, Geom. Funct. Anal. 5 (1995) 731–799 [MR](#) [Zbl](#)
- [6] **M Bucher, R Frigerio, T Hartnick**, *A note on semi-conjugacy for circle actions*, Enseign. Math. 62 (2016) 317–360 [MR](#) [Zbl](#)
- [7] **M Burger, A Iozzi, A Wienhard**, *Higher Teichmüller spaces: from $SL(2, \mathbb{R})$ to other Lie groups*, from “Handbook of Teichmüller theory, IV”, IRMA Lect. Math. Theor. Phys. 19, Eur. Math. Soc., Zürich (2014) 539–618 [MR](#) [Zbl](#)
- [8] **D Calegari**, *Foliations and the geometry of 3-manifolds*, Oxford Univ. Press (2007) [MR](#) [Zbl](#)
- [9] **A Candel**, *The harmonic measures of Lucy Garnett*, Adv. Math. 176 (2003) 187–247 [MR](#) [Zbl](#)
- [10] **B Deroin, V Kleptsyn**, *Random conformal dynamical systems*, Geom. Funct. Anal. 17 (2007) 1043–1105 [MR](#) [Zbl](#)
- [11] **H Federer**, *Geometric measure theory*, Grundle. Math. Wissen. 153, Springer (1969) [MR](#) [Zbl](#)
- [12] **S Frankel**, *Harmonic analysis of surface group representations to $\text{Diff}(S^1)$ and Milnor type inequalities*, preprint 1125, École Polytechnique (1996)
- [13] **L Garnett**, *Foliations, the ergodic theorem and Brownian motion*, J. Funct. Anal. 51 (1983) 285–311 [MR](#) [Zbl](#)
- [14] **É Ghys**, *Groupes d’homéomorphismes du cercle et cohomologie bornée*, from “The Lefschetz centennial conference, III”, Contemp. Math. 58, Amer. Math. Soc., Providence, RI (1987) 81–106 [MR](#) [Zbl](#)
- [15] **P Li**, *Uniqueness of L^1 solutions for the Laplace equation and the heat equation on Riemannian manifolds*, J. Differential Geom. 20 (1984) 447–457 [MR](#) [Zbl](#)
- [16] **P Li, R Schoen**, *L^p and mean value properties of subharmonic functions on Riemannian manifolds*, Acta Math. 153 (1984) 279–301 [MR](#) [Zbl](#)
- [17] **S Matsumoto**, *Some remarks on foliated S^1 bundles*, Invent. Math. 90 (1987) 343–358 [MR](#) [Zbl](#)
- [18] **S Matsumoto**, *The dichotomy of harmonic measures of compact hyperbolic laminations*, Tohoku Math. J. 64 (2012) 569–592 [MR](#) [Zbl](#)
- [19] **S Matsumoto**, *Rigidity theorem for surface group actions on the circle*, preprint (2021)
- [20] **J Milnor**, *On the existence of a connection with curvature zero*, Comment. Math. Helv. 32 (1958) 215–223 [MR](#) [Zbl](#)
- [21] **A Navas**, *Groups of circle diffeomorphisms*, Univ. Chicago Press (2011) [MR](#) [Zbl](#)
- [22] **T Tao**, *An introduction to measure theory*, Graduate Studies in Math. 126, Amer. Math. Soc., Providence, RI (2011) [MR](#) [Zbl](#)
- [23] **W P Thurston**, *Three-manifolds, foliations and circles, II: The transverse asymptotic geometry of foliations*, preprint (1998) Reprinted as pages 413–450 in his “Collected works with commentary, I”, Amer. Math. Soc., Providence, RI (2022)
- [24] **M Tsuji**, *Potential theory in modern function theory*, Maruzen, Tokyo (1959) [MR](#) [Zbl](#)

- [25] **J W Wood**, *Bundles with totally disconnected structure group*, Comment. Math. Helv. 46 (1971) 257–273
[MR](#) [Zbl](#)
- [26] **C B Yue**, *Brownian motion on Anosov foliations and manifolds of negative curvature*, J. Differential Geom. 41 (1995) 159–183 [MR](#) [Zbl](#)

*Department of Mathematics, Shizuoka University
Shizuoka, Japan*

*Department of Mathematical Sciences, Aoyama Gakuin University
Sagamihara, Japan*

*Department of Mathematical Sciences, Ritsumeikan University
Kusatsu, Japan*

adachi.masanori@shizuoka.ac.jp, ymatsuda@math.aoyama.ac.jp,
hnozawa@fc.ritsumei.ac.jp

Received: 1 October 2023 Revised: 15 April 2024

Finite groups of untwisted outer automorphisms of RAAGs

COREY BREGMAN

RUTH CHARNEY

KAREN VOGTMANN

For any right-angled Artin group A_Γ , Charney, Stambaugh and Vogtmann showed that the subgroup $U^0(A_\Gamma) \leq \text{Out}(A_\Gamma)$ generated by Whitehead automorphisms and inversions acts properly and cocompactly on a contractible space K_Γ . We show that any finite subgroup of $U^0(A_\Gamma)$ fixes a point of K_Γ . This generalizes the fact that any finite subgroup of $\text{Out}(F_n)$ fixes a point of outer space, and implies that there are only finitely many conjugacy classes of finite subgroups in $U^0(A_\Gamma)$.

20F28, 20F36, 20F65

1 Introduction

A right-angled Artin group (RAAG) is a finitely generated group whose only defining relations are that some of the generators commute. This can be encoded by forming a finite simplicial graph Γ with one vertex for each generator and an edge between each pair of commuting generators; the associated RAAG is then called A_Γ . The extreme examples are the free group F_n (if Γ has no edges) and the free abelian group $\mathbb{Z}^n$ (if Γ is a complete graph). We are interested in studying finite subgroups of the group $\text{Out}(A_\Gamma)$ of outer automorphisms of A_Γ .

For $A_\Gamma = \mathbb{Z}^n$, it follows from the classical Jordan–Zassenhaus theorem that there are only finitely many conjugacy classes of finite subgroups in $\text{Out}(A_\Gamma) = \text{GL}(n, \mathbb{Z})$ (see eg Curtis and Reiner [6]). Since $\text{GL}(n, \mathbb{Z})$ acts on the symmetric space $\text{GL}(n, \mathbb{R})/O(n)$ preserving a CAT(0) metric, any finite subgroup fixes a point. Since $\text{GL}(n, \mathbb{R})/O(n)$ can be identified with the space of marked lattices $\Lambda \subset \mathbb{R}^n$, where a *marking* is a choice of basis B , which gives an isomorphism $\Lambda \cong \mathbb{Z}^n$, it follows that any finite subgroup $G < \text{GL}(n, \mathbb{Z})$ acts by isometries on a lattice Λ . Equivalently, any finite subgroup $G < \text{GL}(n, \mathbb{Z})$ can be embedded in the isometry group of a flat torus T , so that the induced action on π_1 agrees with G .

For $A_\Gamma = F_n$, there is a realization theorem that says any finite subgroup G of $\text{Out}(F_n)$ can be realized as automorphisms of a finite graph X ; see Culler [5], Khramtsov [10] and Zimmermann [15]. This means one can *mark* the graph by an isomorphism $\pi_1(X) \cong F_n$ so that automorphisms of X induce the elements of G on π_1 . Furthermore, one may assume that all vertices of X have valence at least three. Since there

are only finitely many such graphs, this implies that there are only finitely many conjugacy classes of finite subgroups of $\text{Out}(F_n)$. Thus one can study finite subgroups of $\text{Out}(F_n)$ by studying symmetries of such graphs (see eg Levitt and Nicolas [12] and Smillie and Vogtmann [14]). An equivalent way to state the realization theorem is that the action of the finite group $G \leq \text{Out}(F_n)$ on outer space CV_n has a fixed point.

In previous work we constructed an *outer space* $\mathbb{O}_\Gamma$ for an arbitrary RAAG A_Γ that combines features of both CV_n and symmetric spaces [1]. The group $\text{Out}(A_\Gamma)$ acts on $\mathbb{O}_\Gamma$ with finite stabilizers, and it is proved in [1] that $\mathbb{O}_\Gamma$ is contractible. The group $\text{Out}(A_\Gamma)$ contains a natural *untwisted* subgroup $U^0(A_\Gamma)$, which is the whole group in some cases, including when $A_\Gamma = F_n$. The results in [1] build on the fact that $\mathbb{O}_\Gamma$ contains a subspace K_Γ on which the subgroup $U^0(A_\Gamma)$ acts with compact quotient. The space K_Γ was first defined by Charney, Stambaugh and Vogtmann [4], who proved it to be contractible. Points in K_Γ are special types of cube complexes called Γ -complexes, with special types of markings called *untwisted markings*. We prove the following theorem.

Theorem 8.1 *Let Γ be a simplicial graph, G a finite group and $\rho: G \rightarrow U^0(A_\Gamma)$ a homomorphism. Then there is a Γ -complex X with an untwisted marking $h: X \rightarrow \mathbb{S}_\Gamma$ on which ρ is realized by isometries.*

The following corollary is immediate.

Corollary 8.2 *Any finite subgroup of $U^0(A_\Gamma)$ has a fixed point in K_Γ (and therefore in $\mathbb{O}_\Gamma$).*

We conjecture that the entire fixed-point set is contractible, ie that K_Γ is an $\underline{\text{EG}}$ for $G = U^0(A_\Gamma)$.

Corollary 8.2 is a necessary first step towards this goal.

It is easy to see that there are only a finite number of combinatorial types of Γ -complexes, generalizing the fact that there are only a finite number of combinatorial types of graphs in CV_n . This gives us the following information about finite subgroups of $U^0(A_\Gamma)$.

Corollary 8.3 *The group $U^0(A_\Gamma)$ contains only finitely many conjugacy classes of finite subgroups.*

Extending these theorems to all of $\text{Out}(A_\Gamma)$ presents subtle difficulties that we do not address in this paper. Among these is the problem of including outer automorphisms of A_Γ that are induced by graph automorphisms of Γ . More serious is the fact that the full group $\text{Out}(A_\Gamma)$ may contain finite subgroups of $\text{GL}(n, \mathbb{Z})$ which do not preserve any Γ -complex structure, so that understanding these will require additional techniques involving the action of $\text{GL}(n, \mathbb{Z})$ on the symmetric space $\text{GL}(n, \mathbb{R})/O(n)$.

To prove Theorem 8.1, we use an inductive approach which starts from the realization theorem for $\text{Out}(F_n)$. This was inspired by work of Hensel and Kielak [8], who proved that a finite subgroup G of $U^0(A_\Gamma)$ can be realized on some cube complex, but it is not clear whether this can be taken to be a Γ -complex. We borrow a number of ideas from [8], but our proof is shorter. In particular, much of our proof is independent of the specific group G being considered, depending rather on the combinatorial structure of the defining graph Γ .

Structure of the paper and outline of the proof

In Sections 2 and 3, we review the group $U^0(A_\Gamma)$ and the definition and basic properties of Γ -complexes. The strategy of the proof is to build a marked Γ -complex realizing a finite $G < U^0(A_\Gamma)$ by gluing together marked Δ -complexes for certain subgraphs $\Delta \subseteq \Gamma$. The subgraphs we use are those whose associated special subgroup A_Δ is invariant (up to conjugacy) under $U^0(A_\Gamma)$, which we will call *U^0 -invariant subgraphs*. The argument is inductive, and the induction parameter is the *chain length* of Γ , ie the longest length of a chain of U^0 -invariant subgraphs contained in Γ . In Section 4, we study U^0 -invariant subgraphs Δ , show there is a restriction homomorphism $r_\Delta : U^0(A_\Gamma) \rightarrow U^0(A_\Delta)$, and show that minimal U^0 -invariant subgraphs are discrete, providing a base case for our induction.

In Section 5, we show that a marked Γ -complex that realizes a finite subgroup $G < U^0(A_\Gamma)$ contains a subcomplex associated to each U^0 -invariant subgraph Δ with empty link, and that the restriction of the marking to this subcomplex realizes the restriction of G to $U^0(A_\Delta)$. In Section 6, we address the opposite problem, establishing a necessary condition for extending a Δ -complex realizing the restriction of G to a Γ -complex realizing G .

In Section 7, we show how to build marked Γ -complexes when Γ is a simplicial join or a disjoint union of subgraphs Δ for which we already have marked Δ -complexes. We also show that, if Γ is a join or disjoint union, and one can realize the restriction of a finite subgroup $G < U^0(A_\Gamma)$ on each component, then one can realize all of G .

Finally, in Section 8, we induct on the length of a maximal chain of $U^0(A_\Gamma)$ -invariant subgraphs to construct a marked Γ -complex that realizes G .

Acknowledgements

We would like to thank Dawid Kielak for several helpful discussions and the referee for giving thoughtful comments. Bregman is supported by NSF grant DMS-2052801.

2 Review of RAAGs and the untwisted subgroup of $\text{Out}(A_\Gamma)$

Let Γ be a finite simplicial graph. The *right-angled Artin group* (RAAG) A_Γ is the group generated by the vertices V of Γ with defining relations given by declaring that adjacent vertices commute.

By a subgraph of Γ we will always mean a full (induced) subgraph, unless otherwise specified. Given a subgraph $\Delta \subseteq \Gamma$, we write $x \in \Delta$ if x is a vertex of Δ .

For $x \in \Gamma$, the *link* $\text{lk}(x)$ is the subgraph spanned by vertices adjacent to x . The link of a subgraph $\Delta \subset \Gamma$ is the intersection of the links of all vertices of Δ . The *double link* $\text{dlk}(x)$ is the link of $\Delta = \text{lk}(x)$. The *star* $\text{st}(x)$ is the subgraph spanned by x and $\text{lk}(x)$.

Recall from [4] that a Γ -*Whitehead partition* $\mathcal{P}$ based at $x \in \Gamma$ is a partition of $V^\pm = V \cup V^{-1}$ into three sets $\text{lk}^\pm(x)$, P_1 and P_2 satisfying certain conditions. The sets P_1 and P_2 are called the *sides* of $\mathcal{P}$. A

Γ -Whitehead partition can be most easily described using the *double* $\Gamma^\pm$ of Γ , where the vertices of $\Gamma^\pm$ are $V^\pm$ and two vertices are joined by an edge if they commute but are not inverses of each other. If $\mathcal{P}$ is based at x , $\text{lk}^\pm(x)$ consists of all vertices adjacent to x in $\Gamma^\pm$, and each of P_1 and P_2 is a union of (the vertices in) some connected components of $\Gamma^\pm \setminus \text{lk}^\pm(x)$. Furthermore, we require x and x^{-1} to be in different sides of $\mathcal{P}$, and each side must contain at least one additional element. We will abbreviate Γ -Whitehead partition to simply Γ -partition.

A vertex $y \in V$ is *split* by a Γ -partition $\mathcal{P}$ if y is in one side and y^{-1} is in the other. If y is split by $\mathcal{P}$, then y and y^{-1} must lie in different components of $\Gamma^\pm \setminus \text{lk}^\pm(x)$; hence, $\text{lk}(y) \subseteq \text{lk}(x)$.

A Γ -partition $\mathcal{P}$ based at x determines a *Whitehead automorphism* $\varphi(\mathcal{P}, x)$, defined as follows. Let P_i be the side of $\mathcal{P}$ containing x . If $\mathcal{P}$ splits y , then $\varphi(\mathcal{P}, x)$ sends $y \mapsto yx^{-1}$ if $y \in P_i$, and $y \mapsto xy$ if $y^{-1} \in P_i$. If both y and y^{-1} are in P_i , then $\varphi(\mathcal{P}, x)(y) = xyx^{-1}$. For all other y , $\varphi(\mathcal{P}, x)(y) = y$. The simplest Whitehead automorphisms are the *folds* sending $y \mapsto yx^{-1}$ or $y \mapsto xy$ for some x and y (and fixing all generators other than y), and the *partial conjugations* sending $y \mapsto xyx^{-1}$ for all y in some component C of $\Gamma \setminus \text{lk}(x)$. These correspond to partitions $\mathcal{P} = (\text{lk}^\pm(x) \mid P_1 \mid P_2)$ with $P_1 = \{x, y\}$ or $\{x, y^{-1}\}$ (for a fold) or $P_1 = \{x, C^\pm\}$ (for a partial conjugation). Every Whitehead automorphism is a product of folds and partial conjugations.

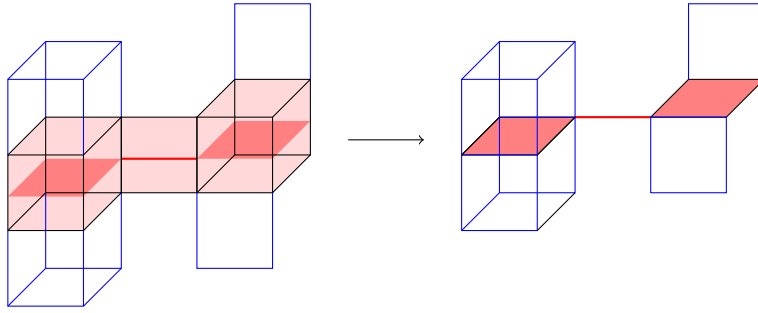
The subgroup of $\text{Out}(A_\Gamma)$ generated by Whitehead automorphisms and by inversions of the generators is denoted by $U^0(A_\Gamma)$. If $A_\Gamma = F_n$, this is the whole group, ie $U^0(F_n) = \text{Out}(F_n)$. If $A_\Gamma = \mathbb{Z}^n$, there are no Whitehead automorphisms, and $\text{Out}(\mathbb{Z}^n) = \text{GL}(n, \mathbb{Z})$ is generated by inversions and *twists*, where a twist sends a generator y to xy for some y with $\text{st}(y) \subseteq \text{st}(x)$ and fixes all other generators. By a theorem of Laurence and Servatius [11; 13], for a general RAAG the group $\text{Out}(A_\Gamma)$ is generated by Whitehead automorphisms, inversions, twists and automorphisms of Γ .

The subgroup generated by $U^0(A_\Gamma)$ and graph automorphisms was called the *untwisted subgroup* and denoted by $U(A_\Gamma)$ in [4; 1].

3 Blowups and Γ -complexes

Let $\mathcal{P} = (\text{lk}^\pm(x) \mid P_1 \mid P_2)$ be a Γ -partition based at x . If $\text{lk}(x) = \text{lk}(y)$ and $\mathcal{P}$ splits y , then y can also serve as a base for $\mathcal{P}$. Specifying a choice of base specifies the corresponding Whitehead automorphism, but we will often use Γ -partitions without specifying a base, in which case we write $\mathcal{P} = (\text{lk}(\mathcal{P}) \mid P_1 \mid P_2)$.

We say Γ -partitions $\mathcal{P}, \mathcal{Q}$ are *adjacent* if some (hence any) base of $\mathcal{P}$ commutes with some (any) base of $\mathcal{Q}$, and they are *compatible* if either they are adjacent or some side of $\mathcal{Q}$ is disjoint from some side of $\mathcal{P}$. A collection Π of Γ -partitions is a *compatible collection* if its elements are distinct and pairwise compatible. In [4], the authors constructed a labeled cube complex $\mathbb{S}_\Gamma^\Pi$ called a *blowup* from a compatible collection Π of Γ -partitions. The underlying (unlabeled) cube complex is called a Γ -complex. In this section we review some facts about special cube complexes and Γ -complexes that we will need.

Figure 1: Hyperplane collapse $X \rightarrow X//H$.

3.1 Special cube complexes, collapsing and duplicating hyperplanes

Recall that a cube complex is called a *special cube complex* if it is locally CAT(0) and has no hyperplanes that selfintersect or are one-sided, selfosculating or interosculating. We refer to the original article by Haglund and Wise [7] for the basic definitions.

Let X be a special cube complex. If H is a hyperplane in X , the *collapse map* $c: X \rightarrow X//H$ collapses the carrier $\kappa(H)$ of H orthogonally onto H . We say the result $X//H$ is obtained from X by a *hyperplane collapse* (see Figure 1).

The edges that intersect a hyperplane H are said to be *dual* to H , and by an *orientation* on H we mean a consistent choice of orientation of the edges dual to H .

If $\mathcal{S}$ is a collection of hyperplanes, we write $X//\mathcal{S}$ for the space obtained by collapsing all hyperplanes in $\mathcal{S}$ (in any order). The collection $\mathcal{S}$ is *acyclic* if the collapse map $X \rightarrow X//\mathcal{S}$ is a homotopy equivalence.

If H is a hyperplane in X with carrier $\kappa(H)$, we can obtain a new cube complex by doubling $\kappa(H)$ (see Figure 2). We will refer to this as *duplicating the hyperplane* H . The resulting cube complex has two

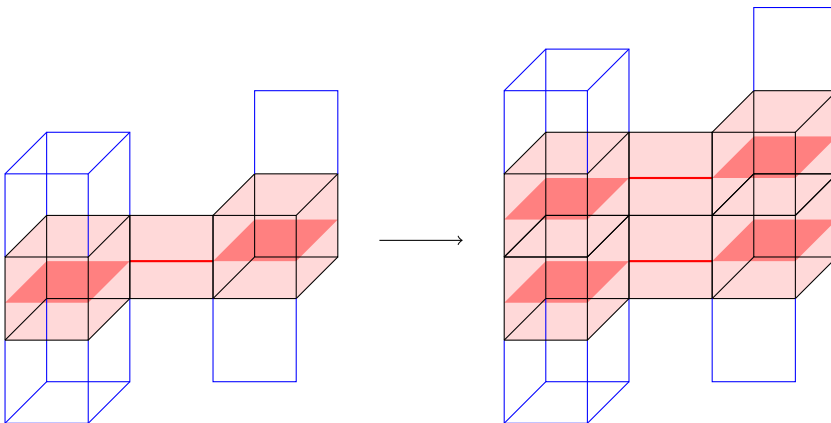


Figure 2: Duplicating a hyperplane.

new hyperplanes H' and H'' , and collapsing either recovers the original complex X . We say H' and H'' are *parallel*. A hyperplane is called a *duplicate* if it is parallel to another hyperplane.

3.2 Blowups

The blowup $\mathbb{S}_\Gamma^\Pi$ associated to a compatible collection Π of Γ -partitions is a special cube complex with no separating hyperplanes and with some extra structure.

If Γ is discrete, then $\mathbb{S}_\Gamma^\Pi$ is a finite connected graph with no separating edges or bivalent vertices, and the extra structure consists of a maximal tree T and an orientation and label on each edge in $\mathbb{S}_\Gamma^\Pi \setminus T$, where the labels are the vertices of Γ . Each edge of T corresponds to a partition in Π , determined by the labels and orientations of the edges not in T .

If Π is empty, then $\mathbb{S}_\Gamma^\Pi$ is the *Salvetti complex* $\mathbb{S}_\Gamma$ associated to A_Γ . Recall that this is a cube complex with a single 0-cell, one oriented 1-cell for each vertex of Γ and one k -torus for each k -clique in Γ . The orientations on the 1-cells, which are labeled by vertices of Γ , determine an isomorphism $\pi_1(\mathbb{S}_\Gamma) \cong A_\Gamma$, and the cubical isomorphisms of $\mathbb{S}_\Gamma$ can be identified with the automorphisms of the graph Γ .

In general, $\mathbb{S}_\Gamma^\Pi$ has

- one hyperplane $H_\mathcal{P}$ for each partition $\mathcal{P} \in \Pi$, and
- one hyperplane H_v for each vertex $v \in \Gamma$.

The hyperplanes H_v are oriented, but the hyperplanes $H_\mathcal{P}$ are not.

The set of hyperplanes labeled by partitions is acyclic, and the complex obtained by collapsing all hyperplanes in this set is isomorphic to the Salvetti complex $\mathbb{S}_\Gamma$.

Collapsing a single hyperplane labeled by $\mathcal{P} \in \Pi$ is equivalent to removing $\mathcal{P}$ from the collection Π . In particular, collapsing every hyperplane in Π other than $\mathcal{P}$ results in a *single blowup* $\mathbb{S}^\mathcal{P}$. This has exactly two vertices x_1 and x_2 , and one can recover $\mathcal{P} = (\text{lk}(\mathcal{P}) \mid P_1 \mid P_2)$ from $\mathbb{S}^\mathcal{P}$ by looking at the (oriented!) edges dual to the hyperplanes H_v : if there is only one edge dual to H_v and it terminates at x_i , then $v \in P_i$; if it originates at x_i , then $v^{-1} \in P_i$. If there are two edges dual to H_v , then $v \in \text{lk}(\mathcal{P})$. The carrier of $H_\mathcal{P}$ in $\mathbb{S}^\mathcal{P}$ is isomorphic to the product of an interval with the Salvetti for $\text{lk}(\mathcal{P})$.

3.3 Γ -complexes

Definition 3.1 A cube complex is called a Γ -*complex* if it is isomorphic to the underlying cube complex of a blowup $\mathbb{S}_\Gamma^\Pi$. A *blowup structure* on a Γ -complex X is a labeling of its hyperplanes that identifies X with a blowup $\mathbb{S}_\Gamma^\Pi$, ie hyperplanes are labeled by Γ -partitions or by vertices of Γ , and the hyperplanes labeled by vertices are oriented. A blowup structure determines a *collapse map* $c_\pi: X \rightarrow \mathbb{S}_\Gamma$ that collapses all hyperplanes labeled by partitions. If $v \in \Gamma$, a *characteristic cycle* for v is a closed edge path which crosses each hyperplane at most once, and whose image under c_π is the loop labeled v .

In general, a Γ -complex may have several different blowup structures. For example, if Γ is discrete, a Γ -complex is a graph, which may have several different maximal trees, and the remaining edges may be oriented and labeled with the vertices of Γ in any way.

Definition 3.2 A set $\mathcal{T}$ of hyperplanes in a Γ -complex is called *treelike* if collapsing $\mathcal{T}$ gives a cube complex isomorphic to $\mathbb{S}_\Gamma$.

The following proposition says that any treelike set of hyperplanes in a Γ -complex is the set of hyperplanes labeled by partitions in at least one blowup structure. The only ambiguity comes from the assignment of labels and orientations to the hyperplanes not in the treelike set, which can be permuted by any automorphism of the graph Γ . If this assignment changes, the partitions labeling the hyperplanes in the tree also change, by the same (signed) permutation of vertices.

Proposition 3.3 Let $\mathcal{T}$ be a treelike set of hyperplanes in a Γ -complex X . Then there is a compatible set of Γ -partitions Π and an isometry $X \cong \mathbb{S}_\Gamma^\Pi$, such that $\mathcal{T}$ is the set of hyperplanes associated to the partitions in Π .

Proof We recall the construction. For complete details see [4, Section 4].

Label the edges dual to each hyperplane $H \in \mathcal{T}$ by H . Choose an isomorphism of $X//\mathcal{T}$ with $\mathbb{S}_\Gamma$; this orients each hyperplane that is not in $\mathcal{T}$ and labels its dual edges by a vertex of Γ . The set of cubes in X with all edge labels in $\mathcal{T}$ forms a CAT(0) subcomplex $\mathbb{C}$ that contains all vertices of X . A hyperplane $H \in \mathcal{T}$ cuts $\mathbb{C}$ into two pieces, so partitions the vertices of X into two sets, $v_1(H)$ and $v_2(H)$. Now form a partition $(\text{lk}(H) \mid U_1 \mid U_2)$ of $V^\pm$ as follows:

- (1) If the hyperplane H_v labeled by v intersects H , then v and v^{-1} are in $\text{lk}(H)$.
- (2) If $H_v \cap H = \emptyset$ and the terminal vertex of an edge dual to H_v is in $v_i(H)$, then $v \in U_i$.
- (3) If $H_v \cap H = \emptyset$ and the initial vertex of an edge dual to H_v is in $v_i(H)$, then $v^{-1} \in U_i$.

Then the partition $(\text{lk}(H) \mid U_1 \mid U_2)$ is a Γ -partition, the set of Γ -partitions for all $H \in \mathcal{T}$ is a compatible collection Π , and X is isomorphic to $\mathbb{S}_\Gamma^\Pi$. $\square$

We note that condition (1) in the proof of [Proposition 3.3](#) is equivalent to saying that, in the universal cover $\widetilde{\mathbb{S}}_\Gamma^\Pi$, some lift of H contains an axis for v . Saying that v and v^{-1} are in different P_i is equivalent to saying that an axis for v in $\widetilde{\mathbb{S}}_\Gamma^\Pi$ intersects some lift of H transversally; in this case we say the axis *skewers* H . Saying v and v^{-1} are in the same P_i is equivalent to saying that no axis for v intersects any lift of H .

If we are given a special cube complex X which we do not know a priori is a Γ -complex, then, to prove that it is, we first need to find an acyclic collection $\mathcal{T}$ of hyperplanes which collapses to give a cube complex isomorphic to $\mathbb{S}_\Gamma$. Choosing an isomorphism $X//\mathcal{T} \cong \mathbb{S}_\Gamma$ gives a labeling and orientation on all of the remaining hyperplanes. We then need to check that each hyperplane $H \in \mathcal{T}$ determines a Γ -partition. We can do this by collapsing all hyperplanes other than H to get a complex with two vertices,

then checking whether the location of the initial and terminal vertices of edges labeled by $v \in \Gamma$ gives a valid Γ -partition. By Proposition 3.3, the partitions for one treelike set $\mathcal{T}$ are all Γ -partitions if and only if this holds for every treelike set.

3.4 Subdividing blowups

In a blowup $\mathbb{S}_\Gamma^\Pi$, no two hyperplanes are parallel, ie there are no duplicate hyperplanes. However, in the arguments that follow we will need to allow cubical subdivisions of blowups that result in duplicate hyperplanes. Duplicating $H_\mathcal{P}$ can be thought of as subdividing its carrier $\kappa(H_\mathcal{P})$, and is equivalent to adding a duplicate copy of $\mathcal{P}$ to Π . We want both of the new hyperplanes we have created to be in the treelike set since we must collapse both to recover $\mathbb{S}_\Gamma$. Subdividing the carrier of H_v is a little subtler; here we want only one of the two new hyperplanes to be added to the treelike set, so that collapsing the treelike set still gives $\mathbb{S}_\Gamma$. In other words, when we duplicate H_v , we want one of the two resulting hyperplanes to be labeled H_v , and the other to correspond to a partition. We also need the new H_v to have the orientation induced from the old H_v . This is accomplished by adding a “singleton partition” to Π ; this is a partition based at v with one side containing only v (if we want the initial segment of the dual edge to retain the v label) or v^{-1} (if we want the terminal segment to retain the v label). To make this a canonical operation, we can consistently use the singleton partition $\mathcal{S}_{v^{-1}} = (\text{lk}^\pm(v) \mid \{v^{-1}\} \mid (V \setminus \text{lk}(v))^\pm \setminus \{v^{-1}\})$, so that the terminal segment always retains the v label.

Note that duplicate partitions fit the definition of “compatible with each other”, and a singleton partition is compatible with every Γ -partition. A set of pairwise compatible partitions that is allowed to have singletons and duplicates will be called a *compatible multiset*. By the above remarks, compatible multisets correspond to subdivided blowups.

4 U^0 -invariant subgraphs

Recall that a *marking* on a Γ -complex is a homotopy equivalence $h: X \rightarrow \mathbb{S}_\Gamma$. Let G be a finite group and $\rho: G \rightarrow \text{Out}(A_\Gamma)$ a homomorphism.

Definition 4.1 A marked Γ -complex (X, h) *realizes* ρ if there is an action $f: G \rightarrow \text{Aut}(X)$ of G on X by cubical automorphisms such that $h \circ f(g) \circ h^{-1}$ induces $\rho(g)$ on $\pi_1(\mathbb{S}_\Gamma) = A_\Gamma$ for all $g \in G$.

Our goal is to build a marked Γ -complex that realizes (the inclusion of) a finite subgroup $G < U^0(A_\Gamma)$. Our approach is inductive. Specifically, we will build our Γ -complex by gluing together marked Δ -complexes for subgraphs Δ which are U^0 -invariant, in the sense that elements of $U^0(A_\Gamma)$ preserve the special subgroup A_Δ (up to conjugacy). We will induct on the length of a maximal chain of U^0 -invariant subgraphs. In this section we prepare for this by establishing some basic facts about U^0 -invariant subgraphs.

Definition 4.2 Given $\phi \in \text{Out}(A_\Gamma)$, a subgraph Δ of Γ is ϕ -invariant if $\hat{\phi}(A_\Delta)$ is conjugate to A_Δ for some lift $\hat{\phi}$ of ϕ to $\text{Aut}(A_\Gamma)$. Since any two such lifts differ by an inner automorphism, this is well defined. A subgraph Δ is U^0 -invariant if it is ϕ -invariant for every $\phi \in U^0(A_\Gamma)$.

The next lemma gives a criterion for U^0 -invariance.

Lemma 4.3 *Let Δ be a subgraph of Γ . Then Δ is U^0 -invariant if and only if the following two conditions hold for all $x, y \in \Gamma$:*

- (i) *If $x \in \Delta$ and $\text{lk}(x) \subseteq \text{lk}(y)$, then $y \in \Delta$.*
- (ii) *If Δ intersects more than one component of $\Gamma \setminus \text{st}(y)$, then $y \in \Delta$.*

Proof Since every subgraph $\Delta \subseteq \Gamma$ is invariant under inversions, a subgraph Δ is U^0 -invariant if and only if it is invariant under the remaining generators of $U^0(A_\Gamma)$, ie all folds and partial conjugations. There is a fold $\tau: x \mapsto xy$ if and only if $\text{lk}(y) \supseteq \text{lk}(x)$, so τ maps $x \in \Delta$ to A_Δ if and only if $y \in \Delta$. If Δ intersects two different components C and C' of $\Gamma \setminus \text{st}(y)$, then A_Δ is sent to a conjugate of itself under the partial conjugation $C \mapsto yCy^{-1}$ if and only if $y \in \Delta$. $\square$

Proposition 4.4 *Let Γ be a simplicial graph.*

- (1) *If Σ is a subgraph of Γ , then $\Delta = \text{lk}(\Sigma)$ is U^0 -invariant.*
- (2) *If Δ_1 and Δ_2 are two U^0 -invariant subgraphs of Γ , then $\Delta_1 \cap \Delta_2$ is U^0 -invariant.*
- (3) *If Δ_1 and Δ_2 are two U^0 -invariant subgraphs of Γ whose join $\Delta_1 * \Delta_2$ is also a subgraph, then $\Delta_1 * \Delta_2$ is U^0 -invariant.*
- (4) *If Σ is a nonsingleton connected component of a U^0 -invariant subgraph Δ , then Σ is U^0 -invariant.*
- (5) *If Δ is U^0 -invariant and $N(\Delta)$ is the subgraph spanned by Δ and all vertices adjacent to Δ , then $N(\Delta)$ is U^0 -invariant.*

Proof In each case we check conditions (i) and (ii) of Lemma 4.3:

- (1) (i) $x \in \Delta = \text{lk}(\Sigma)$ if and only if $\Sigma \subseteq \text{lk}(x)$. If $\text{lk}(x) \subseteq \text{lk}(y)$, then $\Sigma \subseteq \text{lk}(y)$, so $y \in \text{lk}(\Sigma)$.
- (ii) Suppose that $\Delta = \text{lk}(\Sigma)$ intersects two different components C_1 and C_2 of $\Gamma \setminus \text{st}(y)$, say $x_1 \in C_1 \cap \Delta$ and $x_2 \in C_2 \cap \Delta$. If $z \in \Sigma$, then x_1 and x_2 are both connected to z , so z must be in $\text{lk}(y)$. Thus, $\Sigma \subseteq \text{lk}(y)$, so $y \in \text{lk}(\Sigma)$.
- (2) (i) Let $x \in \Delta_1 \cap \Delta_2$. If $\text{lk}(y) \supseteq \text{lk}(x)$, then $y \in \Delta_1$ by invariance of Δ_1 ; similarly, $y \in \Delta_2$, so y is in the intersection.
- (ii) If $\Delta_1 \cap \Delta_2$ intersects two components of $\Gamma \setminus \text{st}(y)$, then the same is true of both Δ_1 and Δ_2 , so y is in both.
- (3) (i) If $x \in \Delta_1$ and $\text{lk}(y) \supseteq \text{lk}(x)$, then $y \in \Delta_1$, and similarly if $x \in \Delta_2$.
- (ii) Suppose $\Delta = \Delta_1 * \Delta_2$ intersects two different components C and C' of $\Gamma \setminus \text{st}(y)$. Let $x \in C \cap \Delta$ and $x' \in C' \cap \Delta$. Then x and x' must be in the same Δ_i , since otherwise there is an edge connecting them. But Δ_i is U^0 -invariant, so $y \in \Delta_i \subseteq \Delta$.
- (4) (i) Suppose $x \in \Sigma$, and $\text{lk}(y) \supset \text{lk}(x)$. Since Σ is not a singleton, $\text{lk}(x) \cap \Sigma$ contains a vertex z . Since z is in the links of both x and y , y is also in Σ .

- (ii) Suppose $x, z \in \Sigma$ are in different components of $\Gamma \setminus \text{st}(y)$. Since Δ is U^0 -invariant, this implies $y \in \Delta$, but then $\text{st}(y)$ cannot separate x from z unless $y \in \Sigma$.
- (5) (i) If $x \in N(\Delta)$, then either $x \in \Delta$ or $\text{lk}(x) \cap \Delta \neq \emptyset$. If the distance from y to Δ is at least 2, then $\text{lk}(y) \cap \Delta = \emptyset$, so $\text{lk}(x) \not\subseteq \text{lk}(y)$.
- (ii) If $y \notin N(\Delta)$, then $\Delta \cap \text{st}(y) = \emptyset$. Since Δ is U^0 -invariant, it lies in a single component of $\Gamma \setminus \text{st}(y)$, so all $x \in N(\Delta) \setminus \text{st}(y)$ must lie in the same component. $\square$

Recall from [1] that two vertices are called *fold-equivalent* if they have the same link in Γ , and we order the set of fold-equivalence classes by inclusion of their links. The following two propositions will allow us to establish the base case of our induction.

Proposition 4.5 *Let $\Delta \subseteq \Gamma$ be a minimal U^0 -invariant subgraph. Then Δ is a maximal fold-equivalence class. In particular, Δ is discrete.*

Proof If $[u]$ is a maximal fold-equivalence class, then it is easy to check using Lemma 4.3 that $[u]$ is U^0 -invariant. On the other hand, if $\Delta \subseteq \Gamma$ is any U^0 -invariant subgraph, then Δ contains a maximal equivalence class $[u]$ by condition (1) of Lemma 4.3, so $[u] = \Delta$ by the minimality of Δ . $\square$

Proposition 4.6 *Let Δ be a U^0 -invariant subgraph of Γ . Then there is a restriction homomorphism $r_\Delta: U^0(A_\Gamma) \rightarrow U^0(A_\Delta)$.*

Proof Let ϕ be an element of $U^0(A_\Gamma)$. Since Δ is U^0 -invariant, there is a lift $\hat{\phi}$ of ϕ to $\text{Aut}^0(A_\Gamma)$ with $\hat{\phi}(A_\Delta) = A_\Delta$. Define $r_\Delta(\phi)$ to be the image in $\text{Out}(A_\Delta)$ of the restriction of $\hat{\phi}$ to A_Δ .

To check that r_Δ is well defined, suppose $\hat{\phi}'$ is another lift of ϕ sending A_Δ to itself. Then $\hat{\phi}' = \iota_g \circ \hat{\phi}$, where ι_g is conjugation by some $g \in A_\Gamma$ that normalizes A_Δ . By [3, Lemma 2.2], the normalizer of A_Δ is $A_\Delta \times A_{\text{lk}(\Delta)}$. Since elements of $A_{\text{lk}(\Delta)}$ act trivially by conjugation, $\hat{\phi}' = \iota_h \circ \hat{\phi}$ for some $h \in A_\Delta$, ie the images of $\hat{\phi}$ and $\hat{\phi}'$ in $\text{Out}(A_\Delta)$ are equal. Moreover, if ϕ_1 and ϕ_2 are two elements of $U^0(A_\Gamma)$ and $\phi = \phi_1 \circ \phi_2$, then $\hat{\phi} = \hat{\phi}_1 \circ \hat{\phi}_2$ is a lift of ϕ which preserves A_Δ . Thus, $r_\Delta: U^0(A_\Gamma) \rightarrow \text{Out}(A_\Delta)$ is a homomorphism.

To see that this lands in $U^0(A_\Delta)$, we check that this is the case for the generators of $U^0(A_\Gamma)$. Let ϕ be a generator that lifts to a fold $\hat{\phi}: x \mapsto xy$, so $\text{lk}(x) \subset \text{lk}(y)$. If $x \notin \Delta$, then $\hat{\phi}$ restricts to the identity on A_Δ . If $x \in \Delta$, then U^0 -invariance of Δ implies that y is also in Δ , so the restriction is the lift of a fold in $U^0(A_\Delta)$. Now suppose $\hat{\phi}$ is a partial conjugation by x . If $x \notin \Delta$, then the fact that Δ is U^0 -invariant implies that Δ lies entirely in one component of $\Gamma \setminus \text{st}_\Gamma(x)$, so the action of $\hat{\phi}$ to A_Δ is trivial. If $x \in \Delta$, then $\text{st}_\Delta(x) \subseteq \text{st}_\Gamma(x)$, so the components of $\Delta \setminus \text{st}_\Delta(x)$ are contained in components of $\Gamma \setminus \text{st}_\Gamma(x)$. Thus, $\hat{\phi}$ restricts to a (product of) partial conjugation(s) by x on A_Δ . $\square$

Definition 4.7 If Δ is a U^0 -invariant subgraph and $f: H \rightarrow U^0(A_\Gamma)$ is any homomorphism, we call $f_\Delta = r_\Delta \circ f: H \rightarrow U^0(A_\Delta)$ the *restriction* of f to Δ .

5 U^0 -invariant subcomplexes of marked Γ -complexes

Throughout this section we assume $\Delta \subseteq \Gamma$ is a U^0 -invariant subgraph with $\text{lk}(\Delta) = \emptyset$ and (X, h) is a Γ -complex with an untwisted marking. This means that, for any blowup structure $\mathbb{S}_\Gamma^\Pi$ on X with associated collapse map $c_\pi: X \rightarrow \mathbb{S}_\Gamma$, the composition $c_\pi h^{-1}$ induces an untwisted automorphism of A_Γ , that is, an element of $U(A_\Gamma)$. The aim is to identify a subcomplex $X_\Delta \subset X$ which is invariant under the action of isometries that induce elements of $U^0(A_\Gamma)$. The reason for the restriction that $\text{lk}(\Delta) = \emptyset$ is that in this case the subcomplex X_Δ is unique. [Remark 5.8](#) discusses the general case.

The following lemma deals with the discrepancy between the subgroups $U(A_\Gamma)$ and $U^0(A_\Gamma)$.

Lemma 5.1 *If X is a Γ -complex with an untwisted marking h , then X has a blowup structure $\mathbb{S}_\Gamma^\Pi$ with collapse map $c_\pi: X = \mathbb{S}_\Gamma^\Pi \rightarrow \mathbb{S}_\Gamma$ such that $(c_\pi h^{-1})_* \in U^0(A_\Gamma)$.*

Proof $U^0(A_\Gamma)$ is normal in the untwisted subgroup $U(A_\Gamma)$ of $\text{Out}(A_\Gamma)$, and the quotient Q is a subgroup of $\text{Aut}(\Gamma)$, the group of graph automorphisms of Γ . The short exact sequence

$$1 \rightarrow U^0(A_\Gamma) \rightarrow U(A_\Gamma) \rightarrow Q \rightarrow 1$$

splits, so $U(A_\Gamma) = U^0(A_\Gamma) \rtimes Q$.

Let $\mathbb{S}_\Gamma^\Pi$ be any blowup structure on X such that $(c_\pi h^{-1})_*$ is untwisted. By the above observation, if $(c_\pi h^{-1})_*$ is not in $U^0(A_\Gamma)$, we can compose it with a graph automorphism α to produce an element of $U^0(A_\Gamma)$. Realize α by an isometry f_α of $\mathbb{S}_\Gamma$. Composing c_π with f_α is equivalent to changing the labels and orientations of the vertex-labeled hyperplanes in $\mathbb{S}_\Gamma^\Pi$ by α ; this changes the partitions by the same relabeling, giving a new set of partitions $\alpha\Pi$. In other words, this gives a blowup structure $\mathbb{S}_\Gamma^{\alpha\Pi}$ on X such that $(h^{-1}c_{\alpha\Pi})_* \in U^0(A_\Gamma)$. $\square$

Now fix a blowup structure $\mathbb{S}_\Gamma^\Pi$ on X such that $(c_\pi h^{-1})_* \in U^0(A_\Gamma)$. In any Γ -partition, all bases have the same link. Since Δ is U^0 -invariant, [Lemma 4.3](#) implies that either all bases of a partition in Π are in Δ , or none are. So we may write $\Pi = \{\mathfrak{D}_1, \dots, \mathfrak{D}_k, \mathcal{P}_1, \dots, \mathcal{P}_l\}$, where the $\mathfrak{D}_i$ are based in Δ and the $\mathcal{P}_j$ are based in $\Gamma \setminus \Delta$.

Lemma 5.2 *Each $\mathcal{P}_i$ has a unique side P_i^Δ such that $\Delta^\pm \subset P_i^\Delta \cup \text{lk}(\mathcal{P})$ and $\Delta^\pm \cap P_i^\Delta \neq \emptyset$. Furthermore, if $\mathcal{P}_i$ and $\mathcal{P}_j$ are not adjacent, then $P_i^\Delta \cap P_j^\Delta \neq \emptyset$.*

Proof Let $y_i \in \Gamma \setminus \Delta$ be a base for $\mathcal{P}_i$. Recall that we have assumed $\text{lk}(\Delta)$ is empty; this implies that $\Delta \setminus \text{st}(y_i)$ is nonempty. Since Δ is U^0 -invariant, it intersects at most one component of $\Gamma \setminus \text{st}(y_i)$. Each side of $\mathcal{P}_i$ is a union of components of $\Gamma \setminus \text{st}(y_j)$. Thus, there is a unique side P_i^Δ such that $\Delta^\pm \cap P_i^\Delta \neq \emptyset$ and $\Delta^\pm \subseteq P_i^\Delta \cup \text{lk}(\mathcal{P}_i)^\pm$.

If $\mathcal{P}_i$ and $\mathcal{P}_j$ are compatible but not adjacent and $P_i^\Delta \cap P_j^\Delta = \emptyset$, then $P_i^\Delta \cap \text{lk}(\mathcal{P}_j) = P_j^\Delta \cap \text{lk}(\mathcal{P}_i) = \emptyset$ as well (see [\[1, Lemma 2.9\]](#)), forcing $\Delta^\pm \subset \text{lk}(\mathcal{P}_i) \cap \text{lk}(\mathcal{P}_j)$. This contradicts our assumption that $\Delta^\pm \cap P_i^\Delta \neq \emptyset$. $\square$

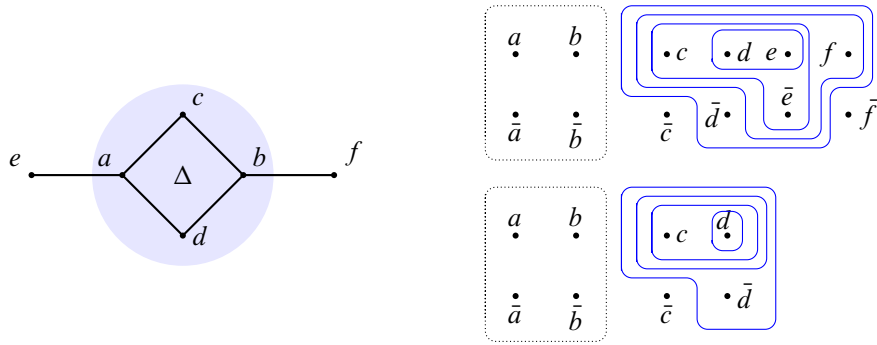


Figure 3: Left: a graph Γ and U^0 -invariant subgraph Δ . Right: four different Γ -partitions with the same link and their restrictions to $\Delta^\pm$. In the restriction, the innermost and outermost partitions become trivial and the middle two become equal.

Recall from [4] that vertices of $\mathbb{S}_\Gamma^\Pi$ correspond to collections $\{Q_1^\times, \dots, Q_k^\times, P_1^\times, \dots, P_l^\times\}$, where the superscript $\times$ indicates a choice of side. Each pair of sides in the collection must be *consistent*, which means either they intersect nontrivially or their bases commute. We define $\mathbb{K}_\Delta$ to be the subcomplex of $\mathbb{S}_\Gamma^\Pi$ consisting of vertices of the form $\{Q_1^\times, \dots, Q_k^\times, P_1^\Delta, \dots, P_l^\Delta\}$, edges obtained by switching a side of some $\mathfrak{Q}_i$ or labeled by some $v \in \Delta$, and all higher-dimensional cubes spanned by these edges.

Proposition 5.3 *The subcomplex $\mathbb{K}_\Delta$ is a subdivided blowup of $\mathbb{S}_\Delta$.*

Proof Let $\{\mathfrak{Q}_1, \dots, \mathfrak{Q}_k\}$ be the partitions of Π based in Δ , let Ω be the multiset of Δ -partitions obtained by intersecting the $\mathfrak{Q}_i$ with $\Delta^\pm$ (see Figure 3), and let $\mathbb{S}_\Delta^\Omega$ be the corresponding subdivided blowup. If the bases of $\mathfrak{Q}_i$ and $\mathfrak{Q}_j$ don't commute, then exactly one pair of sides has empty intersection (see [4, Lemma 3.6]). The $\mathfrak{Q}_i \cap \Delta^\pm$ are Δ -partitions, and the corresponding pair of sides still has empty intersection. This means that they are compatible, and all other pairs of sides must intersect. Since the $\mathfrak{Q}_i$ are based in Δ , this means that $\{Q_1^\times, \dots, Q_k^\times\}$ is consistent in $\Gamma^\pm$ if and only if $\{Q_1^\times \cap \Delta^\pm, \dots, Q_k^\times \cap \Delta^\pm\}$ is consistent in $\Delta^\pm$, ie defines a vertex of $\mathbb{S}_\Delta^\Omega$. Note that, for each $j = 1, \dots, k$, either $\mathfrak{P}_i$ is adjacent to $\mathfrak{Q}_j$ or P_i^Δ intersects both sides of $\mathfrak{Q}_j$ nontrivially, since $\mathfrak{Q}_j$ splits its own base, which is in $\Delta \setminus \text{lk}(\mathfrak{P}_i)$. Therefore, the map sending

$$\{Q_1^\times \cap \Delta^\pm, \dots, Q_k^\times \cap \Delta^\pm\} \rightarrow \{Q_1^\times, \dots, Q_k^\times, P_1^\Delta, \dots, P_l^\Delta\}$$

is well defined, and induces an isomorphism of $\mathbb{S}_\Delta^\Omega$ with $\mathbb{K}_\Delta$. □

A priori, our marking $h: X \rightarrow \mathbb{S}_\Gamma$ maps $\mathbb{K}_\Delta$ to $\mathbb{S}_\Gamma$, but we can adjust it within its homotopy class to map $\mathbb{K}_\Delta$ to $\mathbb{S}_\Delta$. This is because $h = u \circ c_\pi$, where $u: \mathbb{S}_\Gamma \rightarrow \mathbb{S}_\Gamma$ induces an element $\phi \in U^0(A_\Gamma)$. We can choose a representative $\hat{\phi} \in \text{Aut}(A_\Gamma)$ that sends A_Δ to itself, so, adjusting u by a homotopy, we get a map sending $\mathbb{S}_\Delta$ to itself. Since c_π sends $\mathbb{K}_\Delta$ to $\mathbb{S}_\Delta$, we may assume the composition $h = u \circ c_\pi$ restricts to a marking on $\mathbb{K}_\Delta$.

We constructed the subcomplex $\mathbb{K}_\Delta \subset X$ using the blowup structure that we chose on X . We now show that $\mathbb{K}_\Delta$ is essentially independent of this choice.

Proposition 5.4 *The subcomplex $\mathbb{K}_\Delta$ is independent of the blowup structure $\mathbb{S}_\Gamma^\Pi$ as long as this blowup structure satisfies $(c_\pi h^{-1})_* \in U^0(A_\Gamma)$.*

Proposition 5.4 can be proved combinatorially by keeping track of how the partitions change when we change the treelike set of hyperplanes or the labelings and orientations on the hyperplanes not in the treelike set. Being based in Δ turns out to be a property of hyperplanes in the treelike set, independent of the partitions used to describe them. The same is true for the property of being the Δ -side of a hyperplane based in $\Gamma \setminus \Delta$. Since these properties are what is used to define $\mathbb{K}_\Delta$, the subcomplex $\mathbb{K}_\Delta$ itself is independent of the blowup structure. Furthermore, an isometry $f: X \rightarrow X$ satisfying $(hf h^{-1})_* \in U^0(A_\Gamma)$ preserves these properties, so preserves $\mathbb{K}_\Delta$.

Below we give a different proof of **Proposition 5.4** in terms of the action of A_Γ on $\tilde{X}$ determined by h . This proof is more in the spirit of our previous paper [1] and more amenable to generalization. We continue to assume that $\text{lk}(\Delta) = \emptyset$.

The universal cover $\tilde{X}$ is a CAT(0) cube complex, and we will take advantage of the following facts about isometries of CAT(0) cube complexes. We say that a hyperbolic automorphism g of a CAT(0) cube complex *skewers* a hyperplane H if some axis for g crosses H transversely.

Lemma 5.5 *Let g be a hyperbolic automorphism of a CAT(0) cube complex $\tilde{X}$ and let H be a hyperplane in $\tilde{X}$. If some axis for g skewers H , then every axis for g skewers H . If no axis for g skewers H , then either all axes for g are on the same side of H , or H contains an axis for g .*

Proof Let α_1 and α_2 be two axes for g . Suppose α_1 crosses H but not α_2 does not. Let H^+ denote the half-space containing α_2 and H^- the complementary half-space. Set $\alpha_1^\pm = H^\pm \cap \alpha_1$. Let γ be a geodesic connecting a point in α_1^- to a point in α_2 . Any such path must cross H .

The action of g preserves both axes, so either g or g^{-1} maps α_1^+ into itself. Without loss of generality, assume $g(\alpha_1^+) \subset \alpha_1^+$ or, equivalently, $\alpha_1^- \subset g(\alpha_1^-)$. Consider the action on $\tilde{X}$ by positive powers of g . Since H does not intersect α_2 , the hyperplanes $g^k H$ also do not intersect α_2 . Thus, $g^k H$ separates α_1^- from α_2^+ for all $k > 0$. But this means that the path γ must cross infinitely many hyperplanes, which is impossible.

Now suppose α_1 and α_2 lie on opposite sides of H . The min set of g decomposes as an orthogonal product of an axis α and a convex subspace of $\tilde{X}$. It follows that α_1 and α_2 span a strip $\alpha \times I$ for some interval I . This strip intersects H in a convex set which separates these two axes. Any such set must contain $\alpha \times y$ for some point $y \in I$. $\square$

Proof of Proposition 5.4 Fix a basepoint at a vertex $x_0 \in X$, and let p_0 denote the base vertex of $\mathbb{S}_\Gamma$. We may assume the marking $h: X \rightarrow \mathbb{S}_\Gamma$ sends x_0 to p_0 , so induces an isomorphism $\pi_1(X, x_0) \cong \pi_1(\mathbb{S}_\Gamma, p_0) \equiv A_\Gamma$. The collapse map $c_\pi: X = \mathbb{S}_\Gamma^\Pi \rightarrow \mathbb{S}_\Gamma$ gives another marking, and by construction the composition $\phi = (h \circ c_\pi^{-1})_*: A_\Gamma \rightarrow A_\Gamma$ lies in $U^0(A_\Gamma)$ (as an automorphism). Since A_Δ is U^0 -invariant, ϕ sends A_Δ to a conjugate of itself. We may therefore homotope h , by dragging the basepoint p_0 around a loop in $\mathbb{S}_\Gamma$, so that ϕ sends A_Δ isomorphically to A_Δ . If we choose a lift $\tilde{x}_0$ of x_0 to the universal cover $\tilde{X}$, we obtain two actions of A_Γ on $\tilde{X}$, one from h and the other from c_π . Since ϕ preserves A_Δ , this means that the axes for elements of A_Δ under both actions coincide setwise.

Our goal will now be to characterize $\mathbb{K}_\Delta$ as subcomplex of X purely in terms of the U^0 -marking $h: X \rightarrow \mathbb{S}_\Gamma$, or equivalently, in terms of the action of A_Γ on $\tilde{X}$. The preceding paragraph implies that, if we can characterize $\mathbb{K}_\Delta$ in terms of the set of axes of elements of A_Δ , it does not matter whether we use the action from h or c_π .

Now let $\mathcal{H}$ be the set of hyperplanes in $\tilde{X}$ that are *not* skewered by any element of A_Δ . By Lemma 5.5, this is the same set whether we are considering the action defined by h or by c_π .

Claim *If $H \in \mathcal{H}$, then exactly one half-space $H^+ \subset \tilde{X}$ contains an axis for every element of A_Δ .*

Proof Let $\tilde{\mathbb{K}}_\Delta$ be the lift of $\mathbb{K}_\Delta$ preserved by the action of A_Δ . The lift $\tilde{\mathbb{K}}_\Delta$ is convex, and hence is connected and contains an axis for every element of A_Δ .

Let H be a hyperplane in $\tilde{X}$. We claim that $H \in \mathcal{H}$ if and only if $H \cap \tilde{\mathbb{K}}_\Delta = \emptyset$, so that all of $\tilde{\mathbb{K}}_\Delta$ is on the same side of H , and that side contains an axis for every $g \in A_\Delta$.

If $H \cap \tilde{\mathbb{K}}_\Delta \neq \emptyset$, then $\tilde{\mathbb{K}}_\Delta$ contains an edge dual to H . Lemma 3.10 of [1] implies that every edge in $\tilde{\mathbb{K}}_\Delta$ is in some axis for some $g \in A_\Delta$, so H is skewered by an element of A_Δ , ie H is not in $\mathcal{H}$. Conversely, if H is skewered by an element $g \in A_\Delta$, then, by Lemma 5.5, every axis for g skewers H , so in particular some axis contained in $\tilde{\mathbb{K}}_\Delta$ skewers H , so H intersects $\tilde{\mathbb{K}}_\Delta$.

If both half-spaces determined by H contain axes for every element of A_Δ , then, by Lemma 5.5, H itself contains an axis for every element of A_Δ . Let e_H be an edge dual to H . Then e_H is contained in an axis for some element $w \in A_\Gamma$, since that is true of every edge in $\tilde{X}$, and $w \notin A_\Delta$ since w skewers H . By [1, Lemma 3.10], all of H is contained in the min set for w so w commutes with every element of A_Δ , ie A_Δ has nontrivial centralizer. This contradicts the assumption that $\text{lk}(\Delta) = \emptyset$. $\triangleleft$

We now define

$$\tilde{X}_\Delta = \bigcap_{H \in \mathcal{H}} H_\Delta,$$

where $H_\Delta = H^+ \setminus \kappa(H)$ is the largest subcomplex of $\tilde{X}$ contained in H^+ . This is independent of the blowup structure, and coincides with $\tilde{\mathbb{K}}_\Delta$ for the action defined by c_π . Therefore, the image $\mathbb{K}_\Delta$ of $\tilde{X}_\Delta$ in X is independent of the blowup structure. $\square$

Remark 5.6 In the terminology of [2], the set of hyperplanes $\mathcal{H}$ occurring in the proof of Proposition 5.4 are exactly those which are *inessential* for the action of A_Δ . Indeed, by [2, Proposition 3.2(ii)], the essential core $\text{Ess}(\tilde{X}, A_\Delta)$ for the action of A_Δ on $\tilde{X}$ consists of those hyperplanes skewed by some axis of A_Δ , ie the complement of $\mathcal{H}$. Thus, the proposition asserts the existence of a convex subcomplex of $\tilde{X}$ whose hyperplanes extend exactly to $\text{Ess}(\tilde{X}, A_\Delta)$, and on which A_Δ acts cocompactly.

Notation Since Proposition 5.4 shows that $\mathbb{K}_\Delta$ is independent of the blowup structure, we emphasize this by using the notation X_Δ instead of $\mathbb{K}_\Delta$ for the image of $\tilde{X}_\Delta$ in X .

Corollary 5.7 Let Δ be a U^0 -invariant subgraph with $\text{lk}(\Delta) = \emptyset$, and $G < U^0(A_\Gamma)$ a finite subgroup which is realized on a Γ -complex X with an untwisted marking $h: X \rightarrow \mathbb{S}_\Gamma$. Then the restriction of G is realized on $(X_\Delta, h|_{X_\Delta})$.

Proof An element of G is realized by an isometry $f: X \rightarrow X$. An isometry of X sends any blowup structure to a new blowup structure. If the isometry induces an element of $U^0(A_\Gamma)$, we have shown that X_Δ has not changed, so f must send X_Δ to itself. $\square$

Remark 5.8 We constructed X_Δ assuming that $\text{lk}(\Delta) = \emptyset$. If this is not the case, we can look instead at $\text{st}(\Delta)$, which always has empty link. We will see in Section 7 that the complex for $\text{st}(\Delta) = \Delta * \text{lk}(\Delta)$ breaks into a product $X_\Delta \times X_{\text{lk}(\Delta)}$, where X_Δ and $X_{\text{lk}(\Delta)}$ are Δ - and $\text{lk}(\Delta)$ -complexes, respectively.

6 Extendable Δ -complexes

Let Δ be a U^0 -invariant subgraph of Γ , $\rho: G \rightarrow U^0(A_\Gamma)$ a homomorphism from a finite group G and (X, h) a marked Γ -complex realizing ρ . In the last section we found a G -invariant Δ -complex (possibly subdivided) sitting inside X . In this section we consider the opposite problem: given a marked Δ -complex Y realizing the restriction $\rho_\Delta: G \rightarrow U^0(A_\Gamma)$, when can Y sit equivariantly inside a marked Γ -complex? If there is such a marked Γ -complex, we say Y is *extendable*. To determine when Y is extendable, we first define what it means for a Δ -partition to be extendable.

Definition 6.1 Let $\Delta \subset \Gamma$ be a U^0 -invariant subgraph and let $\mathfrak{Q}$ be a Δ -partition. We say that $\mathfrak{Q}$ is *extendable* if there exists a Γ -partition $\hat{\mathfrak{Q}}$ such that $\hat{\mathfrak{Q}} \cap \Delta^\pm = \mathfrak{Q}$.

Proposition 6.2 Let Δ be a U^0 -invariant subgraph of Γ and $\mathfrak{Q}$ a Δ -partition. Then $\mathfrak{Q}$ is extendable if and only if there is some base m of $\mathfrak{Q}$ such that

- (1) $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(m)$ for every v split by $\mathfrak{Q}$, and
- (2) if v_1 and v_2 are in the same component of $\Gamma \setminus \text{st}_\Gamma(m)$, then $v_1^\pm, v_2^\pm$ are all in the same side of $\mathfrak{Q}$.

Proof The “only if” direction is immediate, since any extension of $\mathfrak{Q}$ is a Γ -partition. Note that any extension of $\mathfrak{Q}$ also splits m . In fact, it has to be based at m since Δ is U^0 -invariant, which implies there is no $v \in \Gamma \setminus \Delta$ with $\text{lk}(v) \supseteq \text{lk}(m)$.

For the converse, suppose $\mathfrak{Q} = (\mathrm{lk}_\Delta^\pm(m) \mid Q_1 \mid Q_2)$ satisfies conditions (1) and (2). We build a Γ -partition $(\mathrm{lk}_\Gamma^\pm(m) \mid \hat{Q}_1 \mid \hat{Q}_2)$ based at m as follows.

If $v \in \mathrm{lk}_\Delta(m)$, then $v \in \mathrm{lk}_\Gamma(m)$. If v (resp. v^{-1}) is in Q_i , put v (resp. v^{-1}) in $\hat{Q}_i$. This determines where to place all $v^{\pm 1}$ for $v \in \Delta$.

Now suppose $v \in \Gamma \setminus (\Delta \cup \mathrm{st}_\Gamma(m))$ and let C be the component of $\Gamma \setminus \mathrm{st}_\Gamma(m)$ containing v . If $C \cap \Delta = \emptyset$, put all vertices of C and their inverses in the same side of $\hat{\mathfrak{Q}}$ (either side will do). If $C \cap \Delta$ is nonempty, then some side Q_i of $\mathfrak{Q}$ contains an element $w \in \Delta$. We must have $w^{-1} \in Q_i$ as well, since otherwise $\mathrm{lk}(w) \subseteq \mathrm{lk}(m)$ by condition (1), which would imply that w was the only vertex of C . By condition (2), the side Q_i is independent of the choice of w . Put all vertices of C and their inverses into $\hat{Q}_i$. $\square$

Definition 6.3 A blowup $\mathbb{S}_\Delta^\Omega$ is *extendable* if every $\mathfrak{Q} \in \Omega$ is extendable. Note we are not assuming the extended partitions are compatible. A Δ -complex is *extendable* if it can be given the structure of an extendable blowup.

Remark 6.4 It is not hard to show that, if a Δ -complex is extendable with respect to one blowup structure, then it is extendable with respect to any blowup structure, but we will not need this fact.

6.1 U^0 -invariant subgraphs and extendability

In this subsection we give a condition that guarantees that a Δ -complex realizing ρ_Δ is extendable.

Definition 6.5 A G -action on a Γ -complex X is *reduced* if no orbit of hyperplanes is contained in any treelike set. A marked Γ -complex (X, h) realizing $\rho: G \rightarrow \mathrm{Out}(A_\Gamma)$ is *reduced* if the associated G -action on X is reduced.

If a marked Γ -complex (X, h) realizing ρ is not reduced, then some orbit $G.H$ is acyclic since it is contained in the treelike (acyclic) set associated to some blowup structure on X . We can collapse every hyperplane in $G.H$ to produce a new marked Γ -complex. The following lemma guarantees that the new Γ -complex still realizes ρ .

Lemma 6.6 Let X be an NPC cube complex and let $G \leq \mathrm{Isom}(X)$ be a subgroup. Suppose $\mathcal{S}$ is a collection of hyperplanes that is acyclic and G -invariant. If $G \rightarrow \mathrm{Out}(\pi_1(X))$ is injective, then the collapse map $c: X \rightarrow X//\mathcal{S}$ induces an injection $G \rightarrow \mathrm{Isom}(X//\mathcal{S})$.

Proof Let $\pi = \pi_1(X)$. Since $\mathcal{S}$ is acyclic, c is a homotopy equivalence. The fact that G preserves $\mathcal{S}$ means there is an induced map $\bar{c}: G \rightarrow \mathrm{Isom}(X//\mathcal{S})$. We obtain a commutative diagram

$$(1) \quad \begin{array}{ccc} G & \longrightarrow & \mathrm{Out}(\pi) \\ \downarrow \bar{c} & & \downarrow \cong \\ \mathrm{Isom}(X//\mathcal{S}) & \longrightarrow & \mathrm{Out}(\pi) \end{array}$$

By assumption, $G \leq \mathrm{Isom}(X)$ injects into $\mathrm{Out}(\pi)$; hence, G also injects under $\bar{c}$. $\square$

Let (X, h) be a marked Γ -complex realizing ρ . By Lemma 6.6, we may continue to collapse G -orbits of hyperplanes until we obtain a reduced marked Γ -complex realizing ρ . Note that the result of this process is not unique, but depends on the set of orbits we choose to collapse.

In Section 5 we produced a subcomplex X_Δ of X for any U^0 -invariant subgraph Δ with empty link. If (X, h) is reduced, then in any blowup structure the orbit of any hyperplane labeled by a partition $\mathcal{P}$ contains a hyperplane labeled by an element of V . Since the action of G preserves the subcomplex X_Δ , the same must be true for orbits in X_Δ , so the action of G on X_Δ is also reduced. As an example, note that, if (X, h) is reduced and the restriction of G to $U^0(A_\Delta)$ is trivial, then the subcomplex X_Δ must be equal to the Salvetti complex $\mathbb{S}_\Delta$. This follows since no nontrivial blowup of $\mathbb{S}_\Delta$ is reduced with respect to the trivial action. Thus, reduced realizations of G may be thought of as “minimal” Γ -complexes realizing G .

Let Δ be a U^0 -invariant subgraph of Γ . The next proposition states that being reduced is sufficient to guarantee extendability for any Δ -complex realizing the restriction ρ_Δ .

Proposition 6.7 *Let (X_Δ, h_Δ) be a marked Δ -complex that realizes ρ_Δ . If the action of G on X_Δ is reduced, then X_Δ is extendable.*

Proof Let V be the set of vertices in Δ . Choose a blowup structure $X_\Delta \cong \mathbb{S}_\Delta^\Omega$. Since the G -action is reduced, for every $\mathfrak{Q} \in \Omega$ there exists $g \in G$ such that $g.H_\mathfrak{Q} = H_w$ for some $w \in V$. By Proposition 6.2, in order to verify extendability we must show:

- (i) There is some base m of $\mathfrak{Q}$ such that $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(m)$ for every v split by $\mathfrak{Q}$.
- (ii) If v_1 and v_2 are in the same component of $\Gamma \setminus \text{st}_\Gamma(m)$, then $v_1^\pm, v_2^\pm$ are all in the same side of $\mathfrak{Q}$.

Proof of (i) Let χ be a characteristic cycle for v in $\mathbb{S}_\Delta^\Omega$. The image $g.\chi$ is a path that crosses each hyperplane of $\mathbb{S}_\Delta^\Omega$ at most once, so the edge labels that are not in Ω spell a cyclic word that is the image of v under g ,

$$(2) \quad g(v) = x_1^{\epsilon_1} \cdots x_k^{\epsilon_k},$$

where all the x_i are distinct and $\epsilon_i = \pm 1$.

Claim 6.8 *If $x_i \in V$ labels an edge in $g.\chi$, then $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(x_i)$. If $\mathfrak{Q} \in \Omega$ labels an edge in $g.\chi$, then $\mathfrak{Q}$ splits some x_i .*

Proof The double link $\text{dlk}(v) = \text{lk}(\text{lk}(v))$ is U^0 -invariant by Proposition 4.4(1). Since $v \in \text{dlk}(v)$, this implies that each x_i appearing in (2) must be in $\text{dlk}(v)$, ie $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(x_i)$. Now suppose $\mathfrak{Q}$ labels an edge of $g.\chi$ for some $\mathfrak{Q} \in \Omega$. We claim that $\mathfrak{Q}$ splits at least one of the x_i . Indeed, after choosing a basepoint $*$ in X_Δ , we see that $g.\chi$ is freely homotopic to a concatenation of edge paths $\eta_1\eta_2 \cdots \eta_k$, where η_i is an edge path based at $*$ representing the element x_i . Each η_i is freely homotopic to a characteristic cycle for x_i ; hence, η_i crosses the hyperplane $H_\mathfrak{Q}$ an odd number of times if and only if $\mathfrak{Q}$ splits x_i . Since $\eta_1 \cdots \eta_k$ is freely homotopic to $g.\chi$ and $g.\chi$ crosses $H_\mathfrak{Q}$ exactly once, we must have that $\mathfrak{Q}$ splits some x_i . $\triangleleft$

The proof of (i) now follows directly from the following claim.

Claim 6.9 *For every $\mathfrak{Q} \in \Omega$, there exists $m \in \text{split}(\mathfrak{Q})$ such that $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(m)$ for every $v \in \text{split}(\mathfrak{Q})$. Moreover, defining $\text{lk}_\Gamma(\mathfrak{Q}) = \text{lk}_\Gamma(m)$, the action of G preserves $\text{lk}_\Gamma(E)$ for every edge label $E \in V \cup \Omega$.*

Proof Define an increasing filtration $\emptyset = V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_r = V$, where, for $i \geq 1$, $V_i \setminus V_{i-1}$ consists of all $v \in \Delta$ such that $\text{lk}_\Gamma(v)$ is maximal among elements of $V \setminus V_{i-1}$. Every partition E in Ω splits some element of V , so we can extend this to an increasing filtration $\emptyset = \mathcal{F}_0 \subsetneq \mathcal{F}_1 \subsetneq \cdots \subsetneq \mathcal{F}_r = V \cup \Omega$, by letting $E \in \mathcal{F}_i$ if H_E splits some generator of V_i . We will prove by induction that:

- (a) For every $\mathfrak{Q} \in \mathcal{F}_i$, there exists $m \in \text{split}(\mathfrak{Q}) \cap V_i$ such that $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(m)$ for every $v \in \text{split}(\mathfrak{Q})$.
- (b) Define $\text{lk}_\Gamma(\mathfrak{Q}) = \text{lk}_\Gamma(m)$; then, for any $A, B \in \mathcal{F}_i$, if $g.H_A = H_B$, then $\text{lk}_\Gamma(A) = \text{lk}_\Gamma(B)$.

The base case $\mathcal{F}_0 = \emptyset$ holds vacuously. Suppose by induction that, for some $i \geq 1$, we have verified (a) and (b) for $\mathcal{F}_{i-1}$. Since the G -orbit of every $H_{\mathfrak{Q}}$ contains H_w for some $w \in V$, it follows from (b) that $\mathcal{F}_{i-1}$ is the union of all G -orbits of elements of V_{i-1} . Consider now $\mathfrak{Q} \in \mathcal{F}_i \setminus \mathcal{F}_{i-1}$. Then any generator in $\text{split}(\mathfrak{Q})$ lies in $V \setminus V_{i-1}$, and $\mathfrak{Q}$ splits some $m \in V_i \setminus V_{i-1}$. Let $g \in G$ be such that $g.H_{\mathfrak{Q}} = H_w$ for some $w \in V$. For any $v \in \text{split}(\mathfrak{Q})$, we know that $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(w)$ by Claim 6.8. In particular, $\text{lk}_\Gamma(m) \subseteq \text{lk}_\Gamma(w)$ and therefore $w \in V_i$. On the other hand, since $\mathcal{F}_{i-1}$ is a union of G -orbits and does not contain $\mathfrak{Q}$, we know that $w \notin V_{i-1}$. Hence, $w \in V_i \setminus V_{i-1}$ and therefore $\text{lk}_\Gamma(w) = \text{lk}_\Gamma(m)$ as m is maximal among all elements of $V \setminus V_{i-1}$. It follows that $\text{lk}_\Gamma(v) \subseteq \text{lk}_\Gamma(w) = \text{lk}_\Gamma(m)$ for any $v \in \text{split}(\mathfrak{Q})$, which proves (a).

Let χ be a characteristic cycle for $v \in V_i \setminus V_{i-1}$. Claim 6.8 implies that any hyperplane crossed by $g.\chi$ has a Γ -link that contains $\text{lk}_\Gamma(v)$, and hence its label is in $\mathcal{F}_i$. Since $\mathcal{F}_{i-1}$ is a union of G -orbits, if some label appearing in $g.\chi$ is not in $\mathcal{F}_{i-1}$, it must be in $\mathcal{F}_i \setminus \mathcal{F}_{i-1}$, and therefore is equivalent to v . In particular, if $g.H_v = H_A$, then $\text{lk}_\Gamma(A) = \text{lk}_\Gamma(v)$. Since every label in $\mathcal{F}_i \setminus \mathcal{F}_{i-1}$ appears in such an orbit, we conclude that G preserves the Γ -link of each element of $\mathcal{F}_i$, which proves (b) and completes the inductive step. $\triangleleft$

Proof of (ii) Suppose $\mathfrak{Q} \in \Omega$ is based at m and $v_1^\pm, v_2^\pm$ lie on opposite sides of $\mathfrak{Q}$. Since X_Δ is a blowup of $\mathbb{S}_\Delta$, we know that v_1 and v_2 lie in different components of $\Delta \setminus \text{st}_\Delta(m)$. We must show that they lie in different components of $\Gamma \setminus \text{st}_\Gamma(m)$.

As shown in [1], the inverse image of the (unique) vertex of $\mathbb{S}_\Delta$ under the collapse map $\mathbb{S}_\Delta^\Omega \rightarrow \mathbb{S}_\Delta$ is a CAT(0) subcomplex of $\mathbb{S}_\Delta^\Omega$, consisting of cubes whose edges are all labeled by partitions. Denote this subcomplex by $\mathbb{C}^\Delta$. Choose a minimal length edge path α in $\mathbb{C}^\Delta$ between any characteristic cycle for v_1 and any characteristic cycle for v_2 . Then α crosses exactly those hyperplanes labeled by partitions containing $v_1^\pm$ and $v_2^\pm$ on opposite sides. In particular, it crosses the hyperplane labeled $\mathfrak{Q}$. For $i = 1, 2$, let χ_i be a characteristic cycle for v_i starting at either end of α , and consider the edge path $\gamma = \chi_1 \alpha \chi_2 \bar{\alpha}$. Under the collapse map to $\mathbb{S}_\Delta$, the loop γ represents the element $v_1 v_2$. Observe that, by the minimality of α , the hyperplanes crossed by α , χ_1 and χ_2 are pairwise disjoint. Given an element $g \in G$, we have

$$g_*(\gamma) = g_*(\chi_1)g_*(\alpha)g_*(\chi_2)g_*(\bar{\alpha}) = g_*(\chi_1)g_*(\alpha)g_*(\chi_2)\overline{g_*(\alpha)}.$$

Since α , χ_1 and χ_2 cross each hyperplane of X_Δ at most once, and cross pairwise distinct sets of hyperplanes, the same is true of $g_*(\alpha)$, $g_*(\chi_1)$ and $g_*(\chi_2)$. Therefore, the hyperplanes crossed by $g_*(\gamma)$ that are not labeled by partitions define a (cyclic) word in the generators that is the image of $v_1 v_2$ under the action of g . Thus, we may write $g_*(v_1 v_2) = w_1 u w_2 u^{-1}$, where

$$u = y_1^{\delta_1} \cdots y_r^{\delta_r}$$

is a word in pairwise distinct generators y_j and $\delta_j = \pm 1$.

Since X_Δ is reduced, there exists some $g \in G$ such that g maps the hyperplane labeled $\mathcal{P}$ to a hyperplane labeled by one of the y_j , and, by [Claim 6.9](#), m and y_j belong to the same Γ -equivalence class. It thus suffices to prove that v_1 and v_2 lie in different components of $\Gamma \setminus \text{st}(y_j)$ for each j .

For any vertex v in Γ , $\text{dlk}(v)$ is U^0 -invariant, up to conjugacy. Thus, the cyclically reduced form of the word $g_*(v_1)$, namely w_1 , must be a word in $\text{dlk}(v_1)$, and similarly w_2 must be a word in $\text{dlk}(v_2)$. Choosing the representative of g_* in $\text{Aut}(A_\Gamma)$ to be one that takes v_1 to w_1 , we then have $g_*(v_2) = u w_2 u^{-1}$, where u is a product of generators y_i such that $\text{st}(y_i)$ separates some element of $\text{dlk}(v_1)$ from some element of $\text{dlk}(v_2)$. But, in this case, $\text{st}(y_i)$ must also separate v_1 from v_2 , as required. $\square$

7 Γ -complexes for joins and disjoint unions

In this section we describe procedures for constructing Γ -complexes realizing G when Γ is either a join or a disjoint union of U^0 -invariant subgraphs realizing the restriction of G . We begin with the case where $\Gamma = \Gamma_1 * \Gamma_2$ is a join, which is straightforward.

Proposition 7.1 *If $\Gamma = \Gamma_1 * \Gamma_2$, then any Γ -complex is a product of a Γ_1 -complex and a Γ_2 -complex. Conversely, any product of a Γ_1 -complex and a Γ_2 -complex is a $\Gamma_1 * \Gamma_2$ -complex.*

Proof Suppose $\mathbb{S}_\Gamma^\Pi$ is a blowup structure on a Γ -complex X . Write

$$\Pi = \{\mathcal{P}_1, \dots, \mathcal{P}_k, \mathcal{Q}_1, \dots, \mathcal{Q}_l\},$$

where $\mathcal{P}_i$ is based at $x_i \in \Gamma_1$ and $\mathcal{Q}_j$ is based at $y_j \in \Gamma_2$. (Note that, if one base is in Γ_i , then all bases are in Γ_i .) Since $\Gamma = \Gamma_1 * \Gamma_2$, every x_i is adjacent to every y_j , intersecting each $\mathcal{P}_i$ with $V^\pm(\Gamma_1)$ gives a Γ_1 -partition $\mathcal{P}_i^1$, and intersecting each $\mathcal{Q}_j$ with $V^\pm(\Gamma_2)$ gives a Γ_2 -partition $\mathcal{Q}_j^2$. Thus, Π is a compatible collection of Γ -partitions if and only if $\Pi_1 = \{\mathcal{P}_1^1, \dots, \mathcal{P}_k^1\}$ is a compatible collection of Γ_1 -partitions and $\Pi_2 = \{\mathcal{Q}_1^2, \dots, \mathcal{Q}_l^2\}$ is a compatible collection of Γ_2 -partitions. We conclude that X is a product of the Γ_i complexes with blowup structures $\mathbb{S}_{\Gamma_1}^{\Pi_1}$ and $\mathbb{S}_{\Gamma_2}^{\Pi_2}$, respectively. $\square$

Proposition 7.2 *Suppose $\Gamma = \Gamma_1 * \Gamma_2$ and let $\rho: G \rightarrow U^0(A_\Gamma)$ be a homomorphism. Then Γ_1 and Γ_2 are U^0 -invariant and, if (X_i, h_i) for $i = 1, 2$ are marked Γ_i -complexes that realize the restrictions ρ_i of ρ to Γ_i , their product $X_\Gamma = X_1 \times X_2$, equipped with the product action of G and the product marking $h = h_1 \times h_2$, is a marked Γ -complex realizing ρ . Moreover, the action of G on X_Γ given by ρ is reduced if and only if the actions on X_1 and X_2 are.*

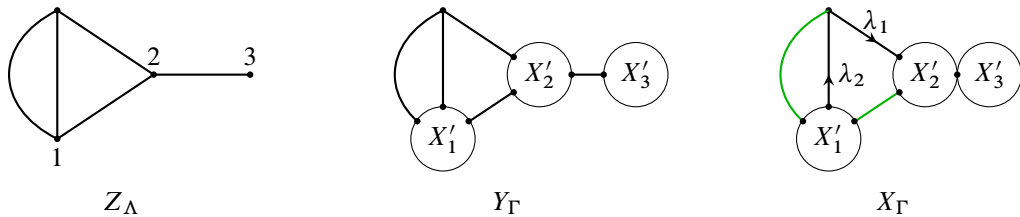


Figure 4: Building a Γ -complex for a disjoint union $\Gamma = \Gamma_1 \sqcup \Gamma_2 \sqcup \Gamma_3 \sqcup \Lambda$.

Proof Observe that $\Gamma_1 = \text{lk}(\Gamma_2)$ and $\Gamma_2 = \text{lk}(\Gamma_1)$; hence, by Proposition 4.4, both are proper U^0 -invariant subgraphs. For $i = 1, 2$, suppose X_i is a Γ_i -complex and that $h_i : X_i \rightarrow \mathbb{S}_{\Gamma_i}$ is a marking which realizes the restriction ρ_i . Define $X_\Gamma = X_1 \times X_2$ and let G act via the product action. Blowup structures on X_1 and X_2 give a blowup structure to $X_1 \times X_2$ by Proposition 7.1, and the product marking $h = h_1 \times h_2 : X_\Gamma \rightarrow \mathbb{S}_\Gamma$ realizes the action of G . The final statement of the lemma follows from the fact that hyperplanes of X_Γ are all of the form $H_1 \times X_2$ for H_1 a hyperplane of X_1 or $X_1 \times H_2$ for H_2 a hyperplane of X_2 , and that the action of G preserves the product decomposition. $\square$

We next consider the case when Γ is a disjoint union of (not necessarily connected) subgraphs. Given a Γ_i -complex X_i for each subgraph Γ_i that is not a singleton, we construct a Γ -complex X that contains each X_i as a subcomplex.

Definition 7.3 (Γ -amalgam) Suppose Γ is a disjoint union $\Gamma_1 \sqcup \cdots \sqcup \Gamma_k \sqcup \Lambda$, where Λ is discrete. Let Z_Λ be a graph satisfying:

- The rank of Z_Λ is equal to $|\Lambda|$.
- k vertices of Z_Λ are labeled by $\{1, \dots, k\}$.
- Any unlabeled vertex of Z_Λ has valence at least 3.

For each $i \in \{1, \dots, k\}$, let X_i be a Γ_i -complex. Form a new cube complex Y_Γ , called a Γ -amalgam, as follows. For each vertex v of Z_Λ , set $X_v = X_i$ if v is labeled by i , and set X_v to be a point otherwise. Now construct a complex by starting from the disjoint union $\bigsqcup X_v$ and attaching an edge from X_v to X_w whenever $\{v, w\}$ is an edge in Z_Λ . When $X_v = X_i$, we may attach the edge anywhere. The resulting complex can be given the structure of a cube complex: if the endpoint of one of the added edges lies at a point p in the interior of a cube $C \subseteq X_i$, we perform the cubical subdivision of C at p . Define the resulting cube complex to be Y_Γ . (See Figure 4.)

Observe that Y_Γ contains a subdivided copy X'_i of X_i as a subcomplex for each i . Moreover, collapsing each of these subcomplexes separately to a point defines a natural map $Y_\Gamma \rightarrow Z_\Lambda$, which is a bijection away from the X_i . Since each X_i is a Γ_i -complex, it does not have any separating hyperplanes. It follows that a hyperplane of Y_Γ is separating if and only if it comes from a separating edge of Z_Λ .

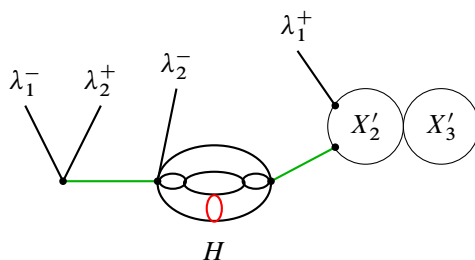


Figure 5: Determining the Γ -partition associated to a hyperplane H in X_Γ . All hyperplanes in the treelike set for X'_1 other than H have been collapsed in this figure (see proof of [Proposition 7.4](#)).

Proposition 7.4 Let $\Gamma = \Gamma_1 \sqcup \cdots \sqcup \Gamma_k \sqcup \Lambda$, where Λ is discrete, and let Y_Γ be a Γ -amalgam formed from Γ_i -complexes X_i and a graph Z_Λ . Let X_Γ be the complex formed from Y_Γ by collapsing all separating edges from Z_Γ . Then X_Γ is a Γ -complex, ie there exists a collection of Γ -partitions Π such that $X_\Gamma \cong \mathbb{S}_\Gamma^\Pi$.

Proof First note that the hyperplanes of X_Γ consist of the hyperplanes of each X'_i (which remain in X'_i) and the midpoints of nonseparating edges of Z_Λ . Choose a subdivided blowup structure on each $X'_i \subset Y_\Gamma$, corresponding to a compatible multiset of partitions Π_i . Then choose a maximal tree T_Λ in Z_Λ and label and orient the edges not in T_Λ by the elements of Λ . Let $\mathcal{T}$ be the union of all hyperplanes with labels in the Π_i and those dual to edges in T_Λ . Collapsing all hyperplanes in $\mathcal{T}$ gives the Salvetti $\mathbb{S}_\Gamma$.

To see that $\mathcal{T}$ is the treelike set for a blowup structure, we need to check that each $H \in \mathcal{T}$ determines a Γ -partition. Cut all edges of Z_Λ that are labeled by elements of Λ , labeling the initial half-edge λ^{-1} and the terminal half-edge λ (see [Figure 5](#)). Each hyperplane in $\mathcal{T}$ now determines an evident partition of the vertices of $\Gamma^\pm$. First consider the partition associated to an edge e of T_Λ . Since we chose T_Λ after collapsing each X'_i to a point, no element of Γ_i is split by this partition. The fact that Z_Λ has no separating edges implies that the partition associated to e separates some $\lambda \in \Lambda$ from its inverse, so this gives a Γ -partition based at λ .

A hyperplane in X_i partitions the vertices of $\Gamma_i^\pm$, and the only new vertices of Γ that might be split are in Λ , so have empty links; hence, the resulting partition is still a valid Γ -partition with its original base.

We must also check that the duplicate partitions in the Π_i give distinct Γ -partitions, and the singleton partitions in the Π_i give rise to legitimate Γ -partitions. Singleton and duplicate partitions in Π_i result from attaching an edge e of Z_Λ to the middle of a cube, which we then subdivide by duplicating all hyperplanes that intersect the cube. The edge e lies on a characteristic cycle for some $w \in \Lambda$.

Let H be one of the hyperplanes that has been duplicated to form new hyperplanes H' and H'' , so that now e terminates at a point between H' and H'' . If H was labeled by a partition $\mathcal{P}$ in Π_i , the new partitions $\mathcal{P}'$ and $\mathcal{P}''$ corresponding to H' and H'' agree on $\Delta^\pm$, but either w or w^{-1} (depending on the orientation of the cycle) lies in opposite sides of the extensions of $\mathcal{P}'$ and $\mathcal{P}''$ to Γ -partitions. If H was labeled by a vertex v , let H' be the duplicate hyperplane that did *not* get the label v , so H' corresponds

to a singleton partition $\mathcal{S}$. Then v (or v^{-1}) and w (or w^{-1}) lie on the same side of the extension of $\mathcal{S}$ to $\Gamma^\pm$, so the corresponding Γ -partition is not a singleton. $\square$

In passing from Y_Γ to X_Γ in the proof of [Proposition 7.4](#), we had to collapse the set of separating edges in Z_Γ . Since the collection of all separating hyperplanes in an NPC cube complex X is acyclic and invariant under $\text{Aut}(X)$, [Lemma 6.6](#) implies that collapsing them has no effect on realizing actions of finite subgroups of $\text{Out}(A_\Gamma)$ by isometries.

Proposition 7.5 *Let $\Gamma = \Gamma_1 \sqcup \cdots \sqcup \Gamma_k \sqcup \Lambda$, where Λ is discrete, and let $\rho: G \rightarrow U^0(A_\Gamma)$ be a homomorphism. If (X_i, h_i) are marked Γ_i -complexes realizing the restriction of ρ to Γ_i , then there exists a marked Γ -complex (X_Γ, h) realizing ρ such that each X_i is a subcomplex of X_Γ and $h|_{X_i} = h_i$.*

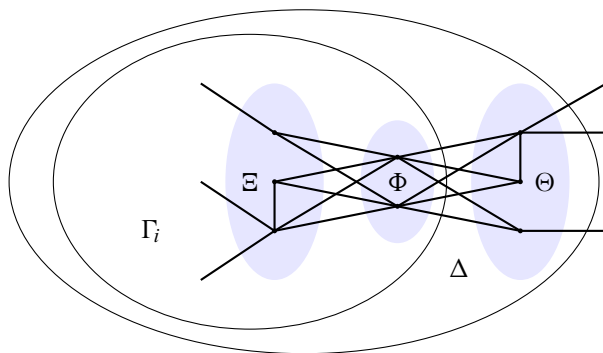
Proof Let $\bar{A}_\Gamma$ and $\bar{A}_{\Gamma_i}$ be the finite extensions of A_Γ and A_{Γ_i} , respectively, determined by G . We apply [\[9, Proposition 7.5\]](#) to $\Gamma = \Gamma_1 \sqcup \cdots \sqcup \Gamma_k \sqcup \Lambda$, with marked complexes (X_i, h_i) as input. The result is a marked cube complex X realizing the action of G on A_Γ . In the construction (see [\[9, Proposition 3.1 and Theorem 4.1\]](#)), the marked complex (X_Γ, h) is formed from a graph-of-groups decomposition of $\bar{A}_\Gamma$. The edge stabilizers are all finite, vertex stabilizers are the corresponding finite extensions of the $\bar{A}_{\Gamma_i}$, and the rank of the underlying graph is $|\Lambda|$. Each vertex is then “blown up” to a (possibly subdivided) copy of X_i equipped with the marking h_i , to which edges are attached (see [\[9, Remark 7.7\]](#)), though we only subdivide where necessary, ie where an added edge meets the interior of a cube. In particular, collapsing each of the X_i separately to points yields a graph of rank $|\Lambda|$. Thus, the hypotheses of [Proposition 7.4](#) are satisfied, and, by [Lemma 6.6](#), the resulting marked Γ -amalgam (X_Γ, h) is a marked Γ -complex which realizes ρ . $\square$

Remark 7.6 In [Proposition 6.7](#), the hypothesis that action of G on X_Δ is reduced is essential. In [Proposition 7.5](#), given a graph $\Gamma = \Gamma_1 \sqcup \cdots \sqcup \Gamma_k \sqcup \Lambda$ and a collection of Γ_i -complexes X_i , we constructed a Γ -amalgam, X_Γ , whose restriction to each Γ_i is a subdivision of the given complexes X_i . While we may assume that the original complexes X_i are reduced, the resulting subdivisions need not be, since a subdivided e_v edge has one segment labeled e_v and the rest labeled by partitions. If X_Γ is not reduced, we can collapse an acyclic set of hyperplanes in X_Γ so that the resulting complex X'_Γ is reduced, and hence its restriction X'_i to each Γ_i -subcomplex is also reduced. We claim that X'_i is still a subdivision of the original X_i . To see this, note that, since the orbit of every edge e_φ in X_i contains an e_v edge, after subdividing, the orbit of at least one segment of this edge will also contain an e_v edge. Thus, this segment will not be collapsed in the reduction process.

8 Realizing finite subgroups of $U^0(A_\Gamma)$

In this section we prove our main theorem.

Theorem 8.1 *Let Γ be a simplicial graph, G a finite group and $\rho: G \rightarrow U^0(A_\Gamma)$ a homomorphism. Then there is a Γ -complex X with an untwisted marking $h: X \rightarrow \mathbb{S}_\Gamma$ on which ρ is realized by isometries.*

Figure 6: Subgraphs of Δ referred to in the proof of [Theorem 8.1](#).

We are especially interested in the case that ρ is an inclusion, but the proof is inductive and the general case is used in the induction. The proof borrows a number of ideas from [\[8\]](#).

Proof For any U^0 -invariant subgraph Δ , we will say a marked Δ -complex (X_Δ, h_Δ) “realizes G ” as shorthand for “realizes the restriction $\rho_\Delta = r_\Delta \circ \rho: G \rightarrow U^0(A_\Delta)$ ”. Here the target of h_Δ is the subcomplex $\mathbb{S}_\Delta \subset \mathbb{S}_\Gamma$.

We proceed by constructing marked Δ -complexes realizing G for larger and larger U^0 -invariant subgraphs Δ of Γ , until we have one for all of Γ . Specifically, we define the *chain length* $\ell = \ell(\Delta)$ to be the length of a maximal chain of U^0 -invariant subgraphs $\emptyset = \Gamma_0 \subsetneq \Gamma_1 \subsetneq \cdots \subsetneq \Gamma_\ell \subsetneq \Delta$ and proceed by induction on $\ell(\Delta)$. At each stage we ensure that the G -action is reduced, so that the complexes we construct are extendable.

If $\ell(\Delta) = 0$, [Proposition 4.5](#) says that Δ is discrete, ie A_Δ is a free group. The classical realization theorem for free groups [\[5; 10; 15\]](#) says we can find a marked graph (X_Δ, h_Δ) on which G is realized by isometries. After collapsing G -invariant forests, we may assume X_Δ is reduced (in particular has no separating edges), so (X_Δ, h_Δ) is the desired marked Δ -complex.

Now suppose $\ell(\Delta) = i \geq 1$. Note that any proper U^0 -invariant subgraph of Δ has chain length strictly smaller than i , so we can construct marked complexes realizing G for any such subgraph by induction. Fix a maximal U^0 -invariant chain $\emptyset = \Gamma_0 \subsetneq \Gamma_1 \subsetneq \cdots \subsetneq \Gamma_i \subsetneq \Delta$ and let $\Theta = \Delta \setminus \Gamma_i$. The next claim follows from maximality of the chain.

Claim If $w, w' \in \Theta$, then $\text{lk}(w) \cap \Gamma_i = \text{lk}(w') \cap \Gamma_i$.

Proof By [Proposition 4.4\(5\)](#), the subgraph spanned by Γ_i and all vertices adjacent to Γ_i is invariant, so, by maximality of the chain, either every element of Θ is adjacent to Γ_i , or none of them are. In the latter case, the claim is vacuously true because $\text{lk}(w) \cap \Gamma_i = \emptyset$ for every $w \in \Theta$. In the former case, choose $w \in \Theta$ such that $\text{lk}(w) \cap \Gamma_i$ is maximal, let $W = \{w' \in \Theta \mid \text{lk}(w') \cap \Gamma_i = \text{lk}(w) \cap \Gamma_i\}$ and let Δ' be the subgraph spanned by Γ_i and W .

We will now show that Δ' is U^0 -invariant, so, by maximality of the chain, $\Delta' = \Delta$ and Θ is the graph spanned by W . By [Lemma 4.3](#), we need to prove:

- (1) If $x \in \Delta'$ and $\text{lk}(x) \subset \text{lk}(y)$, then $y \in \Delta'$.
- (2) If Δ' intersects more than one component of $\Gamma \setminus \text{st}(y)$, then $y \in \Delta'$.

We know that Δ is U^0 -invariant, so, under the hypotheses of either (1) or (2), y must be in Δ . If $y \in \Gamma_i$, we are done, so assume instead that $y \in \Theta$. We need to show that $y \in W$. If $x \in \Delta'$ and $\text{lk}(x) \subset \text{lk}(y)$, the invariance of Γ_i , together with the assumption that $y \notin \Gamma_i$, guarantees that $x \notin \Gamma_i$. That is, $x \in W$, so, by maximality, y is also in W . In case (2), the invariance of Γ_i guarantees that $\text{st}(y)$ does not separate Γ_i , so either it separates Γ_i from W , or it separates two elements of W from each other. In either case, we must have $\text{lk}(w) \subseteq \text{st}(y)$, so $\text{lk}(w) \cap \Gamma_i \subseteq \text{lk}(y) \cap \Gamma_i$; hence, $y \in W$. $\triangleleft$

Now set $\Phi = \text{lk}(\Theta) \cap \Gamma_i$, and let (X_i, h_i) be a reduced marked Γ_i -complex realizing G . Note that Φ is U^0 -invariant by [Proposition 4.4](#).

If $\Phi = \Gamma_i$, then the above claim implies that Δ is the join $\Gamma_i * \Theta$, and that $\Theta = \text{lk}(\Gamma_i) \cap \Delta$, so is U^0 -invariant. By induction, we can find a reduced marked Θ -complex (X_Θ, h_Θ) which realizes the action of G , so, by [Proposition 7.2](#), the product $(X_i \times X_\Theta, h_i \times h_\Theta)$ is a Δ -complex realizing G .

If $\Phi = \emptyset$, then Δ is the disjoint union of Γ_i and Θ , so, by [Proposition 7.5](#), we can build a Δ -complex realizing G using (X_i, h_i) and complexes for the components of Θ that are not singletons (these components are U^0 -invariant by [Proposition 4.4\(4\)](#)).

If Φ is a proper subgraph of Γ_i , let $\Xi = \text{lk}(\Phi) \cap \Gamma_i$ (see [Figure 6](#)). Now

$$\Delta = \Gamma_i \cup \text{st}_\Delta(\Phi) = \Gamma_i \cup (\Phi * (\Xi \sqcup \Theta))$$

and $\Gamma_i \cap \text{st}_\Delta(\Phi) = \Phi * \Xi$. Both Φ and Ξ are U^0 -invariant, so $\Phi * \Xi$ is also U^0 -invariant by [Proposition 4.4\(3\)](#). Since $\text{lk}_{\Gamma_i}(\Phi * \Xi) = \emptyset$, X_i contains a unique invariant (possibly subdivided) $(\Phi * \Xi)$ -complex $X_{\Phi * \Xi} = X_\Phi \times X_\Xi$ realizing G , by [Corollary 5.7](#). We may choose h_i so that it restricts to a marking $h_{\Phi * \Xi} = h_\Phi \times h_\Xi$ on $X_{\Phi * \Xi}$.

The subgraph $\text{st}_\Delta(\Phi)$ is also U^0 -invariant, and we build a $\text{st}_\Delta(\Phi)$ -complex $X_{\text{st}(\Phi)}$ realizing G as follows. We first build a complex $X_{\text{lk}(\Phi)}$ for $\text{lk}_\Delta(\Phi) = \Xi \sqcup \Theta$ using a copy of the complex X_Ξ we already found in X_i and complexes for the components of Θ that are not singletons, as we did in the case that $\Phi = \emptyset$ above. After reducing, [Proposition 6.7](#) ensures that $X_{\text{lk}(\Phi)}$ is extendable. By [Remark 7.6](#), the reduced complex still contains a subdivided copy of X_Ξ . We then take the product of $X_{\text{lk}(\Phi)}$ with a copy of the complex $X_\Phi \subset X_i$ to obtain a complex for $X_{\text{st}(\Phi)}$ realizing G with respect to the product marking.

If $\Xi \neq \emptyset$, then $\text{lk}_{\text{st}(\Phi)}(\Phi * \Xi) = \emptyset$, so the complex $X_{\text{st}(\Phi)}$ that we just built contains a unique (possibly subdivided) $(\Phi * \Xi)$ -complex realizing G . By construction, this is identical to the complex $X_{\Phi * \Xi}$ contained

in X_i , so we may glue X_i to $X_{\text{st}(\Phi)}$ by identifying these subcomplexes, thus forming a new complex X_Δ . The markings h_i and $h_{\text{st}(\Phi)}$ agree on $X_{\Phi * \Xi}$, so we obtain a marked complex (X_Δ, h_Δ) realizing G .

If $\Xi = \emptyset$, then $\text{lk}_{\text{st}(\Phi)}(\Phi) = \Theta$, so $X_{\text{st}(\Phi)} = X_\Phi \times X_\Theta$. The following claim will allow us to pick out a particular slice $X_\Phi \times \{p\} \subset X_\Phi \times X_\Theta$ to glue to the (unique) copy of X_Φ contained in X_i .

Claim *The restriction $\rho_\Theta: G \rightarrow U^0(A_\Theta)$ lifts to a homomorphism $G \rightarrow \text{Aut}(A_\Theta)$.*

Proof We use [3, Lemma 2.2], which says:

- (*) If Σ is a subgraph of Δ , then the normalizer in A_Δ of A_Σ is $A_{\text{st}_\Delta(\Sigma)} = A_\Sigma \times A_{\text{lk}_\Delta(\Sigma)}$.
- (**) If $\Sigma_1, \Sigma_2 \subset \Delta$, then $x A_{\Sigma_2} x^{-1} \leq A_{\Sigma_1}$ if and only if $\Sigma_2 \subset \Sigma_1$ and $x = x_1 x_2$ with $x_1 \in N_{A_\Delta}(A_{\Sigma_1})$ and $x_2 \in N_{A_\Delta}(A_{\Sigma_2})$.

Since all normalizers will be taken with respect to A_Δ , for the rest of this proof we omit A_Δ from the notation for normalizers.

We are assuming $\Xi = \emptyset$, so $\Phi = \text{lk}_\Delta(\Theta)$ and $\Theta = \text{lk}_\Delta(\Phi)$. Then (*) says $N(A_\Phi) = N(A_\Theta) = N(A_{\Phi * \Theta}) = A_{\Phi * \Theta}$. Furthermore, if $\text{lk}_\Delta(\Gamma_i) \neq \emptyset$, then $\Gamma_i = \Phi$ and $\Delta = \Phi * \Theta$. We have already taken care of this case, so we may assume that $\text{lk}_\Delta(\Gamma_i) = \emptyset$ and we have $N(A_{\Gamma_i}) = A_{\Gamma_i}$.

Let $g \in G$. Since Γ_i, Θ, Φ and $\Phi * \Theta$ are all U^0 -invariant, the corresponding special subgroups of A_Γ are all sent to conjugates of themselves by any representative of g in $\text{Aut}(A_\Gamma)$. Pick a representative $\hat{g}$ that sends A_{Γ_i} to itself, and suppose $\hat{g}(A_\Phi) = x A_\Phi x^{-1}$. Since $\hat{g}(A_\Phi) \leq \hat{g}(A_{\Gamma_i}) = A_{\Gamma_i}$, (**) says that $x = x_1 x_2$ with $x_1 \in N(A_{\Gamma_i}) = A_{\Gamma_i}$ and $x_2 \in N(A_\Phi)$. Then $x A_\Phi x^{-1} = x_1 x_2 A_\Phi x_2^{-1} x_1^{-1} = x_1 A_\Phi x_1^{-1}$, so, after composing $\hat{g}$ with conjugation by x_1 , we may assume $\hat{g}$ sends both A_{Γ_i} and A_Φ to themselves.

Now $\Phi \subset \Phi * \Theta$, so $\hat{g}(A_\Phi) \leq \hat{g}(A_{\Phi * \Theta}) = y A_{\Phi * \Theta} y^{-1}$ for some y . Since $\hat{g}(A_\Phi) = A_\Phi$, this gives $y^{-1} A_\Phi y \leq A_{\Phi * \Theta}$, so, by (**), $y \in N(A_{\Phi * \Theta}) N(A_\Phi) = A_{\Phi * \Theta}$, ie $\hat{g}(A_{\Phi * \Theta}) = y A_{\Phi * \Theta} y^{-1} = A_{\Phi * \Theta}$. Finally, $\Theta \subset \Phi * \Theta$, so $\hat{g}(\Theta) = z A_\Theta z^{-1} \leq \hat{g}(\Phi * \Theta) = \Phi * \Theta$ for some z , so (**) gives that $z \in N(A_{\Phi * \Theta}) N(A_\Theta) = A_{\Phi * \Theta} = N(A_\Theta)$, so $\hat{g}(A_\Theta) = A_\Theta$ as well.

Now let $g_1, g_2 \in G$ with $g_1 g_2 = g_3$, and find representatives $\hat{g}_1, \hat{g}_2$ and $\hat{g}_3$ as above. We know that $\hat{g}_1 \hat{g}_2 \hat{g}_3^{-1}$ is inner and preserves A_{Γ_i} and A_Θ , so the conjugating element lives in $N(A_{\Gamma_i}) \cap N(A_\Theta) = A_{\Gamma_i} \cap A_{\Phi * \Theta} = A_\Phi$. But conjugation by an element of A_Φ is trivial on A_Θ , ie the restriction of $\hat{g}_1 \hat{g}_2 \hat{g}_3^{-1}$ to A_Θ is the identity. Thus, $g \mapsto \hat{g}$ gives a lift of G to $\text{Aut}(A_\Theta)$. $\triangleleft$

By the claim, the action of G on X_Θ lifts to an action on $\tilde{X}_\Theta$. Since $\tilde{X}_\Theta$ is CAT(0) this action has a fixed point; projecting this fixed point back down gives a fixed point $p \in X_\Theta$. We now build X_Δ by gluing X_i to $X_{\Phi * \Theta} = X_\Phi \times X_\Theta$ along their common subspace $X_\Phi = X_\Phi \times \{p\}$. As above, the markings h_i and $h_{\Phi * \Theta}$ agree on the overlap, so give a marking $h_\Delta: X_\Delta \rightarrow \mathbb{S}_\Delta$ realizing G .

It remains to check that the complexes X_Δ that we have just built are actually Δ -complexes. We start by choosing a blowup structure $X_i \cong \mathbb{S}_{\Gamma_i}^{\Omega_i}$. Since the action of G on X_i is reduced, this blowup structure is extendable, and induces (possibly subdivided) extendable blowup structures on the subcomplexes X_Φ , X_Ξ and $X_\Phi \times X_\Xi$.

If $\Xi = \emptyset$, then finding a blowup structure is slightly easier, so we do that case first. In this case, Θ is U^0 -invariant, and we can choose an extendable blowup structure $\mathbb{S}_{\Theta}^{\Omega_\Theta}$ on X_Θ . Recall that we may have needed to subdivide X_Θ in order to make the fixed point a vertex; this means that Ω_Θ may contain trivial or duplicate partitions. The fixed point $p \in X_\Theta$ now lies in a *region*, ie a consistent choice of sides for each element of Ω_Θ . The structure on X_Θ together with the blowup structure on X_Φ now give a blowup structure on $X_\Phi \times X_\Theta$, by [Proposition 7.1](#). The partitions in Ω_i and Ω_Θ are all extendable, so in particular can be extended to Δ . We need to find extensions that form a compatible collection of Δ -partitions.

If $\mathcal{P} \in \Omega_i$, we need to decide where to put the vertices $v^\pm \in \Theta^\pm$ in our extension $\hat{\mathcal{P}}$. If $\mathcal{P}$ is based at $m \in \Phi$, they must all go into $\text{lk}(\hat{\mathcal{P}})$. If m is distance at least 2 from Θ , there is some $u \in \Phi$ with $u \notin \text{lk}(m)$ (since $\Xi = \emptyset$), so all vertices of Θ are in the same component of $\Delta \setminus \text{st}(m)$ as u , so all of $\Theta^\pm$ must go into the same side of $\mathcal{P}$ as u . (Note that the extendability of $\mathcal{P}$ guarantees that, since any two choices for u lie in the same component of $\Delta \setminus \text{st}(m)$, they must lie on the same side of $\mathcal{P}$, so there is no ambiguity here.) We also need to extend the partitions $\mathcal{Q} \in \Omega_\Theta$ to $\Delta^\pm$. All vertices in Φ must go into the link of each extension. Since Γ_i is U^0 -invariant, no vertex in Θ has a star that separates Γ_i , so vertices of $\Gamma_i \setminus \Phi$ and their inverses all have to go in the same side of $\mathcal{Q}$ for each $\mathcal{Q} \in \Omega_\Theta$. We put them all into the region determined by the fixed point p . It is now routine to check that all the extensions $\hat{\mathcal{Q}}$ and $\hat{\mathcal{P}}$ we have constructed are compatible. This verifies that X_Δ is a Δ -complex in the case $\Xi = \emptyset$. If the G -action on X_Δ is not reduced, we can reduce it to obtain an extendable Δ -complex.

If $\Xi \neq \emptyset$, we can find a blowup structure $X_{\text{lk}(\Phi)} \cong \mathbb{S}_{\text{lk}(\Phi)}^{\Omega_{\text{lk}(\Phi)}}$ that restricts to the given blowup structure on X_Ξ by [Proposition 7.4](#).

The procedure we used in the case $\Xi = \emptyset$ to extend $\mathcal{P} = (P_1 | P_2 | \text{lk}(\mathcal{P})) \in \Omega_i$ to $\Delta^\pm$ works again unless $\mathcal{P}$ is based at $m \in \Xi$. In this case, each $\mathcal{P} \cap \Xi^\pm$ is also the restriction of a partition $\mathcal{Q} = (Q_1 | Q_2 | \text{lk}(\mathcal{Q})) \in \Omega_{\text{st}(\Phi)}$. We form $\hat{\mathcal{P}}$ by adding $Q_i \cap \Theta^\pm$ to P_i for $i = 1, 2$.

We also need to extend partitions $\mathcal{Q} \in \Omega_{\text{st}(\Phi)}$ that are based at $m \in \Theta$ to Δ -partitions $\hat{\mathcal{Q}}$. The star $\text{st}(m)$ cannot disconnect Ξ or Γ_i since both are U^0 -invariant, so we add all of $\Gamma_i \setminus \Phi$ to the same side of $\hat{\mathcal{Q}}$ as Ξ . The extensions $\hat{\mathcal{P}}$ and $\hat{\mathcal{Q}}$ are now a compatible collection of Δ -partitions, giving X_Δ a blowup structure. Reducing X_Δ if necessary, this completes the induction and concludes the proof of the theorem. $\square$

Recall from [\[4\]](#) that K_Γ is a contractible subspace of $\mathbb{C}_\Gamma$ which is invariant under the action of $U^0(A_\Gamma)$. Points of K_Γ are Γ -complexes with untwisted markings, so, if ρ is an inclusion, [Theorem 8.1](#) gives the following statement.

Corollary 8.2 *The action of any finite subgroup of $U^0(A_\Gamma)$ on K_Γ has a fixed point.*

Finally, we obtain the following information about finite subgroups of $U^0(A_\Gamma)$.

Corollary 8.3 *The group $U^0(A_\Gamma)$ contains only finitely many conjugacy classes of finite subgroups.*

Proof By [Theorem 8.1](#), every finite subgroup of $U^0(A_\Gamma)$ is realized on a Γ -complex. Changing the marking produces a conjugate subgroup, so we may ignore the markings. There are only finitely many Γ -complexes since there are only finitely many partitions of $\Gamma^\pm$, so only finitely many compatible collections of Γ -partitions, each of which determines a unique blowup of $\mathbb{S}_\Gamma$. Finally, there are only finitely many cubical isometries of a finite cube complex. $\square$

References

- [1] **C Bregman, R Charney, K Vogtmann**, *Outer space for RAAGs*, *Duke Math. J.* 172 (2023) 1033–1108 [MR](#) [Zbl](#)
- [2] **P-E Caprace, M Sageev**, *Rank rigidity for CAT(0) cube complexes*, *Geom. Funct. Anal.* 21 (2011) 851–891 [MR](#) [Zbl](#)
- [3] **R Charney, J Crisp, K Vogtmann**, *Automorphisms of 2-dimensional right-angled Artin groups*, *Geom. Topol.* 11 (2007) 2227–2264 [MR](#) [Zbl](#)
- [4] **R Charney, N Stambaugh, K Vogtmann**, *Outer space for untwisted automorphisms of right-angled Artin groups*, *Geom. Topol.* 21 (2017) 1131–1178 [MR](#) [Zbl](#)
- [5] **M Culler**, *Finite groups of outer automorphisms of a free group*, from “Contributions to group theory”, *Contemp. Math.* 33, Amer. Math. Soc., Providence, RI (1984) 197–207 [MR](#) [Zbl](#)
- [6] **C W Curtis, I Reiner**, *Representation theory of finite groups and associative algebras*, *Pure Appl. Math.* 11, Interscience, New York (1962) [MR](#) [Zbl](#)
- [7] **F Haglund, D T Wise**, *Special cube complexes*, *Geom. Funct. Anal.* 17 (2008) 1551–1620 [MR](#) [Zbl](#)
- [8] **S Hensel, D Kielak**, *Nielsen realisation for untwisted automorphisms of right-angled Artin groups*, *Proc. Lond. Math. Soc.* 117 (2018) 901–950 [MR](#) [Zbl](#)
- [9] **S Hensel, D Kielak**, *Nielsen realization by gluing: limit groups and free products*, *Michigan Math. J.* 67 (2018) 199–223 [MR](#) [Zbl](#)
- [10] **D G Khramtsov**, *Finite groups of automorphisms of free groups*, *Mat. Zametki* 38 (1985) 386–392 [MR](#) [Zbl](#) In Russian; translated in *Math. Notes* 38 (1985) 721–724
- [11] **MR Laurence**, *A generating set for the automorphism group of a graph group*, *J. Lond. Math. Soc.* 52 (1995) 318–334 [MR](#) [Zbl](#)
- [12] **G Levitt, J-L Nicolas**, *On the maximum order of torsion elements in $GL(n, \mathbb{Z})$ and $\text{Aut}(F_n)$* , *J. Algebra* 208 (1998) 630–642 [MR](#) [Zbl](#)
- [13] **H Servatius**, *Automorphisms of graph groups*, *J. Algebra* 126 (1989) 34–60 [MR](#) [Zbl](#)
- [14] **J Smillie, K Vogtmann**, *Automorphisms of graphs, p -subgroups of $\text{Out}(F_n)$ and the Euler characteristic of $\text{Out}(F_n)$* , *J. Pure Appl. Algebra* 49 (1987) 187–200 [MR](#) [Zbl](#)
- [15] **B Zimmermann**, *Über Homöomorphismen n -dimensionaler Henkelkörper und endliche Erweiterungen von Schottky-Gruppen*, *Comment. Math. Helv.* 56 (1981) 474–486 [MR](#) [Zbl](#)

*Department of Mathematics, Tufts University
Medford, MA, United States*

*Department of Mathematics, Brandeis University
Waltham, MA, United States*

*Mathematics Institute, University of Warwick
Coventry, United Kingdom*

corey.bregman@tufts.edu, charney@brandeis.edu, k.vogtmann@warwick.ac.uk

Received: 15 October 2023 Revised: 14 February 2024

Computations on cobordism groups of projected immersions

ANDRÁS CSÉPAI

We use a classifying space construction by Szűcs and Terpai to extend Salomonsen's exact sequence from cobordism groups of immersions to cobordism groups of hyperplane projections of immersions with prescribed singularities. We apply this to obtain results on these cobordism groups in small and large codimensional cases.

57R90; 57R42, 57R45

1 Introduction

A *projected immersion*, or shortly *prim map*, is defined as a smooth map $f: M \rightarrow P$ between manifolds such that there is an immersion $g: M \hookrightarrow P \times \mathbb{R}$ for which f is the composition $\text{pr}_P \circ g$, where pr_P denotes projection to P . Prim maps are maps with singularities, and their cobordism groups are natural generalisations of cobordism groups of immersions and fit into the larger theory of cobordisms of singular maps; see eg Ando [1], Kalmár [6], Sadykov [9], Szűcs [16] and Szűcs and Terpai [19] for different approaches.

Throughout this paper G will denote either of the groups O and SO (although everything goes through without change if G is any stable group in the sense of Wall [21, Section 8.2], ie a direct limit of linear groups with certain properties which imply that it makes sense to talk about cobordism groups with G -structure). We shall denote by γ_{k+1}^G the tautological vector bundle over $BG(k+1)$ and by ζ_r^G the pullback of γ_{k+1}^G by the projection of $S(r+1)\gamma_{k+1}^G$, that is, the sphere bundle of $\gamma_{k+1}^G \oplus \cdots \oplus \gamma_{k+1}^G$ with $r+1$ summands. Note that here $r = \infty$ also makes sense: ζ_∞^G is simply the bundle γ_{k+1}^G . The cobordism groups of immersions and prim maps $M^n \rightarrow \mathbb{R}^{n+k}$ with normal bundle induced from a vector bundle ξ in the case of immersions, and with normal G -structure and at most Σ^{1r} -type singularities in the case of prim maps, will be denoted by $\text{Imm}^\xi(n, k)$ and $\text{Prim}_r^G(n, k)$, respectively. Note that if $s < r$, there is a natural forgetful homomorphism $\varphi_s^r: \text{Prim}_s^G(n, k) \rightarrow \text{Prim}_r^G(n, k)$.

Szűcs and Terpai [19] proved the isomorphism $\text{Prim}_r^G(n, k) \cong \text{Imm}^{\zeta_r^G}(n, k+1)$ for all r . Using this, we will prove:

Theorem 1.1 *For all $0 \leq s < r \leq \infty$ the forgetful homomorphism φ_s^r fits into a long exact sequence*

$$\begin{aligned} \cdots \rightarrow \text{Prim}_s^G(n, k) \xrightarrow{\varphi_s^r} \text{Prim}_r^G(n, k) \rightarrow \text{Imm}^{(s+2)\zeta_{r-s-1}^G}(n - (s+1)(k+1), (s+2)(k+1)) \\ \rightarrow \text{Prim}_s^G(n-1, k) \rightarrow \cdots \end{aligned}$$

The paper is organised as follows: in [Section 2](#) we recall the definitions of singularities and cobordism groups of immersions and prim maps with prescribed singularities; in [Section 3](#) we prove [Theorem 1.1](#) and make a few remarks on it; then in [Section 4](#) we apply this theorem to complete the computation of the cobordism groups $\text{Prim}_r^G(n, k)$ modulo small torsion groups where k is either at most 1 or at least $\frac{1}{2}(n-2)$.

2 Singularities and cobordism of prim maps

Definition 2.1 For a smooth map $f: M \rightarrow P$ and a number r , we define $\Sigma^r(f) \subset M$ to be the set of points $p \in M$ where the corank of the differential of f (ie $\dim \ker df_p$) is r . For generic maps f the subset $\Sigma^r(f) \subset M$ is a manifold; hence we can consider for any $s \leq r$ the subset $\Sigma^{r,s}(f) := \Sigma^s(f|_{\Sigma^r(f)})$, and iterating this process yields the definition of the subset $\Sigma^S(f)$ for any weakly decreasing sequence S of nonnegative integers. This classification of singular points defines the *Thom–Boardman singularity type* Σ^S .

Definition 2.2 A map $f: M \rightarrow P$ is said to be a *Morin map* if $\dim \ker df_p \leq 1$ for any point $p \in M$. In this case, each point $p \in M$ belongs to $\Sigma^{1r}(f)$ for some $r \geq 0$, where Σ^{1r} denotes $\Sigma^{1,\dots,1,0}$ with r ones. A Morin map f is called a Σ^{1r} -map if $\Sigma^{1r+1}(f)$ is empty.

Remark 2.3 Morin [\[8\]](#) proved that for a Morin map $f: M \rightarrow P$, the germs of f at two points $p, q \in M$ coincide up to local coordinate change precisely if they both belong to $\Sigma^{1r}(f)$ for some r . This local singularity defined by Σ^{1r} is also commonly denoted by A_r .

Now if $f: M \rightarrow P$ is a prim map, then it is Morin. Moreover, the kernel line bundle $\ker df$ over $\Sigma^1(f)$ is trivial. Here we assume prim maps f to be equipped with a fixed trivialisation of $\ker df$ which corresponds to fixing the regular homotopy class of their immersion lifts $g: M \looparrowright P \times \mathbb{R}$ (the maps such that $f = \text{pr}_P \circ g$); we will consider projections of immersions with different trivialisations of the kernel bundle to be different prim maps.

Remark 2.4 It is not hard to see that the opposite of the above observation also holds, namely a Morin map with trivialised kernel line bundle is prim; see eg [\[13\]](#).

Definition 2.5 Fix a number $0 \leq r \leq \infty$. Two prim maps $f_0: M_0^n \rightarrow P^{n+k}$ and $f_1: M_1^n \rightarrow P^{n+k}$ with closed source manifolds and at most Σ^{1r} -type singularities are said to be *prim Σ^{1r} -cobordant* if

- (i) there is a manifold W^{n+1} with boundary $\partial W = M_0 \sqcup M_1$,
- (ii) there is a prim Σ^{1r} -map $F: W \rightarrow P \times [0, 1]$ such that for $i = 0, 1$ the restriction $F|_{M_i}$ is $f_i: M_i \rightarrow P \times \{i\}$.

If the stable normal bundles of f_0 , f_1 and F above are equipped with G -structures, then we call them *prim G - Σ^{1r} -cobordant*. This cobordism is an equivalence relation and we denote the cobordism class represented by a map f by $[f]$ (suppressing G and r in the notation).

Definition 2.6 For two prim Σ^{1r} -maps $f: M^n \rightarrow \mathbb{R}^{n+k}$ and $g: N^n \rightarrow \mathbb{R}^{n+k}$ with normal G -structures, forming their disjoint union defines an abelian group operation $[f] + [g] := [f \sqcup g]$ on the set of prim G - Σ^{1r} -cobordism classes of maps of n -manifolds to $\mathbb{R}^{n+k}$; this group will be denoted by $\text{Prim}_r^G(n, k)$.

Definition 2.7 For $0 \leq s < r \leq \infty$, assigning to a prim G - Σ^{1s} -cobordism class $[f]$ the prim G - Σ^{1r} -cobordism class of f defines the *forgetful homomorphism* $\varphi_s^r: \text{Prim}_s^G(n, k) \rightarrow \text{Prim}_r^G(n, k)$.

The analogue of the cobordism group $\text{Prim}_r^G(n, k)$ for general (nonprim) Morin maps can be defined analogously to Definitions 2.5 and 2.6; we will denote this group by $\text{Mor}_r^G(n, k)$.

The special case $r = 0$ of Definition 2.5 yields the cobordism of immersions with stable normal G -structures whose cobordism group is denoted by $\text{Imm}^G(n, k)$. In this case not only the stable, but also the actual normal bundle can be specified; if $r = 0$ and the normal bundles of f_0 , f_1 and F in Definition 2.5 are induced from a vector bundle ξ , we get the definition of cobordism of immersions with normal ξ -structures, ie ξ -immersions. This cobordism group is denoted by $\text{Imm}^\xi(n, k)$. Note that we have $\text{Imm}^G(n, k) = \text{Imm}^{\gamma_k^G}(n, k)$ where γ_k^G is the tautological vector bundle over $BG(k)$.

3 Proof of Theorem 1.1

Recall that for any vector bundle ξ , there is a Pontryagin–Thom type isomorphism (due to Wells [22]) $\text{Imm}^\xi(n, k) \cong \pi_{n+k}(\Gamma T\xi) \cong \pi_{n+k}^s(T\xi)$, where $T\xi$ denotes the Thom space of ξ and Γ is the functor $\Omega^\infty S^\infty$. In other words, the classifying space of cobordisms of ξ -immersions is $\Gamma T\xi$. Similarly, cobordisms of singular maps also have classifying spaces, ie spaces X such that the appropriate cobordism group of maps $M^n \rightarrow \mathbb{R}^{n+k}$ is isomorphic to $\pi_{n+k}(X)$; see eg [9] or [16] for general constructions. In [19] Szűcs and Terpai identified the classifying space of the prim cobordism groups $\text{Prim}_r^G(n, k)$ for $0 \leq r \leq \infty$ with $\Omega \Gamma T \zeta_r^G$, yielding

$$\text{Prim}_r^G(n, k) \cong \pi_{n+k+1}^s(T \zeta_r^G) \cong \text{Imm}^{\zeta_r^G}(n, k+1).$$

Recall that ζ_r^G is defined as the pullback $p_{r+1}^* \gamma_{k+1}^G$, where $p_{r+1}: S(r+1) \gamma_{k+1}^G \rightarrow BG(k+1)$ is the projection of the sphere bundle. For $r = \infty$ this bundle has contractible fibre, and hence in this case p_{r+1} is a homotopy equivalence. The geometric interpretation of ζ_r^G -immersions $g: M^n \looparrowright \mathbb{R}^{n+k+1}$ is that they are immersions with normal G -structure, endowed with $r+1$ sections $\sigma_1, \dots, \sigma_{r+1}$ of ν_g which have no common zeros. The identification of the cobordism group $\text{Imm}^{\zeta_r^G}(n, k+1)$ with $\text{Prim}_r^G(n, k)$ on the level of representatives of cobordism classes assigns to the tuple $(g, \sigma_1, \dots, \sigma_{r+1})$ a prim map $f: M \rightarrow \mathbb{R}^{n+k}$ such that for each i the set $\Sigma^{1i}(f)$ is $\sigma_1^{-1}(0) \cap \dots \cap \sigma_i^{-1}(0)$; hence $\Sigma^{1r+1}(f)$ is empty.

Lemma 3.1 For any $0 \leq s < r \leq \infty$ there is a cofibration

$$T\zeta_s^G \hookrightarrow T\zeta_r^G \rightarrow T(s+2)\zeta_{r-s-1}^G.$$

Proof Consider the disk bundle $D\zeta_r^G$ of ζ_r^G which is the pullback $p_{r+1}^* D\gamma_{k+1}^G$. Now $S(r+1)\gamma_{k+1}^G$ is the union of the disk bundles of the pullbacks of $(s+1)\gamma_{k+1}^G$ and $(r-s)\gamma_{k+1}^G$ over $S(r-s)\gamma_{k+1}^G$ and $S(s+1)\gamma_{k+1}^G$, respectively. That is,

$$S(r+1)\gamma_{k+1}^G = D(s+1)\zeta_{r-s-1}^G \cup D(r-s)\zeta_s^G$$

glued along their common sphere bundles. Thus $D\zeta_r^G$ is the union of $D(s+2)\zeta_{r-s-1}^G$ and $D(r-s+1)\zeta_s^G$. We obtain the Thom space $T\zeta_r^G$ by factoring $D\zeta_r^G$ by $S\zeta_r^G$, and if we also factor this Thom space by the subspace $D(r-s+1)\zeta_s^G / S\zeta_r^G|_{D(r-s)\zeta_s^G}$ we get $T(s+2)\zeta_{r-s-1}^G$. But $D(r-s+1)\zeta_s^G$ deformation retracts to $D\zeta_s^G$, and this retraction takes the subspace $S\zeta_r^G|_{D(r-s)\zeta_s^G}$ to $S\zeta_s^G$; hence $T(s+2)\zeta_{r-s-1}^G$ is the factor of $T\zeta_r^G$ by $T\zeta_s^G$. $\square$

The long exact sequence of the cofibration in Lemma 3.1 in stable homotopy groups is

$$\begin{aligned} \cdots \rightarrow \text{Imm}^{\zeta_s^G}(n, k+1) &\rightarrow \text{Imm}^{\zeta_r^G}(n, k+1) \\ &\rightarrow \text{Imm}^{(s+2)\zeta_{r-s-1}^G}(n - (s+1)(k+1), (s+2)(k+1)) \rightarrow \text{Imm}^{\zeta_s^G}(n-1, k+1) \rightarrow \cdots \end{aligned}$$

Here the homomorphism $\text{Imm}^{\zeta_s^G}(n, k+1) \rightarrow \text{Imm}^{\zeta_r^G}(n, k+1)$ is induced by the natural inclusion $T\zeta_s^G \subset T\zeta_r^G$, and hence it assigns to the cobordism class of an immersion equipped with $s+1$ sections without common zeros $(g, \sigma_1, \dots, \sigma_{s+1})$ the cobordism class of $(g, \sigma_1, \dots, \sigma_{s+1}, \tau_1, \dots, \tau_{r-s})$ for some sections $\tau_1, \dots, \tau_{r-s}$. This, when considered as a map $\text{Prim}_s^G(n, k) \rightarrow \text{Prim}_r^G(n, k)$, is clearly the forgetful homomorphism φ_s^r and so Theorem 1.1 is proved.

Remark 3.2 From the proof of Lemma 3.1 it is also apparent that the homomorphism $\text{Prim}_r^G(n, k) \rightarrow \text{Imm}^{(s+2)\zeta_{r-s-1}^G}(n - (s+1)(k+1), (s+2)(k+1))$ in Theorem 1.1 assigns to the cobordism class of a prim Σ^1_r -map $f: M^n \rightarrow \mathbb{R}^{n+k}$ with immersion lift g the cobordism class of $g|_{\Sigma^{1_{s+1}}(f)}$ with its natural normal structure. Geometrically this means that the only obstruction to removing the singularities more complicated than Σ^{1_s} of a prim Σ^1_r -map by a cobordism is the cobordism class of its restriction to the closure of the $\Sigma^{1_{s+1}}$ -locus.

Remark 3.3 The extreme cases $s = r-1$ and $r = \infty$ of Theorem 1.1 are worthwhile to spell out; in fact these are the only cases we will need in the next section and they were (explicitly or implicitly) already obtained in [19]. Firstly, if $s = r-1$, we get the long exact sequence

$$\begin{aligned} \cdots \rightarrow \text{Prim}_{r-1}^G(n, k) &\rightarrow \text{Prim}_r^G(n, k) \rightarrow \text{Imm}^{(r+1)(\gamma_k^G \oplus \varepsilon^1)}(n - r(k+1), (r+1)(k+1)) \\ &\rightarrow \text{Prim}_{r-1}^G(n-1, k) \rightarrow \cdots, \end{aligned}$$

where ε^1 is the trivial line bundle and $\text{Imm}^{(r+1)(\gamma_k^G \oplus \varepsilon^1)}(n - r(k+1), (r+1)(k+1))$ is naturally isomorphic to $\text{Imm}^{(r+1)\gamma_k^G}(n - r(k+1), (r+1)k)$. This is the sequence coming from the analogue

of the key fibration (see [16]) for prim cobordisms which was used in [19] to prove its main theorem. Secondly, if $r = \infty$, we get the long exact sequence

$$\cdots \rightarrow \text{Prim}_s^G(n, k) \rightarrow \text{Imm}^G(n, k+1) \rightarrow \text{Imm}^{(s+2)\gamma_{k+1}^G}(n - (s+1)(k+1), (s+2)(k+1)) \\ \rightarrow \text{Prim}_s^G(n-1, k) \rightarrow \cdots,$$

which was shown to be rationally split in the case $G = \text{SO}$ in [19] and was used to determine the ranks of the cobordism groups $\text{Prim}_s^{\text{SO}}(n, k)$.

Remark 3.4 For $s = 0$, $r = \infty$ and $k > 0$ the exact sequence in Theorem 1.1 is an exact sequence of Salomonsen [10]: the cobordism groups in this case are

$$\text{Prim}_0^G(n, k) = \text{Imm}^G(n, k), \quad \text{Prim}_\infty^G(n, k) = \text{Imm}^G(n, k+1), \\ \text{Imm}^{2\xi_\infty^G}(n-k-1, 2k+2) = \text{Imm}^{2\gamma_{k+1}^G}(n-k-1, 2k+2),$$

and the forgetful homomorphism φ_0^∞ is the map induced by the embedding of a hyperplane $\mathbb{R}^{n+k} \subset \mathbb{R}^{n+k+1}$. Salomonsen's proof of his sequence is a special case of our proof here.

4 Computations

In the following, for any prime p , we denote by $\mathcal{C}_p$ the Serre class of abelian groups of order a power of p , and, for any number m , by $\mathcal{C}_{\leq m}$ the class of groups of order some combination of primes at most m . The symbol $\cong_{\mathcal{C}}$ for a Serre class $\mathcal{C}$ will denote isomorphism modulo $\mathcal{C}$.

In [18] the classifying spaces of cobordisms of codimension-0 oriented and unoriented and codimension-1 oriented prim maps were identified with familiar spaces, which yields the correspondences

$$\text{Prim}_r^O(n, 0) \cong_{\mathcal{C}_2} \pi_{n+1}^s(\mathbb{R}P^{r+1}) \cong_{\mathcal{C}_2} \begin{cases} 0 & \text{if } r \text{ is even or } \infty, \\ \pi^s(n-r) & \text{if } r \text{ is odd,} \end{cases} \\ \text{Prim}_r^{\text{SO}}(n, 0) \cong \pi^s(n) \oplus \pi^s(n-r), \\ \text{Prim}_r^{\text{SO}}(n, 1) \cong_{\mathcal{C}_{\leq r+1}} \pi_{n+1}^s(\mathbb{C}P^{r+1}) \cong \bigoplus_{i=0}^r \pi^s(n-2i).$$

Moreover, the groups $\text{Mor}_r^G(n, 1)$ for $G = O, \text{SO}$ were also computed modulo small torsion groups in [12] and [17], respectively:

$$\text{Mor}_r^O(n, 1) \cong_{\mathcal{C}_2} 0 \quad \text{and} \quad \text{Mor}_{2r+1}^{\text{SO}}(n, 1) \cong_{\mathcal{C}_2} \text{Mor}_{2r}^{\text{SO}}(n, 1) \cong_{\mathcal{C}_{\leq 2r+1}} \bigoplus_{i=0}^r \pi^s(n-4i).$$

We can now get an analogous description of codimension-1 unoriented prim cobordism groups:

Proposition 4.1 For all n and r we have

$$\text{Prim}_{2r}^O(n, 1) \cong_{\mathcal{C}_2} \text{Prim}_{2r-1}^O(n, 1) \cong_{\mathcal{C}_{\leq 2r}} \bigoplus_{i=0}^{r-1} \pi^s(n-2-4i).$$

Proof The exact sequence of [Theorem 1.1](#) in the case $G = O$, $k = 1$ and $s = r - 1$ is of the form

$$\cdots \rightarrow \text{Prim}_{r-1}^O(n, 1) \rightarrow \text{Prim}_r^O(n, 1) \rightarrow \text{Imm}^{(r+1)\gamma_1^O}(n - 2r, r + 1) \rightarrow \text{Prim}_{r-1}^O(n - 1, 1) \rightarrow \cdots,$$

and since the Thom space $T(r + 1)\gamma_1^O$ is $\mathbb{R}P^\infty/\mathbb{R}P^r$, we have

$$\text{Imm}^{(r+1)\gamma_1^O}(n - 2r, r + 1) \cong \pi_{n-r+1}^s(\mathbb{R}P^\infty/\mathbb{R}P^r).$$

Observe that if r is even, then the homology of $\mathbb{R}P^\infty/\mathbb{R}P^r$ with $\mathbb{Z}_p$ coefficients vanishes for all odd primes p ; hence by Serre's extension of the Whitehead theorem (see [\[11\]](#)), $\pi_{n-r+1}^s(\mathbb{R}P^\infty/\mathbb{R}P^r) \cong_{\mathcal{C}_2} 0$. This implies the isomorphism $\text{Prim}_r^O(n, 1) \cong_{\mathcal{C}_2} \text{Prim}_{r-1}^O(n, 1)$ for all even r .

The case of r odd is strongly inspired by [\[17\]](#). Note that if r is odd, then the embedding $S^{r+1} \hookrightarrow \mathbb{R}P^\infty/\mathbb{R}P^r$ as a “fibre” of this Thom space is an isomorphism in homology with $\mathbb{Z}_p$ coefficients for all odd primes p . Hence, again by Serre's theorem, it is also an isomorphism in stable homotopy groups modulo $\mathcal{C}_2$. Thus the long exact sequence above takes the form

$$\cdots \rightarrow \text{Prim}_{r-1}^O(n, 1) \rightarrow \text{Prim}_r^O(n, 1) \rightarrow \pi^s(n - 2(r + 1)) \rightarrow \text{Prim}_{r-1}^O(n - 1, 1) \rightarrow \cdots$$

modulo $\mathcal{C}_2$. We will prove that this sequence has an m -splitting $\pi^s(n - 2(r + 1)) \rightarrow \text{Prim}_r^O(n, 1)$, ie a homomorphism whose composition with the map $\text{Prim}_r^O(n, 1) \rightarrow \pi^s(n - 2(r + 1))$ in the sequence is the multiplication by m in $\pi^s(n - 2(r + 1))$, such that m has no prime divisors larger than $r + 1$. If such an m -splitting exists, then the long exact sequence splits modulo $\mathcal{C}_{\leq r+1}$, and hence an induction on r implies our statement.

Now the following steps will conclude the proof:

(i) For any r , if there is an immersion $M^{2r} \looparrowright \mathbb{R}^{2r+2}$, where M is oriented, such that the algebraic number of $(r+1)$ -tuple points (the points counted with signs) is m , then there exists an m -splitting $\sigma: \pi^s(n - 2(r + 1)) \rightarrow \text{Prim}_r^O(n, 1)$ of the above sequence.

(ii) For any r , the order $m(r)$ of the cokernel of the stable Hurewicz homomorphism

$$h_{2r-2}: \pi_{2r-2}^s(\mathbb{C}P^\infty) \rightarrow H_{2r-2}(\mathbb{C}P^\infty)$$

is such that there is an immersion $M^{2r} \looparrowright \mathbb{R}^{2r+2}$ where M is oriented and whose algebraic number of $(r+1)$ -tuple points is $m(r)$.

(iii) For any r , the number $m(r)$ defined above has no prime divisors greater than $r + 1$.

All of these statements were essentially proved in [\[17, Parts 2 and 3\]](#), so we only spell out what is different in our case. For the proof of (i) the only thing missing in [\[17\]](#) is that the map constructed as a representative of the image $\sigma(x)$ for an element $x \in \pi^s(n - 2(r + 1))$ is actually a prim map; this is not spelled out there, but it is apparent from its construction and [Remark 2.4](#). Part (ii) is only stated in [\[17\]](#) in the case where r is even; however, its proof goes through for any r by just changing the normal Pontryagin number $\langle p_1(v_f)^{r/2}, [M] \rangle$ (for an immersion $f: M^{2r} \looparrowright \mathbb{R}^{2r+2}$) to the normal Euler number $\langle e(v_f)^r, [M] \rangle$. Finally, the proof of (iii) in [\[17\]](#) goes through without change. $\square$

Remark 4.2 The above proposition implies that the cobordism group $\text{Mor}_r^{\text{SO}}(n, 1)$ is isomorphic to $\text{Prim}_{r+1}^O(n+2, 1)$ modulo $\mathcal{C}_{\leq r+1}$, and the proof suggests an even stronger connection between them. It would be interesting to know if there is a geometric explanation for this.

This completes the computation of cobordism groups $\text{Prim}_r^G(n, k)$ (modulo small torsion) when the codimension k is at most 1. Next we will consider the cases where $r = 1$ and the codimension is large, more precisely when n is either $2k + 1$ or $2k + 2$. Note that for $n \leq 2k$ we have $\text{Prim}_1^G(n, k) \cong \text{Imm}^G(n, k+1)$ since in this case a generic hyperplane projection of an immersion $M^n \hookrightarrow \mathbb{R}^{n+k+1}$ only has $\Sigma^{1,0}$ -type singularities, and this also holds for cobordisms $W^{n+1} \hookrightarrow \mathbb{R}^{n+k+1} \times [0, 1]$ between them.

The cobordism groups $\text{Mor}_1^G(n, k)$ were computed for $n = 2k + 1, 2k + 2$ in [4] and [20], respectively (see also [2, Remark 10.11]), in terms of $\mathfrak{N}_n$ for $G = O$ and Ω_n for $G = \text{SO}$. In our computation of their prim analogues we will use the group $\text{Imm}^G(n, k+1)$ instead, which was computed modulo small torsion in [13] and [14] for $G = O, \text{SO}$, respectively,

$$\text{Imm}^O(n, k+1) \cong_{\mathcal{C}_2} \begin{cases} \Omega_{n-k-1} & \text{if } k \text{ is odd,} \\ 0 & \text{if } k \text{ is even,} \end{cases} \quad \text{Imm}^{\text{SO}}(n, k+1) \cong_{\mathcal{C}_{\leq 3}} \begin{cases} P_{n,k+1} \oplus \Omega_{n-k-1} & \text{if } k \text{ is odd,} \\ P_{n,k+1} & \text{if } k \text{ is even,} \end{cases}$$

whenever $n < 3k + 3$, where $P_{n,k+1}$ is the subgroup of Ω_n consisting of cobordism classes $[M]$ whose each Pontryagin number $p_I[M]$ such that p_I is divisible by a normal Pontryagin class $\bar{p}_i$ with $2i > k + 1$ is zero.

Lemma 4.3 For all k there is an exact sequence

$$0 \rightarrow \text{Prim}_1^G(2k+2, k) \rightarrow \text{Imm}^G(2k+2, k+1) \rightarrow A \rightarrow \text{Prim}_1^G(2k+1, k) \rightarrow \text{Imm}^G(2k+1, k+1) \rightarrow 0,$$

where $A \cong \mathbb{Z}$ if $G = \text{SO}$, $A \cong \mathbb{Z}_2$ if $G = O$ and the arrow $\text{Imm}^G(2k+2, k+1) \rightarrow A$ is the homomorphism $\# \Sigma^{12}$ which assigns to the cobordism class of an immersion the algebraic number of Σ^{12} -points of its generic hyperplane projection (the number of points with their natural signs if $A \cong \mathbb{Z}$ and modulo 2 if $A \cong \mathbb{Z}_2$).

Proof The exact sequence of Theorem 1.1 with $r = \infty$ and $s = 1$ contains the portion

$$\begin{aligned} \cdots \rightarrow \text{Imm}^{3\gamma_{k+1}^G}(1, 3k+3) \rightarrow \text{Prim}_1^G(2k+2, k) \rightarrow \text{Imm}^G(2k+2, k+1) \rightarrow \text{Imm}^{3\gamma_{k+1}^G}(0, 3k+3) \\ \rightarrow \text{Prim}_1^G(2k+1, k) \rightarrow \text{Imm}^G(2k+1, k+1) \rightarrow 0. \end{aligned}$$

Here we can use [20, Lemmas 1 and 2] to determine the cobordism group $\text{Imm}^{3\gamma_{k+1}^G}(i, 3k+3) \cong \pi_{3k+3+i}^s(T3\gamma_{k+1}^G)$ for $i = 0, 1$; this yields that for $i = 0$ it is the group A defined above and for $i = 1$ it is trivial. The fact that $\text{Imm}^G(2k+2, k+1) \rightarrow A$ is $\# \Sigma^{12}$ follows immediately from Remark 3.2. $\square$

Proposition 4.4 For all k we have

$$\text{Prim}_1^O(2k+1, k) \cong \text{Imm}^O(2k+1, k+1), \quad \text{Prim}_1^O(2k+2, k) < \text{Imm}^O(2k+2, k+1) \text{ is of index 2.}$$

Proof We use [Lemma 4.3](#) with $G = O$, by which it is sufficient to prove that the homomorphism $\# \Sigma^{12}: \text{Imm}^O(2k+2, k+1) \rightarrow \mathbb{Z}_2$ is nontrivial. In other words, we want to show that there is an immersion $f: M^{2k+2} \rightarrow \mathbb{R}^{3k+3}$ such that its hyperplane projection has an odd number of Σ^{12} -points. But by [\[15\]](#) this number is (modulo 2) the same as the number of triple points of f , and by [\[5\]](#) for any k there is an immersion f as above with an odd number of triple points. $\square$

In the following $\ell(n)$ will denote the largest number l such that 3^l divides $\binom{3n}{n}$. For two groups A and B we will use $A ? B$ to denote an extension of A by B , ie a group for which there is an exact sequence $0 \rightarrow B \rightarrow A ? B \rightarrow A \rightarrow 0$.

Proposition 4.5 *For all k we have*

$$\text{Prim}_1^{\text{SO}}(2k+1, k) \cong \begin{cases} \text{Imm}^{\text{SO}}(2k+1, k+1) ? \mathbb{Z}_6 & \text{if } k = 1, 3, \\ \text{Imm}^{\text{SO}}(2k+1, k+1) ? \mathbb{Z}_{3^{\ell((k+1)/2)}} & \text{if } k > 3 \text{ is odd,} \\ \text{Imm}^{\text{SO}}(2k+1, k+1) ? \mathbb{Z} & \text{if } k \text{ is even,} \end{cases}$$

$$\text{Prim}_1^{\text{SO}}(2k+2, k) \cong \text{Imm}^{\text{SO}}(2k+2, k+1).$$

Proof We now use [Lemma 4.3](#) with $G = \text{SO}$; in this case we want to determine the homomorphism $\# \Sigma^{12}: \text{Imm}^{\text{SO}}(2k+2, k+1) \rightarrow \mathbb{Z}$.

If k is odd, then for any oriented immersion $f: M^{2k+2} \rightarrow \mathbb{R}^{3k+3}$ the triple points of f (in the source M) have natural signs, and by [\[15\]](#) their algebraic number is the same as $\# \Sigma^{12}[f]$. The minimal positive number of these triple points was determined by Eccles and Mitchell [\[3\]](#) and Li [\[7\]](#). Using their notation it is $3\lambda_{(k+1)/2}$, where λ_m is 2 for $m = 1, 2$ and $3^{\ell(m)-1}$ for $m > 2$ (here we need its triple since we are counting triple points in the source and not in the target). Thus the image of $\# \Sigma^{12}$ is $3\lambda_{(k+1)/2}\mathbb{Z}$ and our claim for k odd follows.

The case of k even is simpler, since again by [\[15\]](#), if $f: M^{2k+2} \rightarrow \mathbb{R}^{3k+3}$ is an oriented immersion where k is even, then the oriented cobordism class of its Σ^{12} -points is of order 2, ie it is trivial. Hence in this case the homomorphism $\# \Sigma^{12}$ is zero and the claim again follows. $\square$

References

- [1] **Y Ando**, *Cobordisms of maps with singularities of given class*, *Algebr. Geom. Topol.* 8 (2008) 1989–2029 [MR](#) [Zbl](#)
- [2] **A Csépai**, *Analogues of the Atiyah–Wall exact sequences for cobordism groups of singular maps*, *J. Singul.* 26 (2023) 128–169 [MR](#) [Zbl](#)
- [3] **P J Eccles**, **W P R Mitchell**, *Triple points of immersed orientable $2n$ -manifolds in $3n$ -space*, *J. Lond. Math. Soc.* 39 (1989) 335–346 [MR](#) [Zbl](#)
- [4] **T Ekholm**, **A Szűcs**, **T Terpai**, *Cobordisms of fold maps and maps with a prescribed number of cusps*, *Kyushu J. Math.* 61 (2007) 395–414 [MR](#) [Zbl](#)

- [5] **J F Hughes**, *Triple points of immersed $2n$ -manifolds in $3n$ -space*, Q. J. Math. 34 (1983) 427–431 [MR](#) [Zbl](#)
- [6] **B Kalmár**, *Fold maps and immersions from the viewpoint of cobordism*, Pacific J. Math. 239 (2009) 317–342 [MR](#) [Zbl](#)
- [7] **G-S Li**, *Triple points of immersed orientable $4n$ -manifolds in $6n$ -space*, Bull. Lond. Math. Soc. 28 (1996) 413–416 [MR](#) [Zbl](#)
- [8] **B Morin**, *Formes canoniques des singularités d’une application différentiable*, C. R. Acad. Sci. Paris 260 (1965) 5662–5665, 6503–6506 [MR](#) [Zbl](#)
- [9] **R Sadykov**, *Bordism groups of solutions to differential relations*, Algebr. Geom. Topol. 9 (2009) 2311–2347 [MR](#) [Zbl](#)
- [10] **H A Salomonsen**, *On the homotopy groups of Thom-complexes and unstable bordism*, from “Proceedings of the Advanced Study Institute on Algebraic Topology, III”, Various Publ. Series 13, Aarhus Univ. (1970) 476–494 [MR](#) [Zbl](#)
- [11] **J-P Serre**, *Groupes d’homotopie et classes de groupes abéliens*, Ann. of Math. 58 (1953) 258–294 [MR](#) [Zbl](#)
- [12] **E Szabó, A Szűcs, T Terpai**, *On bordism and cobordism groups of Morin maps*, J. Singul. 1 (2010) 134–145 [MR](#) [Zbl](#)
- [13] **A Szűcs**, *On the cobordism groups of immersions and embeddings*, Math. Proc. Cambridge Philos. Soc. 109 (1991) 343–349 [MR](#) [Zbl](#)
- [14] **A Szűcs**, *Cobordism groups of immersions of oriented manifolds*, Acta Math. Hungar. 64 (1994) 191–230 [MR](#) [Zbl](#)
- [15] **A Szűcs**, *On the singularities of hyperplane projections of immersions*, Bull. Lond. Math. Soc. 32 (2000) 364–374 [MR](#) [Zbl](#)
- [16] **A Szűcs**, *Cobordism of singular maps*, Geom. Topol. 12 (2008) 2379–2452 [MR](#) [Zbl](#)
- [17] **A Szűcs**, *On the cobordism groups of cooriented, codimension one Morin maps*, J. Singul. 4 (2012) 196–205 [MR](#) [Zbl](#)
- [18] **A Szűcs, T Terpai**, *Singularities and stable homotopy groups of spheres, II*, J. Singul. 17 (2018) 28–57 [MR](#) [Zbl](#)
- [19] **A Szűcs, T Terpai**, *Classifying spaces for projections of immersions with controlled singularities*, Acta Math. Hungar. 160 (2020) 273–291 [MR](#) [Zbl](#)
- [20] **T Terpai**, *Cobordisms of fold maps of $2k+2$ -manifolds into $\mathbb{R}^{3k+2}$* , from “Geometry and topology of caustics: CAUSTICS ’06”, Banach Center Publ. 82, Polish Acad. Sci. Inst. Math., Warsaw (2008) 209–213 [MR](#) [Zbl](#)
- [21] **C T C Wall**, *Differential topology*, Cambridge Stud. Adv. Math. 156, Cambridge Univ. Press (2016) [MR](#) [Zbl](#)
- [22] **R Wells**, *Cobordism groups of immersions*, Topology 5 (1966) 281–294 [MR](#) [Zbl](#)

Institute of Mathematics, Eötvös Loránd University
Budapest, Hungary

csepai.andras112358@gmail.com

Received: 30 October 2023

Rank-preserving additions for topological vector bundles, after a construction of Horrocks

MORGAN P OPIE

We produce group structures on certain sets of topological vector bundles of fixed rank. In particular, we put a group structure on complex rank 2 bundles on $\mathbb{C}P^3$ with fixed first Chern class. We show that this binary operation coincides with a construction on locally free sheaves due to Horrocks, provided the latter is defined. Using similar ideas, we give group structures on certain sets of rank 3 bundles on $\mathbb{C}P^5$.

These groups arise from the study of relative infinite loop space structures on truncated diagrams. Specifically, we show that the $(2n-2)$ -truncation of an n -connective map $X \rightarrow Y$ with a section is a highly structured group object over the $(2n-2)$ -truncation of Y . Applying these results to classifying spaces yields the group structures of interest.

[55P99](#), [55R25](#), [55R35](#); [55P47](#), [55Q05](#)

1. Introduction	2451
2. Background and setup	2457
3. Comparing group structures on rank 2 bundles	2464
4. Group structures on rank 3 bundles on $\mathbb{C}P^5$	2469
Appendix. The pointed Grothendieck construction for spaces	2472
References	2475

1 Introduction

Despite the importance of vector bundles in geometry and topology, there are few explicit methods to produce them. On complex projective spaces, the simplest complex bundles to write down are sums of line bundles. Indecomposable bundles are difficult to describe explicitly, but there are some famous examples: one is that of the Horrocks–Mumford bundle of rank 2 on $\mathbb{C}P^4$ [7]; another is the Horrocks bundle of rank 3 on $\mathbb{C}P^5$ [6].

In [5], Horrocks takes another approach and constructs new algebraic vector bundles from given ones using a modified extension group procedure. Horrocks’ construction takes as input complex rank 2 algebraic

bundles V and W on $\mathbb{C}P^3$, which must have the same first Chern class and satisfy some technical hypotheses, and outputs another complex rank 2 algebraic vector bundle with the same first Chern class. We will write $V +_H W$ for the output bundle, although the construction does not define a group structure.

Atiyah and Rees show that Horrocks' construction produces essentially all topological equivalence classes of complex rank 2 bundles on $\mathbb{C}P^3$ from the simplest ones:

Theorem 1.1 [2, Theorem 1.1] *Any complex rank 2 topological vector bundle on $\mathbb{C}P^3$ can be obtained from a sum of line bundles by iteratively*

- *tensoring by a line bundle, and*
- *applying Horrocks' construction.*

Remark 1.2 This result shows that every topological equivalence class of complex rank 2 vector bundles on $\mathbb{C}P^3$ admits an algebraic representative.

The algebra involved in Horrocks' construction is quite specialized and does not directly generalize to produce algebraic vector bundles of other ranks or on other spaces. However, we explore Horrocks' construction from a homotopical vantage point and show that, from this perspective, it does generalize.

Proposition 1.3 (topological Horrocks addition) *Fix an integer $a_1 \in \mathbb{Z}$. Let $\mathcal{G}_{a_1}$ denote the set of topological isomorphism classes of complex rank 2 topological bundles on $\mathbb{C}P^3$ with first Chern class equal to a_1 .*

- (i) *$\mathcal{G}_{a_1}$ carries an abelian group structure $+_{a_1}$, via an explicit construction on classifying spaces.*
- (ii) *The identity is $L \oplus \underline{\mathbb{C}}$, where L is the complex line bundle determined by $c_1(L) = a_1$ and $\underline{\mathbb{C}}$ is the trivial complex line bundle.*
- (iii) *The second Chern class defines a homomorphism $c_2: \mathcal{G}_{a_1} \rightarrow H^4(\mathbb{C}P^3; \mathbb{Z})$.*

Remark 1.4 The existence of a group structure on the underlying set of $\mathcal{G}_{a_1}$ can also be deduced by identifying elements of $\mathcal{G}_{a_1}$ with twisted symplectic K -theory classes. It is not hard to show that this group structure agrees with ours, but our description provides a framework for comparison with Horrocks' construction and for generalization.

We show $+_{a_1}$ addition agrees with Horrocks' construction as far as possible:

Theorem 1.5 *Suppose that V and W are two rank 2 algebraic vector bundles such that $c_1(V) = c_1(W) = a_1$ and such that the Horrocks sum $V +_H W$ is defined. Then*

$$V +_H W \simeq V +_{a_1} W$$

as topological vector bundles.

The proof of [Theorem 1.5](#) is conceptually simple. In [\[2\]](#), Atiyah and Rees show that a topological rank 2 bundle on $\mathbb{C}P^3$ is determined by its first and second Chern classes together with a $\mathbb{Z}/2$ -valued invariant called α . So, to prove that $V +_H W$ and $V +_{a_1} W$ are topologically isomorphic, it suffices to show their Chern classes and α -invariants agree.

The groups of [Proposition 1.3](#) generalize as follows:

Theorem 1.6 (topological Horrocks addition for rank 3 bundles) *Fix a complex rank 2 topological bundle V_0 on $\mathbb{C}P^5$. Let $\mathcal{G}_{V_0}$ denote the set of topological equivalence classes of complex rank 3 topological bundles on $\mathbb{C}P^5$ with the same first and second Chern classes as V_0 .*

- (i) $\mathcal{G}_{V_0}$ carries an abelian group structure, via an explicit construction on classifying spaces.
- (ii) The identity is $V_0 \oplus \underline{\mathbb{C}}$.
- (iii) The third Chern class defines a homomorphism $c_3: \mathcal{G}_{V_0} \rightarrow H^6(\mathbb{C}P^5; \mathbb{Z})$.

Remark 1.7 The underlying set of $\mathcal{G}_{V_0}$ depends only on a choice of first and second Chern classes of V_0 . However, we need an auxiliary rank 2 bundle to define the group operation.

1.1 Methods

[Proposition 1.3](#) and [Theorem 1.6](#) follow from a more general study of relative infinite loop space structures on truncated diagrams. We remind the reader of the following terminology:

Definition 1.8 Let Top denote the category of topological spaces.

- (a) The m -truncation functor $\tau_m: \text{Top} \rightarrow \text{Top}_{\leq m}$ is reflection onto the full subcategory $\text{Top}_{\leq m}$ of spaces X with $\pi_i X = 0$ for $i > m$. The functor τ_m is left adjoint to the inclusion $\text{Top}_{\leq m} \hookrightarrow \text{Top}$.
- (b) A map $f: X \rightarrow Y$ in Top is n -connective if the homotopy fiber of f has first nonzero homotopy group in degree n .
- (c) In any category, given objects X and Y and a morphism $s: Y \rightarrow X$, let $X \cup_Y X$ denote the pushout of the diagram

$$\begin{array}{ccc} Y & \xrightarrow{s} & X \\ \downarrow s & & \\ X & & \end{array}$$

Let $1: X \rightarrow X$ denote the identity. The map $\text{fold}: X \cup_Y X \rightarrow X$ is the dotted arrow making the following diagram commute:

$$\begin{array}{ccccc} Y & \xrightarrow{s} & X & & \\ \downarrow s & & \downarrow & \searrow 1 & \\ X & \longrightarrow & X \cup_Y X & \xrightarrow{\text{fold}} & X \\ & \searrow 1 & & \nearrow & \end{array}$$

In the homotopy category of spaces, our key technical result can be stated as follows:

Proposition 1.9 *Let $f: X \rightarrow Y$ be a map of pointed, simply connected spaces, with section $s: Y \rightarrow X$. Suppose that f is n -connective and let $m := 2n - 2$. Let τ_m denote the m -truncation functor. Then*

$$\tau_m X \xrightarrow{\tau_m f} \tau_m Y$$

$\nwarrow \tau_m s \nearrow$

is a group object in the homotopy category of m -truncated spaces over $\tau_m Y$, with binary operation given by truncating the diagram

$$X \times_Y X \xleftarrow{(1 \cup sf) \times (sf \cup 1)} X \cup_Y X \xrightarrow{\text{fold}} X,$$

$\nwarrow \quad \nearrow$

where the dotted arrow exists only after truncating.

Using the skeleton–truncation adjunction on the homotopy category of spaces, we obtain:

Corollary 1.10 *Let $f: X \rightarrow Y$, $s: Y \rightarrow X$, n and m be as in Proposition 1.9. Let C be an m -skeletal space and fix $c: C \rightarrow Y$. Let $[C, X]_c$ denote homotopy classes of maps $g: C \rightarrow X$ making*

$$(1-1) \quad \begin{array}{ccc} C & \xrightarrow{g} & X \\ & \searrow c & \downarrow f \\ & & Y \end{array}$$

homotopy commutative. Then $[C, X]_c$ is a group with identity $s \circ c$.

Remark 1.11 This corollary can be viewed as a relative version of Borsuk’s classical construction, as given in Spanier [13]. More recently, such ideas have been used by Asok and Fasel [1], for example.

It is rather tedious to prove Proposition 1.9 in a point–set way. Instead, we prove an $(\infty, 1)$ -categorical refinement. As a consequence of n -connectivity, the fibers of $\tau_m f: \tau_m X \rightarrow \tau_m Y$ are infinite loop spaces and the automorphisms of the fibers induced by paths in the base are infinite loop maps; this data should assemble to make $\tau_m X \rightarrow \tau_m Y$ an infinite loop space over $\tau_m Y$. The machinery of an $(\infty, 1)$ -categorical Grothendieck construction, for example straightening and unstraightening as in Lurie [9], allows us to make this precise.

Proposition 1.12 *Let $f: X \rightarrow Y$ be a map of pointed simply connected spaces with section $s: Y \rightarrow X$. Suppose f has homotopy fiber with first nonzero homotopy group in degree n and let $m := 2n - 2$. The functor of infinity categories*

$$\mathrm{St}_*(\tau_m f): \tau_m Y \rightarrow \mathrm{Top}_*,$$

given by applying pointed straightening St_* to the m -truncation of f , factors as

$$\begin{array}{ccc} \tau_m Y & \xrightarrow{\mathrm{St}_*(\tau_m f)} & \mathrm{Top}_* \\ \exists! \downarrow & \nearrow \Omega^\infty & \\ \mathrm{Sp} & & \end{array}$$

where $\mathbf{Sp}$ denotes the category of spectra. In particular, $\mathrm{St}_*(\tau_m f)$ takes values in infinite loop spaces and infinite loop maps.

By applying the inverse of St_* to $\mathrm{St}_*(\tau_m f)$, we recover the binary operation given in [Proposition 1.9](#) on the homotopy category. Moreover, [Proposition 1.12](#) shows that the group object

$$\tau_m X \xrightarrow[\tau_m s]{\tau_m f} \tau_m Y$$

as in [Proposition 1.9](#) is a grouplike $\mathbb{E}_\infty$ -space over $\tau_m Y$, rather than just an H -group object.

Paper outline

[Section 2](#) serves as a setup for the rest of the paper, including both technical results and key examples. In [Section 2.1](#), we give proofs of [Corollary 1.10](#) and [Propositions 1.9](#) and [1.12](#). [Section 2.1](#) relies on $(\infty, 1)$ -categorical machinery, so we summarize the key facts before our proofs. The geometrically inclined reader may choose to take the proofs of [Corollary 1.10](#) and [Proposition 1.9](#) for granted and begin their perusal with [Section 2.2](#).

[Section 2.2](#) includes the key examples of group structures on sets of vector bundles, which are the focus of the rest of the paper. These include the groups $\mathcal{G}_{a_1}$ of [Proposition 1.3](#) and $\mathcal{G}_{V_0}$ of [Theorem 1.6](#). The proof of [Theorem 1.5](#) uses certain elementary comparisons between different groups, which we include in [Section 2.3](#).

In [Section 3](#), we focus on rank 2 bundles on $\mathbb{C}P^3$. [Section 3.1](#) summarizes the Atiyah–Rees α -invariant, topological classification of rank 2 bundles on $\mathbb{C}P^3$, and Horrocks’ construction. In [Section 3.2](#), we complete the proof of [Proposition 1.3](#) by showing that c_2 is a group homomorphism. [Section 3.3](#) includes the proof of [Theorem 1.5](#).

[Section 4](#) focuses on group structures on isomorphism classes of rank 3 bundles on $\mathbb{C}P^5$. In [Section 4.1](#) we prove part (iii) of [Theorem 1.6](#). We study group structures on rank 3 bundles in more detail in [Section 4.2](#). We show that, in some of the groups $\mathcal{G}_{V_0}$ from [Theorem 1.6](#), indecomposable bundles are generated by sums of line bundles under the addition operation in the group. We view this as a proof of concept: our addition operation produces interesting bundles from simple ones. We end [Section 4.2](#) with naturally arising questions about these groups of unstable vector bundles.

We also include an [appendix](#), which supplies details of the pointed, $(\infty, 1)$ -categorical Grothendieck construction for spaces. This material is used in [Section 2.1](#), although the section can be read without the [appendix](#). This material is not new, but our exposition consolidates the necessary background and provides a reference to more comprehensive literature.

Conventions

- “Vector bundle” refers to a complex, topological vector bundle unless otherwise stated. “Rank” refers to complex rank.
- We write $\mathbf{Top}$ for the category of spaces and $\mathbf{Sp}$ for the category of spectra. We write $\mathbf{Top}_*$ for pointed topological spaces.
- Limits and colimits are implicitly homotopy limits and colimits.
- We write $\mathcal{O}(k)$ for the vector bundle on $\mathbb{C}P^n$ with first Chern class equal to k .
- Given a classifying space $BU(n)$, we write γ_n for the universal bundle on it.
- We write $\underline{\mathbb{C}}$ for the trivial rank 1 vector bundle on any space.
- Given an $(\infty, 1)$ -category C and objects X, Y in C , we write $\mathrm{Maps}_C(X, Y)$ for the space of maps from X to Y .
- The m -truncation functor $\tau_m: \mathbf{Top} \rightarrow \mathbf{Top}_{\leq m}$ is reflection onto the full subcategory $\mathbf{Top}_{\leq m}$ of spaces with X with $\pi_i X = 0$ for $i > m$, ie the left adjoint to the inclusion $\mathbf{Top}_{\leq m} \hookrightarrow \mathbf{Top}$.
- Given spaces X and Y , we write $[X, Y]$ for homotopy classes of maps with domain X and codomain Y .
- A space Y is said to be m -skeletal if, for all $X \in \mathbf{Top}$, $[Y, X] \xrightarrow{\cong} [Y, \tau_m X]$ via postcomposition with the m -truncation map $X \rightarrow \tau_m X$. Examples of m -skeletal spaces include CW complexes with cells of dimension at most m and m -dimensional manifolds.
- Fix $f: X \rightarrow Y$ in $\mathbf{Top}$. Given $Y \in s$ and $c: C \rightarrow Y$, let $[C, X]_c$ denote homotopy classes of maps $g: C \rightarrow X$ such that $f \circ g = c$.
- We say a map $f: X \rightarrow Y$ is n -connective if the homotopy fiber of f has first nonzero homotopy group in degree n .

Acknowledgements

First and foremost I want to thank my PhD advisor, Mike Hopkins, for suggesting this project and for his wisdom and support throughout my PhD program. I am also grateful to Haynes Miller for his mentorship during my time in graduate school, and to both Haynes and Elden Elmanto for serving on my dissertation committee. Mike Hill’s guidance and encouragement — mathematical and practical — were invaluable for me while improving and revising this work. This work benefited greatly from my conversations with Aravind Asok, Lukas Brantner, Hood Chatham, Jeremy Hahn, Yang Hu, Brian Shin, Alexander Smith and Dylan Wilson.

I would also like to thank the referee for extremely detailed and helpful comments on earlier drafts of this work.

While working on this project, the author was supported by the National Science Foundation under award 2202914.

2 Background and setup

This section has two goals: to prove necessary technical results about constructing group objects and to introduce key examples of rank-preserving group structures on sets of vector bundles. These examples will be the focus of the rest of the paper; the reader primarily interested in vector bundles may safely begin with [Section 2.2](#), referring to [Section 2.1](#) as needed.

In [Section 2.1](#), we prove [Proposition 1.12](#) first and then deduce [Proposition 1.9](#) and [Corollary 1.10](#). Much of this relies on aspects of the Grothendieck construction, or straightening and unstraightening, as described in detail by Lurie [9, 2.2 and 3.2] and summarized in the [appendix](#).

[Section 2.2](#) applies [Corollary 1.10](#) to maps from $\mathbb{C}P^3$ (respectively $\mathbb{C}P^5$) into diagrams involving classifying spaces of rank 2 (respectively rank 3) bundles, yielding group structures on collections of rank 2 bundles on $\mathbb{C}P^3$ (respectively rank 3 bundles on $\mathbb{C}P^5$). We also show that our methods give multiple a priori distinct group structures on the same set of bundles.

In [Section 2.3](#), we investigate the relationship between two different examples from the previous section. The results in this subsection are needed for the proof of [Theorem 1.5](#).

2.1 Technical preliminaries

In this subsection we prove [Corollary 1.10](#) and [Propositions 1.9](#) and [1.12](#), starting with the latter. We also record several other facts about the additive structures in question, which we will use later to study specific examples.

Construction 2.1 For Y a space, let $(\mathrm{Top}/Y)_*$ denote the $(\infty, 1)$ -category of pointed objects in the $(\infty, 1)$ -category of spaces over Y , ie spaces over Y together with a choice of section. Let $(\mathrm{Top}_*)^Y$ denote the $(\infty, 1)$ -category of functors from the Poincaré infinity groupoid of Y to Top_* . The following facts are justified in the [appendix](#):

- (1) Given a map of spaces $f: X \rightarrow Y$ with section $s: Y \rightarrow X$, we can naturally associate a functor $\mathrm{St}_*(f): Y \rightarrow \mathrm{Top}_*$, the *pointed straightening of f* . Heuristically, $\mathrm{St}_*(f)$ takes a point p in Y to the homotopy fiber $f^{-1}(p)$ pointed by $s(p)$.
- (2) The pointed straightening functor St_* participates in an equivalence of $(\infty, 1)$ -categories

$$\begin{array}{ccc} & \mathrm{St}_* & \\ & \curvearrowright & \\ (\mathrm{Top}/Y)_* & & (\mathrm{Top}_*)^Y \\ & \curvearrowleft & \\ & \mathrm{Un}_* & \end{array}$$

(See [Corollary A.7](#).)

- (3) There is a functor $\text{forget}: (\text{Top}/Y)_* \rightarrow \text{Top}/Y$ given on objects by $(Y \rightarrow x) \mapsto x$ and a free basepoint functor $+: \text{Top}/Y \rightarrow (\text{Top}/Y)_*$ given on objects by $x \mapsto (Y \rightarrow Y \sqcup x)$, participating in an adjunction of $(\infty, 1)$ -categories

$$\begin{array}{ccc} & \text{forget} & \\ & \curvearrowright & \\ \text{Top}/Y & \top & (\text{Top}/Y)_* \\ & \curvearrowleft & \\ & + & \end{array}$$

(See [Lemma A.8](#).)

Proof of Proposition 1.12 For nonnegative integers $a \leq b$, let $\text{Sp}_{[a,b]}$ denote the full subcategory of spectra with homotopy groups zero outside the range $[a, b]$. Similarly, let $\text{Top}_{*,[a,b]}$ denote the full subcategory of pointed spaces with homotopy groups zero outside the range $[a, b]$ (ie a -connective, b -truncated spaces).

Observe that $\text{St}_*(\tau_m f)$ takes values in $\text{Top}_{*,[n,2n-2]}$, the subcategory of pointed spaces that are n -connective and $(2n-2)$ -truncated. By [\[10, Theorem 5.1.2\]](#),

$$(2-1) \quad \Omega^\infty: \text{Sp}_{[n,2n-2]} \rightarrow \text{Top}_{*,[n,2n-2]}$$

is an equivalence of $(\infty, 1)$ -categories. So the factorization is automatic and there exists some functor of $(\infty, 1)$ -categories $A: Y \rightarrow \text{Sp}$ such that there is an equivalence

$$(2-2) \quad \Omega^\infty A \simeq \text{St}_*(\tau_m f)$$

in the $(\infty, 1)$ -category of such functors. This factorization is given by applying the inverse of Ω^∞ as in (2-1) to $\text{St}_*(\tau_m f)$ and is therefore functorial. $\square$

Corollary 2.2 With setup as in [Proposition 1.12](#), the diagram

$$\begin{array}{ccc} \tau_m X & \xrightarrow{\tau_m f} & \tau_m Y \\ & \swarrow \tau_m s & \end{array}$$

is an group object in the category of m -truncated spaces over $\tau_m Y$. The binary operation is given by applying Un_* to the operation on the group object $\text{St}_*(\tau_m f)$.

Proof of Corollary 1.10 Let $c: C \rightarrow Y$ be a map from an m -skeletal space C to our base space Y , where $m = 2n - 2$. Let c_+ denote the object

$$C \sqcup \tau_m Y \xrightarrow[c]{} \tau_m Y$$

given by applying $+$ as in [Construction 2.1\(3\)](#) to the object $\tau_m c: C \rightarrow \tau_m Y$.

[Proposition 1.12](#) implies that the space

$$\text{Maps}_{(\text{Top}_*)^{\tau_m Y}}(\text{St}_*(c_+), \text{St}_*(\tau_m f))$$

possesses the structure of a grouplike $\mathbb{E}_\infty$ -space, ie an infinite loop space.

By [Construction 2.1\(2\)](#), we have natural equivalences of grouplike $\mathbb{E}_\infty$ -spaces

$$\begin{aligned} \mathrm{Maps}_{(\mathrm{Top})_*}^{\tau_m Y}(\mathrm{St}_*(c_+), \mathrm{St}_*(\tau_{\leq m} f)) &\simeq \mathrm{Maps}_{(\mathrm{Top}/\tau_m Y)_*}(\mathrm{Un}_* \mathrm{St}_*(c_+), \mathrm{Un}_* \mathrm{St}_*(\tau_m f)) \\ &\simeq \mathrm{Maps}_{(\mathrm{Top}/\tau_m Y)_*}(C \sqcup \tau_m Y, \tau_m X) \\ &\simeq \mathrm{Maps}_{\mathrm{Top}/\tau_m Y}(C, \tau_m X), \end{aligned}$$

where the third step is the definition of c_+ and [Construction 2.1\(3\)](#).

This implies

$$\pi_0 \mathrm{Maps}_{\mathrm{Top}/\tau_m Y}(C, \tau_m X) \simeq \pi_0 \mathrm{Maps}_{\mathrm{Top}/Y}(\mathrm{sk}^m C, X) \simeq \pi_0 \mathrm{Maps}_{\mathrm{Top}/Y}(C, X) \simeq [C, X]_c,$$

giving the group structure on $[C, X]_c$. Note that the identity of $\pi_0 \mathrm{Maps}_{\mathrm{Top}/\tau_m Y}(C, \tau_m X)$ is obtained by postcomposing $c: C \rightarrow \tau_m Y$ with $s: \tau_m Y \rightarrow \tau_m X$. Tracing through the isomorphisms above, this shows that $s \circ c: C \rightarrow X$ is the identity in $[C, X]_c$. $\square$

We record one additional result that will be useful later:

Lemma 2.3 Suppose that we have a commutative solid diagram

$$\begin{array}{ccc} X & \xrightarrow{h} & X' \\ & \searrow f & \swarrow f' \\ & Y & \end{array} \quad \begin{array}{ccc} & \nearrow s & \nwarrow s' \\ & & \end{array}$$

where f and f' are both n -connective, and s and s' are sections of f and f' , respectively. Let $m = 2n - 2$. Then $\mathrm{St}_*(\tau_m h): \mathrm{St}_*(\tau_m f) \rightarrow \mathrm{St}_*(\tau_m f')$ is a natural transformation through infinite loop maps.

Proof Note that the functor $\mathrm{St}_*(\tau_m -): (\mathrm{Top}/Y)_* \rightarrow (\mathrm{Top}_*)^{\tau_m Y}$ factors through $\mathrm{Sp}^{\tau_m Y}$. $\square$

This immediately implies:

Corollary 2.4 With setup as in [Lemma 2.3](#), let C be an m -skeletal space and let $c: C \rightarrow Y$ be given. The map $h \circ -: [C, X]_c \rightarrow [C, X']_c$ is a group homomorphism.

We now explicitly describe the group operation of [Corollary 2.2](#) in the homotopy category of spaces by unraveling the binary operation on $\mathrm{Un}_* \mathrm{St}_*(\tau_m f)$.

Proof of Proposition 1.9 Recall the fold map from [Definition 1.8\(c\)](#). From the proof of [Proposition 1.12](#), equation (2-2) supplied a functor $A: \tau_m Y \rightarrow \mathrm{Sp}$ such that $\mathrm{St}_*(\tau_m f) = \Omega^\infty A$. The group object structure on A is given by

$$A \times A \simeq A \sqcup A \xrightarrow{\text{fold}} A.$$

Applying Un_* , we have a homotopy commutative diagram

$$(2-3) \quad \begin{array}{ccccc} \tau_m X \times_{\tau_m Y} \tau_m X & \xrightarrow{\simeq} & \text{Un}_* \Omega^\infty(A \times A) & \xleftarrow{\simeq} & \text{Un}_* \Omega^\infty(A \sqcup A) & \xrightarrow{\text{Un}_* \Omega^\infty \text{ fold}} & \text{Un}_* \Omega^\infty A \\ \downarrow \simeq & & & & \downarrow \simeq & & \downarrow \simeq \\ \tau_m(X \times_Y X) & & & & \tau_m X \cup_{\tau_m Y} \tau_m X & \xrightarrow{\text{fold}} & \tau_m X \\ \uparrow \simeq \tau_m(sf \times 1 \cup 1 \times sf) & & & & \downarrow \dagger & & \downarrow \dagger \\ \tau_m(X \cup_Y X) & \xrightarrow{\quad * \quad} & \tau_m(\tau_m X \cup_{\tau_m Y} \tau_m X) & \xrightarrow{\tau_m \text{ fold}} & \tau_m X \end{array}$$

where the maps labeled $\dagger$ above are the natural map from a space into its truncation and the map labeled $*$ is the τ_m -truncation of the map on pushouts induced by the map of spans

$$(2-4) \quad \begin{array}{ccccc} X & \xleftarrow{s} & Y & \xrightarrow{s} & X \\ \downarrow & & \downarrow & & \downarrow \\ \tau_m X & \xleftarrow{\tau_m s} & \tau_m Y & \xrightarrow{\tau_m s} & \tau_m X \end{array}$$

where the vertical arrows are again the natural maps from a space to its truncation.

In (2-3), arrow $\tau_m(sf \times 1 \cup 1 \times sf)$ is an equivalence by the hypothesis that f is n -connective. Inspecting the diagram, it is clear that the two right rectangles labeled $\star$ and $\star\star$ commute up to homotopy, by naturality. One then checks that starting at the bottom left corner, going up and around the large left rectangle produces the map $*$ induced by (2-4).

The composite obtained by starting at the upper left-hand corner of the outer rectangle, going down by two, and then right by two computes the proposed operation on $\tau_m X$ over $\tau_m Y$. On the other hand, starting in the upper left-hand corner but going first right by two and then down by two computes the group object structure obtained in Corollary 2.2. $\square$

From this description of the group structure, we obtain group homomorphisms from homotopy classes of maps over one base to homotopy classes of maps over another base. First note the following elementary lemma:

Lemma 2.5 Consider a diagram

$$(2-5) \quad \begin{array}{ccc} X & \xrightarrow{h} & X' \\ s \uparrow \downarrow f & & f' \downarrow \uparrow s' \\ Y & \xrightarrow{j} & Y' \end{array}$$

such that $hs \simeq s'j$, $jf \simeq f'h$, and s and s' are sections of f and f' , respectively. Such a diagram gives rise to a homotopy commutative diagram

$$\begin{array}{ccccc} X \times_Y X & \longleftarrow & X \cup_Y X & \longrightarrow & X \\ \downarrow & & \downarrow & & \downarrow \\ X' \times_{Y'} X' & \longleftarrow & X' \cup_{Y'} X' & \longrightarrow & X' \end{array}$$

Corollary 2.6 Let $h: X \rightarrow X'$ be as in (2-5). Suppose also that f and f' are n -connective and let $m = 2n - 2$. Let C be m -skeletal and let $c: C \rightarrow Y$ be given. Then postcomposition with h induces a group homomorphism $h \circ -: [C, X]_c \rightarrow [C, X']_{j \circ c}$.

2.2 Group structures on vector bundles

The goal of this section is to introduce and explore examples of groups comprising unstable vector bundles. First we note the following general procedure for obtaining group structures:

Example 2.7 Let $g: X \rightarrow Y$ be a map of connected spaces. Consider the diagram

$$\begin{array}{ccc} X \times_Z Y & \longrightarrow & Y \\ \begin{array}{c} \nearrow s_g \\ \downarrow p_1 \end{array} & & \downarrow \tau_k \\ X & \xrightarrow{\tau_k \circ g} & \tau_k Y \end{array}$$

where $Z := \tau_k Y$, p_1 is the natural projection, and the section s_g from X to the pullback is induced by the identity on X and $g: X \rightarrow Y$. The map p_1 is $k+1$ -connective since τ_k is. By Corollary 2.2, $\tau_{2k} f: \tau_{2k}(X \times_Z Y) \rightarrow \tau_{2k} X$ is a group object in $2k$ -truncated spaces over $\tau_{2k} X$. So, for any $2k$ -skeletal space C with $c: C \rightarrow X$, the set $[C, X]_c$ inherits a group structure.

The next two examples are the ones that relate to Horrocks' construction [5].

Example 2.8 The map $\det(\gamma_2): BU(2) \rightarrow BU(1)$ is 4-connective and admits a section $(\gamma_1 \otimes \mathbb{O}(-b)) \oplus \mathbb{O}(b)$, where $b \in \mathbb{Z}$ is arbitrary. If we fix $a_1 \in \mathbb{Z}$, this gives a group structure on isomorphism classes of rank 2 bundles on $\mathbb{C}P^3$ with first Chern class equal to a_1 . The identity is $\mathbb{O}(a_1 - b) \oplus \mathbb{O}(b)$.

As a special case, setting $b = 0$, we obtain the following:

Example 2.9 The determinant map $\det(\gamma_2): BU(2) \rightarrow BU(1)$ is 4-connective and admits a section represented by the bundle $\gamma_1 \oplus \mathbb{C}$ on $BU(1)$. If we fix $\mathbb{O}(a_1): \mathbb{C}P^3 \rightarrow BU(1)$ for $a_1 \in \mathbb{Z}$, this gives a group structure on isomorphism classes of rank 2 bundles on $\mathbb{C}P^3$ with first Chern class equal to a_1 . The identity is $\mathbb{O}(a_1) \oplus \mathbb{C}$.

We will explore the previous two examples and how they are related in Section 2.3, so we establish notation for each.

Definition 2.10 Let $+_{a_1, b}$ denote the operation on rank 2 bundles on $\mathbb{C}P^3$ with first Chern class a_1 given in Example 2.8. We write $\mathcal{G}_{a_1, b}$ for this group. In the case $b = 0$ (ie Example 2.9), we write $+_{a_1}$ for the operation and $\mathcal{G}_{a_1}$ for the group.

Construction 2.11 Let $b = -\frac{1}{2}a_1$. Consider the commutative diagram

$$(2-6) \quad \begin{array}{ccc} BU(2) & \xrightarrow{(-) \otimes \mathbb{C}(-b)} & BU(2) \\ (-) \oplus \mathbb{C} \downarrow \swarrow \det(-) & & \det(-) \downarrow \swarrow (-) \otimes \mathbb{C}(-b) \oplus \mathbb{C}(b) \\ BU(1) & \xrightarrow{\text{id}} & BU(1) \end{array}$$

where for maps we write the functor represented on vector bundles. By [Corollary 2.6](#), (2-6) induces a group isomorphism $\phi: \mathcal{G}_{a_1} \rightarrow \mathcal{G}_{0,b}$ given on bundles by

$$V \mapsto V \otimes \mathbb{C}(-b).$$

Next, we introduce group structures for isomorphism classes of rank 3 bundles on $\mathbb{C}P^5$, which will be our subject in [Section 4](#).

Example 2.12 Consider the diagram

$$(2-7) \quad \begin{array}{ccc} \widetilde{BU}(3) & \longrightarrow & BU(3) \\ s \downarrow \swarrow v & \lrcorner & \downarrow c_1 \times c_2 \\ BU(2) & \xrightarrow{c_1 \times c_2} & K(\mathbb{Z}, 2) \times K(\mathbb{Z}, 4) \end{array}$$

where

- $c_1 \times c_2$ is the product of the first Chern class and second Chern class;
- $\widetilde{BU}(3)$ is the indicated homotopy pullback; and
- the dotted arrow s to the pullback is induced by the identity map on $BU(2)$ and by $\gamma_2 \oplus \mathbb{C}: BU(2) \rightarrow BU(3)$.

The map v is 6-connective. By [Corollary 1.10](#), for any fixed $V_0: \mathbb{C}P^5 \rightarrow BU(2)$, the set of homotopy classes of lifts $[\mathbb{C}P^5, \widetilde{BU}(3)]_{V_0}$ carries a group structure. A priori, an element in $[\mathbb{C}P^5, \widetilde{BU}(3)]_{V_0}$ consists of a homotopy class $\mathbb{C}P^5 \rightarrow BU(3)$ and a homotopy witnessing $c_i(V) = c_i(V_0)$ for $i = 1, 2$. Such homotopies form a torsor for $H^1(\mathbb{C}P^5; \mathbb{Z}) \times H^3(\mathbb{C}P^5; \mathbb{Z}) \simeq 0$. Therefore the underlying set of this group is in fact isomorphism classes of rank 3 bundles on $\mathbb{C}P^5$ with the same first and second Chern classes as V_0 . The identity is $V_0 \oplus \mathbb{C}$.

[Example 2.12](#) is our focus in [Section 4](#), so we introduce the following terminology:

Definition 2.13 Let V_0 be a fixed rank 2 bundle on $\mathbb{C}P^5$. Let $+_{V_0}$ denote the operation on rank 3 bundles on $\mathbb{C}P^5$ with first and second Chern classes equal to those of V_0 , as defined by [Example 2.12](#). We write $\mathcal{G}_{V_0}$ for this group.

Convention 2.14 We will refer to the binary operations of [Definitions 2.10](#) and [2.13](#) as *topological Horrocks additions*.

Remark 2.15 To generalize topological Horrocks addition from rank 2 bundles to rank 3 bundles, one might hope to consider isomorphism classes of rank 3 vector bundles on $\mathbb{C}P^5$ with fixed c_1 and c_2 .

However, the additional choice of V_0 is necessary. To apply our setup, we would need a section σ of $c_1 \times c_2: BU(3) \rightarrow K(\mathbb{Z}, 2) \times K(\mathbb{Z}, 4)$.

In $H^*(BU(3); \mathbb{Z}/3\mathbb{Z})$, $P^1(c_2) = c_1^2 c_2 + c_2^2 - c_1 c_3$. In $H^*(K(\mathbb{Z}, 2) \times K(\mathbb{Z}, 4); \mathbb{Z}/3\mathbb{Z})$, $P^1(\iota_4) = 0$. Since $(c_1 \times c_2)^*(\iota_{2i}) = c_i$ for $i = 1, 2$, existence of σ would force $\sigma^*(P^1(c_2)) = P^1(\iota_4)$ and $\iota_2^2 \iota_4 + \iota_4^2 = 0$, a contradiction.

2.3 Properties of $+_{a_1}$ and $+_{a_1, b}$

Recall from Definition 2.10 that $+_{a_1}$ and $+_{a_1, b}$ are binary operations, with different identities, on the set of isomorphism classes of rank 2 vector bundles over $\mathbb{C}P^3$ with first Chern class a_1 .

Lemma 2.16 Let V, W and Z be rank 2 vector bundles on $\mathbb{C}P^3$ with first Chern class a_1 . Then

$$(V +_{a_1} W) +_{a_1, b} Z = V +_{a_1} (W +_{a_1, b} Z).$$

Proof Consider the commutative diagram

$$(2-8) \quad \begin{array}{ccccc} & BU(1) & & BU(1) & \\ & \swarrow \gamma_1 \oplus \underline{\mathbb{C}} & \searrow \gamma_1 \oplus \underline{\mathbb{C}} & \swarrow g' & \searrow g' \\ BU(2) & & BU(2) & & BU(2) \\ & \searrow \det \gamma_2 & \swarrow \det \gamma_2 & \searrow \det \gamma_2 & \swarrow \det \gamma_2 \\ & BU(1) & & BU(1) & \end{array}$$

where $g' := \gamma_1 \otimes \mathbb{O}(-b) \oplus \mathbb{O}(b)$. To further simplify notation, let

$$g := \gamma_1 \oplus \underline{\mathbb{C}}, \quad d := \det \gamma_2, \quad B_2 := BU(2), \quad B_1 := BU(1).$$

Diagram (2-8) gives a homotopy commutative diagram

$$\begin{array}{c} \begin{array}{c} \text{---} B_2 \text{---} \\ \text{fold} \swarrow \quad \searrow \text{fold} \\ B_2 \cup_{B_1, g'} B_2 \quad B_2 \cup_{B_1, g} B_2 \\ \text{fold} \cup \text{id} \uparrow \quad \text{id} \cup \text{fold} \uparrow \\ (B_2 \cup_{B_1, g} B_2) \cup_{B_1, g'} B_2 \xleftarrow{\simeq} B_2 \cup_g B_2 \cup_{g'} B_2 \xrightarrow{\simeq} B_2 \cup_{B_1, g} (B_2 \cup_{B_1, g'} B_2) \\ \downarrow \simeq_{\leq 6} \quad \downarrow \simeq_{\leq 6} \\ (B_2 \times_{B_1, d} B_2) \times_{B_1, d} B_2 \quad B_2 \times_{B_1, d} (B_2 \times_{B_1, d} B_2) \\ \swarrow \simeq \quad \downarrow \quad \searrow \simeq \\ B_2 \times_{B_1, d} B_2 \times_{B_1, d} B_2 \end{array} \end{array} \quad \begin{array}{l} x \\ y \end{array}$$

where $\simeq_{\leq 6}$ indicates a map that becomes an equivalence after applying the 6-truncation functor and the dashed arrows x and y are the composites defined after applying τ_6 to the entire diagram. The fact that

$$\pi_0 \text{Maps}_{\text{Top}}(\mathbb{C}P^3, -)(x) \cong \pi_0 \text{Maps}_{\text{Top}}(\mathbb{C}P^3, -)(y)$$

proves the lemma. $\square$

This implies:

Corollary 2.17 *Let V and W be rank 2 vector bundles on $\mathbb{C}P^3$ with $c_1(V) = c_1(W) = a_1$. Then*

$$V +_{a_1, b} W = V +_{a_1} W -_{a_1} (\mathbb{O}(a_1 - b) \oplus \mathbb{O}(b)).$$

Proof This is a special case of a general fact. Suppose that $+$ and $*$ are two abelian group structures on a set S such that, for all $x, y, z \in S$,

$$(x + y) * z = x + (y * z).$$

Let e_+ and e_* be the respective identities. Then

$$(x + e_*) * (y + e_*) = (x + e_*) * (e_* + y) = x + (e_* * e_*) + y = x + y + e_*.$$

Set $x = x + (-e_*)$ and $y = y + (-e_*)$ in the above formula, where $-e_*$ is the additive inverse of e_* . This yields

$$x * y = x - e_* + y - e_* + e_* = x + y - e_*.$$

We get the result by applying this to the set of isomorphism classes of vector bundles with first Chern class a_1 , letting $+$ be $+$ and $*$ be $+$, and noting that $e_* = \mathbb{O}(a_1 - b) \oplus \mathbb{O}(b)$. $\square$

3 Comparing group structures on rank 2 bundles

In the previous section, we defined the group structures $\mathcal{G}_{a_1}$ and $\mathcal{G}_{a_1, b}$ on the set of isomorphism classes of rank 2 bundles on $\mathbb{C}P^3$ with first Chern class equal to a_1 (see Examples 2.9 and 2.8 and Definition 2.10). Our project in this section is to relate these group structures to Horrocks' construction and prove Theorem 1.5. This will involve Atiyah and Rees' classification of rank 2 bundles on $\mathbb{C}P^3$, so we begin in Section 3.1 with a review of relevant material from [2], including discussion of the α -invariant. We also recall Horrocks' construction for locally free sheaves and its necessary properties.

In Section 3.2, we prove that $c_2: \mathcal{G}_{a_1} \rightarrow H^2(\mathbb{C}P^3; \mathbb{Z})$ is a group homomorphism, completing the proof of Proposition 1.3. Verifying additivity of c_2 involves only the definition of the group structure and functoriality results about the group structures of interest, as given in Corollary 2.4. We also show that the α -invariant defines a group homomorphism $\alpha: \mathcal{G}_0 \rightarrow \mathbb{Z}/2$.

In Section 3.3, we show that topological Horrocks addition and Horrocks' construction produce the same underlying topological isomorphism class, when both are defined (Theorem 1.5). We do this by verifying that both constructions have the same effect on complete invariants of isomorphism classes of rank 2 bundles on $\mathbb{C}P^3$.

3.1 The classification of rank 2 bundles on $\mathbb{C}P^3$ and Horrocks' construction

In [2], Atiyah and Rees study complex rank 2 topological vector bundles on $\mathbb{C}P^3$. They define a $\mathbb{Z}/2$ -valued invariant α for bundles with even first Chern class.

Construction 3.1 (the α -invariant [2]) Let V be rank 2 bundle on $\mathbb{C}P^3$ with $c_1 = 0$. Such bundles are classified by $BSU(2)$. The accidental isomorphism $BSU(2) \simeq BSp(1)$ composed with stabilization gives

$$\tilde{\alpha}: BSU(2) \rightarrow BSp \simeq \Omega^\infty \Sigma^4 KO,$$

where KO denotes the real K -theory spectrum. Thus we have a class $\tilde{\alpha} \in KO^4(BSU(2))$ and we define

$$\alpha_0(V) := p_* V^*(\tilde{\alpha}),$$

where p_* is the KO -theory pushforward for the spin manifold $\mathbb{C}P^3$.

Atiyah and Rees extend α to all bundles with $c_1(V) \equiv 0 \pmod{2}$ by letting

$$(3-1) \quad \alpha(V) := \alpha_0(V \otimes \mathbb{O}(-\tfrac{1}{2}c_1(V))).$$

Atiyah and Rees give various formulas for computing the α -invariant of special bundles, eg:

Proposition 3.2 [2, Theorem 7.2] For V a rank 2 bundle on $\mathbb{C}P^3$ with $c_1(V) \equiv 0 \pmod{2}$, let

$$\Delta(V) := \tfrac{1}{4}(c_1^2 - 4c_2).$$

If V extends to $\mathbb{C}P^4$, then $\alpha(V) \equiv \tfrac{1}{12}\Delta(\Delta - 1) \pmod{2}$.

The utility of the α -invariant is as follows:

Theorem 3.3 [2, Theorems 1.1, 2.8 and 3.3] Given $a_1, a_2 \in \mathbb{Z}$ with $a_1 a_2 \equiv 0 \pmod{2}$, the number of isomorphism classes of rank 2 bundles on $\mathbb{C}P^3$ with i^{th} Chern class a_i is

- equal to 2 if $a_1 \equiv 0 \pmod{2}$, and
- equal to 1 otherwise.

In the first case, a rank 2 vector bundle on $\mathbb{C}P^3$ is determined by c_1, c_2 and α .

Remark 3.4 The condition $a_1 a_2 \equiv 0 \pmod{2}$ is necessary and sufficient for two integers to be the Chern classes of complex rank 2 topological bundles on $\mathbb{C}P^3$.

We now review Horrocks' construction for rank 2 algebraic vector bundles.

Construction 3.5 (Horrocks' construction [5]) Let V_1 and V_2 be rank 2 locally free sheaves on $\mathbb{C}P^3$. Suppose that

- we have isomorphisms $\wedge^2 V_1 \simeq \wedge^2 V_2 \simeq \mathbb{O}(a)$ for some $a \in \mathbb{Z}$;
- we have regular¹ sections $s_i: \mathbb{O} \rightarrow V_i^*$; and
- the sheaves $\mathcal{R}_i := \text{coker}(s_i^*: \mathcal{F}_i \rightarrow \mathbb{O})$ have disjoint supports.

¹A section s of V is regular if its vanishing locus is of codimension equal to the rank of V .

For each $i = 1, 2$, the Koszul complex relative to s_i has the form $0 \rightarrow \mathcal{O}(a) \rightarrow V_i \rightarrow \mathcal{O}$ and is exact since s_i is regular. By definition, we have exact sequences

$$0 \rightarrow \mathcal{O}(a) \rightarrow V_i \rightarrow \mathcal{O} \rightarrow \mathcal{R}_i \rightarrow 0.$$

The sheaf $V_1 +_H V_2$ is defined by the diagram

$$(3-2) \quad \begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{O}(a) \oplus \mathcal{O}(a) & \longrightarrow & V_1 \oplus V_2 & \longrightarrow & \mathcal{O} \oplus \mathcal{O} & \longrightarrow & \mathcal{R}_1 \oplus \mathcal{R}_2 & \longrightarrow & 0 \\ & & \uparrow \simeq & & \uparrow & \lrcorner & \uparrow \Delta & & \uparrow \simeq & & \\ 0 & \longrightarrow & \mathcal{O}(a) \oplus \mathcal{O}(a) & \longrightarrow & \mathcal{W} & \longrightarrow & \mathcal{O} & \longrightarrow & \mathcal{R}_1 \oplus \mathcal{R}_2 & \longrightarrow & 0 \\ & & \downarrow \nabla & & \downarrow & & \downarrow \simeq & & \downarrow \simeq & & \\ 0 & \longrightarrow & \mathcal{O}(a) & \longrightarrow & V_1 +_H V_2 & \longrightarrow & \mathcal{O} & \longrightarrow & \mathcal{R}_1 \oplus \mathcal{R}_2 & \longrightarrow & 0 \end{array}$$

where $\mathcal{W}$ is the indicated pullback along the diagonal Δ and $V_1 +_H V_2$ is the pushout along the fold map ∇ . The middle and bottom rows in (3-2) are both exact and $V_1 +_H V_2$ is locally free of rank 2 [5, Theorem 1]. The bottom exact sequence of (3-2) implies that

$$c_1(V_1 +_H V_2) = c_1(V_1) = c_1(V_2) = a.$$

Atiyah and Rees and Horrocks study the effect of $+_H$ on algebraic invariants:

Theorem 3.6 [5, Theorem 1; 2, Corollary 5.7] *Let V_1 and V_2 be rank 2 algebraic bundles on $\mathbb{C}P^3$ with $c_1(V_1) = c_1(V_2) = -m$, with $m \geq 0$.² Then*

$$c_2(V_1 +_H V_2) = c_2(V_1) + c_2(V_2).$$

Furthermore, suppose that $m = 2n$ with $n \geq 0$.

- If n is odd or $n \equiv 0 \pmod{4}$, then $\alpha(V_2 +_H V_2) = \alpha(V_1) + \alpha(V_2) \in \mathbb{Z}/2\mathbb{Z}$.
- If $n \equiv 2 \pmod{4}$, then $\alpha(V_1 +_H V_2) = \alpha(V_1) + \alpha(V_2) + 1$.

3.2 Proof of Proposition 1.3

In the previous section, we defined the group structures $\mathcal{G}_{a_1}$ on the set of isomorphism classes of rank 2 bundles on $\mathbb{C}P^3$ with first Chern class equal to a_1 (see Example 2.9 and Definition 2.10). This gives the group structure required by part (i) of Proposition 1.3.

For part (ii), note that, for any $f: X \rightarrow Y$ with section $s: X \rightarrow Y$ satisfying the hypotheses of Proposition 1.12, and any $c: C \rightarrow Y$ with C an m -skeletal space, the identity element is given by $s \circ c$. In this case, $s \circ c \simeq \mathcal{O}(a_1) \oplus \underline{\mathbb{C}}$.

Part (iii) is a consequence of the following result:

²This is necessary for V_i to admit regular sections.

Proposition 3.7 Let $V, W: \mathbb{C}P^3 \rightarrow BU(2)$ be two bundles with $\det V \simeq \det W \simeq \mathbb{O}(a_1)$. Let $+_{a_1}$ denote topological Horrocks addition as in [Definition 2.10](#). Then

- for any $a \in \mathbb{Z}$, the second Chern class c_2 is a homomorphism for $+_{a_1}$; and
- if $a_1 = 0$, then α is a homomorphism for $+_0$.

Proof Let $K(\mathbb{Z}, 4)$ denote the Eilenberg–Mac Lane space that is determined by $\pi_4(K(\mathbb{Z}, 4)) = \mathbb{Z}$ and $\pi_i(K(\mathbb{Z}, 4)) = 0$ for $i \neq 4$. Note that the diagram

$$\begin{array}{ccc} BU(2) & \xrightarrow{c_2} & K(\mathbb{Z}, 4) \\ \downarrow \text{det} & & \downarrow \\ BU(1) & \xrightarrow{0} & * \end{array}$$

The solid diagram is homotopy commutative and the subdiagram with only dashed vertical arrows is homotopy commutative. By [Corollary 2.6](#), we get a group homomorphism $c_2: [\mathbb{C}P^3, BU(2)]_c \rightarrow H^4(\mathbb{C}P^3, \mathbb{Z})$.

For the second item, consider the diagram

$$(3-3) \quad \begin{array}{ccccc} BU(2) & \xleftarrow{\iota} & BSU(2) & \xrightarrow{\alpha} & \Omega^\infty \Sigma^4 KO \\ \downarrow c_1 & & \downarrow & & \downarrow \\ K(\mathbb{Z}, 2) & \xleftarrow{\quad} & * & \xrightarrow{\quad} & * \end{array}$$

where the solid diagram is homotopy commutative and the subdiagram with only dashed vertical arrows is homotopy commutative. All vertical homotopy fibers are 4-connective, so we can apply [Corollary 1.10](#) to all vertical diagrams to obtain group structures on homotopy classes of maps from $\mathbb{C}P^3$ into each diagram. By [Corollary 2.6](#), we get group homomorphisms

$$(3-4) \quad [\mathbb{C}P^3, BU(2)]_{c_1=0} \xleftarrow{\iota \circ (-)} [\mathbb{C}P^3, BSU(2)] \xrightarrow{\alpha \circ (-)} [\mathbb{C}P^3, \Omega^\infty \Sigma^4 KO] \simeq KO^4(\mathbb{C}P^3).$$

The map $\iota \circ -$ is an isomorphism. For $V: \mathbb{C}P^3 \rightarrow BU(2)$ with $c_1(V) = 0$,

$$p_*(\alpha \circ ((\iota \circ (-))^{-1}(V))) = \alpha(V).$$

Since pushforward on cohomology is a group homomorphism, we conclude that α is additive for $+_0$, as was to be shown. $\square$

3.3 Proof of [Theorem 1.5](#)

To show that $V +_{a_1} W \simeq V +_H W$ when both $+_{a_1}$ and $+_H$ are defined, we check that $V +_{a_1} W$ and $V +_H W$ have the same Chern classes and α -invariant. Both $+_{a_1}$ and $+_H$ fix the first Chern class, so it suffices to show that

- $c_2(V +_{a_1} W) = c_2(V +_H W)$, and
- $\alpha(V +_{a_1} W) = \alpha(V +_H W)$.

Item (i) follows from additivity of c_2 for both operations, by [Theorem 3.6](#) and [Proposition 3.7](#).

Checking (ii) is complicated by the fact that the α -invariant does not play well with $+_{a_1}$ when $a_1 \neq 0$. We bootstrap from the case $a_1 = 0$ to obtain a formula for $\alpha(V +_{a_1} W)$ in terms of $\alpha(V)$ and $\alpha(W)$. This involves the study of the groups $\mathcal{G}_{0,b}$ for b nonzero (see [Example 2.8](#)).

Let $a_1 \in \mathbb{Z}$ and suppose that V and W are rank 2 bundles on $\mathbb{C}P^3$ with $c_1(V) = c_1(W) = a_1$. By [Proposition 1.3\(iii\)](#), $c_2(V +_{a_1} W) = c_2(V) + c_2(W) = c_2(V +_H W)$. In the case $a_1 \equiv 1 \pmod{2}$, [Theorem 3.3](#) implies $V +_{a_1} W \simeq V +_H W$ and we are done.

Now suppose $a_1 \equiv 0 \pmod{2}$. Recall [Construction 2.11](#) and the associated group isomorphism $\phi: \mathcal{G}_{a_1} \rightarrow \mathcal{G}_{0,b}$. Using the definition of α (see (3-1)), [Proposition 3.7](#) and [Corollary 2.17](#), we have

$$\begin{aligned} \alpha(V +_{a_1} W) &= \alpha(\phi(V +_{a_1} W)) \\ &= \alpha(\phi(V) +_{0,b} \phi(W)) \\ &= \alpha(\phi(V) +_0 \phi(W) -_0 \mathbb{O}(-b) \oplus \mathbb{O}(b)) \\ &= \alpha(V) + \alpha(W) - \alpha(\mathbb{O}(-b) \oplus \mathbb{O}(b)). \end{aligned}$$

To compute $\alpha(\mathbb{O}(b) \oplus \mathbb{O}(-b))$, we use [Proposition 3.2](#). Since $\mathbb{O}(-b) \oplus \mathbb{O}(b)$ extends to $\mathbb{C}P^4$ and $\Delta(\mathbb{O}(b) \oplus \mathbb{O}(-b)) = b^2$, we see that

$$\alpha(\mathbb{O}(-b) \oplus \mathbb{O}(b)) \equiv \frac{1}{12} b^2 (b^2 - 1) \equiv \frac{1}{4} b^2 (b + 1)(b - 1) \pmod{2}.$$

So we must determine when $b^2(b + 1)(b - 1)$ is divisible by 8.

- If $b \equiv 0 \pmod{4}$, then 16 divides $b^2(b + 1)(b - 1)$, so $\alpha(\mathbb{O}(b) \oplus \mathbb{O}(-b)) = 0$.
- If b is odd, then both $b + 1$ and $b - 1$ are even and one of $b + 1$ or $b - 1$ is divisible by 4, so $\alpha(\mathbb{O}(b) \oplus \mathbb{O}(-b)) = 0$.
- If $b \equiv 2 \pmod{4}$, then both $b + 1$ and $b - 1$ are odd and b^2 is divisible by 4 but not 8, so $\alpha(\mathbb{O}(b) \oplus \mathbb{O}(-b)) = 1$.

Thus,

$$\alpha(V +_a W) = \alpha(V) + \alpha(W) + \epsilon(a), \quad \text{where} \quad \epsilon(a) := \begin{cases} 0 & \text{if } a \not\equiv 4 \pmod{8}, \\ 1 & \text{if } a \equiv 4 \pmod{8}. \end{cases}$$

Comparing this with [Theorem 3.6](#) yields the conclusion $\alpha(V +_{a_1} W) = \alpha(V +_H W)$. □

While this completes the argument that algebraic and topological Horrocks addition produce the same underlying topological isomorphism class, our method is indirect. The reader might hope for a more geometric relationship between the constructions.

Problem 3.8 Is there an explicit, bundle-theoretic comparison between topological Horrocks addition $+_{a_1}$ and Horrocks' original construction for locally free sheaves?

4 Group structures on rank 3 bundles on $\mathbb{C}P^5$

We now explore additive structures on rank 3 bundles on $\mathbb{C}P^5$. Recall that, given a fixed rank 2 bundle V_0 on $\mathbb{C}P^5$, [Example 2.12](#) gives a group $\mathcal{G}_{V_0}$ of isomorphism classes of rank 3 vector bundles with the same first and second Chern classes as V_0 . The identity is $V_0 \oplus \underline{\mathbb{C}}$. Our primary interest here is to better understand the groups $\mathcal{G}_{V_0}$ and their properties.

We complete the proof of [Theorem 1.6](#) in [Section 4.1](#) by showing that c_3 is a group homomorphism. In [Section 4.2](#), we explore the structure of $\mathcal{G}_{V_0}$ for $V_0 = \mathcal{O}(a) \oplus \mathcal{O}(b)$. We show that $+_{V_0}$ allows for the construction of interesting bundles from simple ones, as did Horrocks' construction on rank 2 bundles on $\mathbb{C}P^3$ (see [Theorem 1.1](#)).

Proposition 4.1 *For infinitely many isomorphism classes of rank 2 bundles $V_0 := \mathcal{O}(a) \oplus \mathcal{O}(b)$ on $\mathbb{C}P^5$, there exists a bundle $W = \mathcal{O}(x) \oplus \mathcal{O}(y) \oplus \mathcal{O}(z)$ with $W \in \mathcal{G}_{V_0}$ such that the subgroup generated by W contains bundles that are not sums of line bundles. Moreover, the subgroup generated by W is of finite index in $\mathcal{G}_{V_0}$.*

The proof of this result is elementary, involving only the study of possible Chern classes of sums of line bundles and additivity of c_3 .

4.1 The third Chern class and the structure of $\mathcal{G}_{V_0}$ for rank 3 bundles

We first prove the additivity of $+_{V_0}$ for c_3 and complete a theorem stated in the introduction.

Proof of Theorem 1.6 We have already established the group structure in [Example 2.12](#). The identity element is $V_0 \oplus \underline{\mathbb{C}}$ by [Corollary 1.10](#).

To show that the third Chern class c_3 is a homomorphism, consider the diagram

$$\begin{array}{ccc} \widetilde{BU}(3) & \xrightarrow{c_3} & K(\mathbb{Z}, 6) \\ \downarrow s & & \downarrow \\ BU(2) & \xrightarrow{0} & * \end{array}$$

where the solid diagram is homotopy commutative and the subdiagram with only dashed vertical arrows is homotopy commutative. By [Corollary 2.6](#), we get a group homomorphism $c_3: [\mathbb{C}P^3, BU(3)]_c \rightarrow H^6(\mathbb{C}P^5, \mathbb{Z})$. $\square$

In [\[11\]](#), we prove that rank 3 bundles on $\mathbb{C}P^5$ are determined by Chern classes unless $c_1(V) \equiv c_2(V) \equiv 0 \pmod{3}$, in which case there are three pairwise nonisomorphic bundles with the same Chern classes as V . This proves the following result:

Corollary 4.2 The kernel of $c_3: \mathcal{G}_{V_0} \rightarrow \mathbb{Z}$ is

- (1) trivial if $c_1(V_0) \not\equiv 0 \pmod{3}$ or $c_2(V_0) \not\equiv 0 \pmod{3}$, and
- (2) isomorphic to $\mathbb{Z}/3$ if $c_1(V_0) \equiv c_2(V_0) \equiv 0 \pmod{3}$.

Remark 4.3 The image of c_3 is an infinite subgroup of $\mathbb{Z}$, since the conditions for three integers to be the Chern classes of a rank 3 bundle on $\mathbb{C}P^5$ are a finite number of congruences (see [11, Lemma 2.16]). In the first case in Corollary 4.2, $\mathcal{G}_{V_0}$ is abstractly isomorphic to $\mathbb{Z}$; in the second, $\mathcal{G}_{V_0} \simeq \mathbb{Z} \oplus \mathbb{Z}/3$.

In [11], we define a $\mathbb{Z}/3$ -valued invariant ρ for rank 3 bundles V on $\mathbb{C}P^5$ such that $c_1(V) \equiv c_2(V) \equiv 0 \pmod{3}$. We prove that c_1, c_2, c_3 and ρ are complete invariants of isomorphism classes of complex rank 3 topological bundles on $\mathbb{C}P^5$. Additivity of c_3 leads to the following:

Problem 4.4 For general V_0 with first and second Chern classes divisible by 3, what is a formula for $\rho(V +_{V_0} W)$ in terms of $\rho(V)$, $\rho(W)$ and $\rho(V_0)$?

Note that a necessary condition for ρ to be a group homomorphism is that $\rho(V_0 \oplus \underline{\mathbb{C}}) = 0$.

4.2 Constructing rank 3 bundles on $\mathbb{C}P^5$ from sums of line bundles

We begin this subsection by showing that there exist infinitely many isomorphism classes $V_0 = \mathbb{O}(a) \oplus \mathbb{O}(b)$ on $\mathbb{C}P^3$ such that $\mathcal{G}_{V_0}$ contains another sum of line bundles that is not the group identity. This is a preliminary to prove Proposition 4.1. First, we simplify notation.

Definition 4.5 Given $n_1, \dots, n_k \in \mathbb{Z}$, let $\mathbb{O}(n_1, \dots, n_k) := \mathbb{O}(n_1) \oplus \dots \oplus \mathbb{O}(n_k)$.

An element $\mathbb{O}(x, y, z) \in \mathcal{G}_{\mathbb{O}(a,b)}$ is equivalent to an integer solution to the equations

$$(4-1) \quad x + yz = a + b, \quad xy + yz + zx = ab.$$

Note that, since the identity of $\mathcal{G}_{\mathbb{O}(a,b)}$ is $\mathbb{O}(a, b) \oplus \underline{\mathbb{C}}$, a sum of line bundles $\mathbb{O}(x, y, z) \in \mathcal{G}_{\mathbb{O}(a,b)}$ is a nonidentity element if and only if $c_3(\mathbb{O}(x, y, z)) = xyz$ is nonzero.

Let $c := a - x$ and $d := b - y$. Substituting in (4-1), we obtain

$$Q := c^2 + d^2 - bd - ac + cd = 0.$$

This is the equation of a quadric hypersurface in $\mathbb{P}_{a,b,c,d}^3$, whose rational points can be found via standard methods. For any $u, l, v, w \in \mathbb{Z}$, we get solutions to (4-1)

$$\begin{aligned} a &= w(u^2 + v^2 + uv - lv), & x &= w(v^2 + uv - lv), \\ b &= wul, & y &= wu(l - v), \\ & & z &= wu(u + v). \end{aligned}$$

Remark 4.6 The previous computation shows that there are infinitely many $\mathcal{G}_{V_0}$, where $V_0 = \mathbb{O}(a, b)$, such that $W = \mathbb{O}(x, y, z)$ is a nonidentity element of $\mathcal{G}_{V_0}$.

Example 4.7 Consider $W = \mathcal{O}(2, -1, 2)$ and $V_0 = \mathcal{O}(3, 0)$. By construction, $W \in \mathcal{G}_{V_0}$. Since c_1 is even and c_1 and c_2 are both divisible by 3, the Schwarzenberger conditions as in [11, Lemma 2.16] imply that there exists a bundle $W' \in \mathcal{G}_{V_0}$ with $c_3(W') = a$ if and only if a is even. By Remark 4.3, $\mathcal{G}_{V_0} \simeq \mathbb{Z} \oplus \mathbb{Z}/3$ and, under this identification, c_3 is projection onto the first factor. Since $c_3(W) = 4$, this implies that the subgroup generated by W is of index 6.

Proof of Proposition 4.1 Let $V_0 = \mathcal{O}(a, b)$ and $W = \mathcal{O}(x, y, z)$ be as in Remark 4.6 with $c_3(W) \neq 0$. Inductively, let

$$+_{V_0}^n W := W +_{V_0} (+_{V_0}^{n-1} W).$$

Assume for a contradiction that $+_{V_0}^n W$ is a sum of line bundles for all positive integers n . Let $a_1, a_2, a_3 \in \mathbb{Z}$ denote the Chern classes of W . Since $+_{V_0}$ preserves c_1 and c_2 , and $c_3: \mathcal{G}_{V_0} \rightarrow H^6(\mathbb{C}P^5; \mathbb{Z})$ is a group homomorphism,

$$c_1(+_{V_0}^n W) = a_1, \quad c_2(+_{V_0}^n W) = a_2 \quad \text{and} \quad c_3(+_{V_0}^n W) = na_3.$$

Thus, the equations

$$X + Y + Z = a_1, \quad XYZ = na_3, \quad XY + YX + ZX = a_2$$

have a solution $(X, Y, Z) = (x_n, y_n, z_n)$ for all $n \in \mathbb{Z}$. Let p be a prime such that $p > |3a_i|$ for $i = 1, 3$. Take $n = p$ in the above equations. Since $c_3(W) = a_3 \neq 0$, the third equation implies that p divides one of x_p, y_p or z_p . Without loss of generality, p divides x_p . But note that

$$|z_p y_p| \leq |a_3| < \frac{1}{3}p$$

and so $|z_p| < \frac{1}{3}p$, and $|y_p| < \frac{1}{3}p$. Since $|x_p|$ is divisible by p , it is larger than $|y_p|, |z_p|$ and $|y_p + z_p|$. Therefore,

$$|x_p + y_p + z_p| \geq |x_p| - |y_p| - |z_p| \geq \frac{1}{3}p.$$

On the other hand, $|x_p + y_p + z_p| = |a_1| < \frac{1}{3}p$, a contradiction.

Now consider the subgroup generated W . Since $c_3(W) \neq 0$, $\ker(c_3) \cap \langle W \rangle$ is trivial and we have an exact sequence of groups

$$\ker(c_3) \rightarrow \mathcal{G}_{V_0}/\langle W \rangle \rightarrow \mathbb{Z}/c_3(W).$$

Finally, $\ker(c_3)$ is finite by Remark 4.3 and $\mathbb{Z}/c_3(W)$ is finite, so $\mathcal{G}_{V_0}/H$ is also finite. □

This section gives insight into the groups $\mathcal{G}_{V_0}$, but many questions remain.

Problem 4.8 (1) For groups $\mathcal{G}_{V_0} \simeq \mathbb{Z} \oplus \mathbb{Z}/3$, what are generators for 3-torsion?

(2) Are the Horrocks bundles of rank 3 on $\mathbb{C}P^5$ elements of some $\mathcal{G}_{V_0}$? If so, what subgroups do they generate?

(3) What can be said about the structure of $\mathcal{G}_{V_0}$ for V_0 not a sum of line bundles?

Appendix The pointed Grothendieck construction for spaces

This appendix describes the pointed version straightening and unstraightening for functors to spaces, combining facts from [9]. The particular goal is to make precise the statement that a diagram of spaces

$$(A-1) \quad X \begin{array}{c} \xrightarrow{f} \\ \xleftarrow[s]{\quad} \end{array} Y$$

where f is n -connective can equivalently be viewed as a functor from Y to infinite loop spaces, after $(2n-2)$ -truncation.

Notation A.1 We establish terminology and conventions and provide references.

- (1) Let $\mathbf{sSet}$ denote the category of simplicial sets and $\mathbf{Cat}_\Delta$ denote the category of simplicially enriched categories.
- (2) By $(\infty, 1)$ -category, we will mean a simplicial set that satisfies the inner horn condition (a quasi-category [9, 1.1.2.4]).
- (3) Let $\mathfrak{C}$ denote the left adjoint to the homotopy coherent nerve functor $N : \mathbf{Cat}_\Delta \rightarrow \mathbf{sSet}$ [9, 1.1.5]. There is a Quillen equivalence of simplicial model categories

$$\mathbf{Cat}_\Delta \begin{array}{c} \xleftarrow{\mathfrak{C}} \\ \xrightarrow{N} \end{array} \mathbf{sSet}$$

between simplicial sets with the Joyal model structure and simplicially enriched categories with the Bergner model structure (see [9, Section 1.1.5] for discussion, and [4; 3; 8] for a proof).

- (4) Let $\mathbf{sSet}/_S$ denote the simplicial model category of simplicial sets over S , endowed with the contravariant model structure. We have an associated simplicial category $\mathbf{RFib}(S)$ obtained by taking fibrant–cofibrant objects.
- (5) Let $\mathbf{Kan}$ denote the $(\infty, 1)$ -category of Kan complexes obtained by taking fibrant–cofibrant objects in the simplicial model category of simplicial sets with the Kan model structure and applying N . Given a Kan complex S , let $\mathbf{Kan}/_S$ denote the $(\infty, 1)$ -overcategory of Kan complexes over S .
- (6) Given a $(\infty, 1)$ -category C and a simplicial set S , we write $\mathbf{Fun}(S, C)$ for the simplicial set of maps from S to C . This is an $(\infty, 1)$ -category since C is, and models functors from S to C [9, 1.2.7.2].
- (7) We ignore set-theoretic issues and refer the concerned reader to [9, 1.2.15].

The fundamental result that we need is the following:

Theorem A.2 (Lurie [9]) *Let S be a Kan complex. There is an equivalence of $(\infty, 1)$ -categories*

$$(A-2) \quad \mathbf{Kan}/_S \begin{array}{c} \xleftarrow{\text{Un}} \\ \xrightarrow{\text{St}} \end{array} \mathbf{Fun}(S, \mathbf{Kan}).$$

Proof By [9, 2.2.3.11], for any simplicial set we have an equivalence of simplicial categories

$$(A-3) \quad \text{RFib}(S) \begin{array}{c} \xleftarrow{\text{Un}} \\ \xrightarrow{\text{St}} \end{array} (\text{sSet}^{\text{c}[S]^{\text{op}}})^{\circ},$$

where sSet has the Kan model structure, $\text{sSet}^{\text{c}[S]^{\text{op}}}$ has the projective model structure, and $(-)^{\circ}$ indicates we take fibrant-cofibrant objects. Since S is a Kan complex, [9, 3.1.5.1(A3)] implies $\text{Kan}_S \simeq N(\text{RFib}(S))$ as $(\infty, 1)$ -categories.

By [9, 4.2.4.4], applying the homotopy coherent nerve to the right side of (A-3) recovers $\text{Fun}(S^{\text{op}}, \text{Kan})$. By [12, Section 57] combined with [12, Proposition 14.14], every Kan complex is equivalent to its opposite as an $(\infty, 1)$ -category. $\square$

To obtain a version for spaces over a base with a section, we take pointed objects on either side of Theorem A.2.

Definition A.3 Given an $(\infty, 1)$ -category C with final object t [9, 1.2.12], let C_* denote the full $(\infty, 1)$ -subcategory of $\text{Fun}(\Delta^1, C)$ spanned by functors that restrict to t on 0. Explicitly, if fib_t denotes the (1-categorical) fiber in sSet at t , then

$$C_* := \text{fib}_t(\text{Fun}(\Delta^1, C) \rightarrow \text{Fun}(\Delta^0, C)),$$

where the map is given by precomposition with $\{0\} \simeq \Delta^0 \hookrightarrow \Delta^1$.

Remark A.4 Given an $(\infty, 1)$ -category C with final object t , the infinity category C_* is also modeled by the $(\infty, 1)$ -undercategory $C_{t/}$. See [9, 1.2.9] for a discussion of $(\infty, 1)$ -undercategories and [9, 4.2.1.5 and 7.2.2.8] for the equivalence.

Since equivalent $(\infty, 1)$ -categories support the same category theory, the equivalences St and Un will induce one on pointed objects.

Lemma A.5 An equivalence of $(\infty, 1)$ -categories induces an equivalence upon taking pointed objects.

Proof This is the dual statement of [9, 1.2.9.3] (dualized, for example, following the method of [9, 1.2.9.5]), specialized to the case of a final object. $\square$

We can apply Lemma A.5 directly to the equivalences of Theorem A.2, but to understand the result we first describe pointed objects on the left side of (A-2).

Lemma A.6 For any $(\infty, 1)$ -category C with a final object t , the $(\infty, 1)$ -category $\text{Fun}(S, C)$ has as final object the constant functor at t . There is an equivalence of $(\infty, 1)$ -categories $\text{Fun}(S, C)_* \simeq \text{Fun}(S, C_*)$.

$$\mathrm{Fun}(\Delta^1, \mathrm{Fun}(S, C)) \rightarrow \mathrm{Fun}(\Delta^0, \mathrm{Fun}(S, C))$$
$$\begin{aligned} \mathrm{fib}(\mathrm{Fun}(\Delta^1, \mathrm{Fun}(S, C)) \rightarrow \mathrm{Fun}(\Delta^0, \mathrm{Fun}(S, C))) &\simeq \mathrm{fib}(\mathrm{Fun}(\Delta^1 \times S, C) \rightarrow \mathrm{Fun}(\Delta^0 \times S, C)) \\ &\simeq \mathrm{fib}(\mathrm{Fun}(S, \mathrm{Fun}(\Delta^1, C)) \rightarrow \mathrm{Fun}(S, \mathrm{Fun}(\Delta^0, C))). \end{aligned}$$
$$\mathrm{fib}(\mathrm{Fun}(S, \mathrm{Fun}(\Delta^1, C)) \rightarrow \mathrm{Fun}(S, \mathrm{Fun}(\Delta^0, C))) \simeq \mathrm{Fun}(S, \mathrm{fib}(\mathrm{Fun}(\Delta^1, C) \rightarrow \mathrm{Fun}(\Delta^0, C))).$$
☐

Lemma A.8 *Let $\mathcal{C}$ be a $(\infty, 1)$ -category with a final object t and with coproducts. Then there is an adjunction of $(\infty, 1)$ -categories*

$$\begin{array}{ccc} & \text{forget} & \\ C & \xleftarrow{\quad \top \quad} & C_* \\ & \xrightarrow{\quad + \quad} & \end{array}$$

$$u: \text{Id}_C \rightarrow F \circ +$$
$$\text{Maps}_{C_*}(t \rightarrow c \sqcup t, t \rightarrow y) \xrightarrow{\text{apply } F} \text{Maps}_C(F(t \rightarrow c \sqcup t), F(t \rightarrow y)) \xrightarrow{u_c \circ -} \text{Maps}_C(c, y)$$
$$\mathrm{Maps}_{C^*}(t \rightarrow c \sqcup t, t \rightarrow y) \simeq \mathrm{hofib}(\mathrm{Maps}_C(c \sqcup t, y) \rightarrow \mathrm{Maps}_C(t, y)),$$

Algebraic & Geometric Topology, Volume 25 (2025)

commutative diagram of spaces

$$\begin{array}{ccc} \mathrm{Maps}_C(c \sqcup t, y) & \longrightarrow & \mathrm{Maps}_C(t, y) \\ \downarrow \simeq & & \parallel \\ \mathrm{Maps}_C(c, y) \times \mathrm{Maps}_C(t, y) & \xrightarrow{\pi_2} & \mathrm{Maps}_C(t, y) \end{array}$$

This implies

$$\mathrm{Maps}_{C_*}(t \rightarrow c \sqcup t, t \rightarrow y) \simeq \mathrm{hofib}(\pi_2: \mathrm{Maps}_C(c, y) \times \mathrm{Maps}_C(t, y) \rightarrow \mathrm{Maps}_C(t, y)) \simeq \mathrm{Maps}_C(c, y).$$

We have a homotopy commutative diagram

$$\begin{array}{ccccc} \mathrm{Maps}_{C_*}(t \rightarrow c \sqcup t, t \rightarrow y) & \xrightarrow{\simeq} & Z & \xrightarrow{\simeq} & \mathrm{Maps}_C(c, y) \\ \text{apply } F \downarrow & & \downarrow & \swarrow 1 \times (t \rightarrow y) & \\ \mathrm{Maps}_C(c \sqcup t, y) & \xrightarrow{\simeq} & \mathrm{Maps}_C(c, y) \times \mathrm{Maps}_C(t, y) & & \\ u_c \circ - \downarrow & & \pi_1 \downarrow & \searrow 1 & \\ \mathrm{Maps}_C(c, y) & \xlongequal{\quad} & \mathrm{Maps}_C(c, y) & & \end{array}$$

where $Z := \mathrm{fib}((\mathrm{Maps}_C(c, y) \times \mathrm{Maps}_C(t, y)) \xrightarrow{\pi_2} \mathrm{Maps}_C(t, y))$. The rightmost map is a weak homotopy equivalence, as are all the horizontal arrows, so the left vertical composite is too. $\square$

References

- [1] **A Asok, J Fasel**, *Euler class groups and motivic stable cohomotopy*, J. Eur. Math. Soc. 24 (2022) 2775–2822 [MR](#)
- [2] **M F Atiyah, E Rees**, *Vector bundles on projective 3-space*, Invent. Math. 35 (1976) 131–153 [MR](#)
- [3] **J E Bergner**, *A model category structure on the category of simplicial categories*, Trans. Amer. Math. Soc. 359 (2007) 2043–2058 [MR](#)
- [4] **J E Bergner**, *A survey of $(\infty, 1)$ -categories*, from “Towards higher categories” (J C Baez, J P May, editors), IMA Vol. Math. Appl. 152, Springer (2010) 69–83 [MR](#)
- [5] **G Horrocks**, *A construction for locally free sheaves*, Topology 7 (1968) 117–120 [MR](#)
- [6] **G Horrocks**, *Examples of rank three vector bundles on five-dimensional projective space*, J. London Math. Soc. 18 (1978) 15–27 [MR](#)
- [7] **G Horrocks, D Mumford**, *A rank 2 vector bundle on P^4 with 15 000 symmetries*, Topology 12 (1973) 63–81 [MR](#)
- [8] **A Joyal**, *Quasi-categories vs simplicial categories*, unpublished (2007) [https://math.uchicago.edu/~may/IMA/Incoming/Joyal/QvsDJan9\(2007\).pdf](https://math.uchicago.edu/~may/IMA/Incoming/Joyal/QvsDJan9(2007).pdf)
- [9] **J Lurie**, *Higher topos theory*, Annals of Mathematics Studies 170, Princeton Univ. Press (2009) [MR](#) [Zbl](#)
- [10] **A Mathew, V Stojanoska**, *The Picard group of topological modular forms via descent theory*, Geom. Topol. 20 (2016) 3133–3217 [MR](#)

- [11] **M Opie**, *A classification of complex rank 3 vector bundles on $\mathbb{C}P^5$* , Adv. Math. 455 (2024) art. id. 109878 [MR](#)
- [12] **C Rezk**, *Introduction to quasicategories*, lecture notes (2022) <https://rezk.web.illinois.edu/quasicats.pdf>
- [13] **E Spanier**, *Borsuk's cohomotopy groups*, Ann. of Math. 50 (1949) 203–245 [MR](#)

Department of Mathematics, UCLA
Los Angeles, CA, United States

mopie@math.ucla.edu

<https://www.math.ucla.edu/~mopie/>

Received: 6 November 2023 Revised: 30 November 2023

Power sum elements in the G_2 skein algebra

BODIE BEAUMONT-GOULD

ERIK BRODSKY

VIJAY HIGGINS

ALAINA HOGAN

JOSEPH M MELBY

JOSHUA PIAZZA

We study the skein algebras of surfaces associated to the exceptional Lie group G_2 , using Kuperberg webs. We identify two 2-variable polynomials, $P_n(x, y)$ and $Q_n(x, y)$, and use threading operations along knots to construct a family of central elements in the G_2 skein algebra of a surface, $S_q^{G_2}(\Sigma)$, when the quantum parameter q is a $2n^{\text{th}}$ root of unity. We verify these elements are central using elementary skein-theoretic arguments. We also obtain a result about the uniqueness of the so-called transparent polynomials P_n and Q_n . Our methods involve a detailed study of the skein modules of the annulus and the twice-marked annulus.

[57K31](#); [17B37](#)

1 Introduction

The subject of skein theory lies at the intersection of low-dimensional topology and the representation theory of quantum groups. In skein theory, one works with linear relations involving pictures of skeins, which are hybrid objects that can be readily interpreted either in the language of topology or in the language of algebra. The simplest examples of these skein relations are the Kauffman bracket skein relations:

$$\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} = q \left(+ q^{-1} \begin{array}{c} \diagup \diagdown \\ \diagup \diagdown \end{array} \right), \quad \bigcirc = -q^2 - q^{-2}.$$

On one hand, these relations can be used to define link invariants, like the Jones polynomial, or to construct the 3-manifold invariants of Witten, Reshetikhin, and Turaev; see Lickorish [18, Chapter 13]. On the other hand, these relations faithfully describe the representation theory of quantum SL_2 ; see eg Kuperberg [15].

We are interested in the skein theory associated to the rank-2 exceptional Lie group, G_2 . In this setting, skeins are certain graphs called *webs*. The skein theory for G_2 webs, and for the cases of other rank-2 Lie groups, was developed in the seminal work of Kuperberg [14; 15]. Webs for higher-rank type A_n were introduced by Murakami, Ohtsuki, and Yamada [23] for the purpose of a graphical definition of the colored HOMFLY polynomial. Later, it was shown by Cautis, Kamnitzer, and Morrison [6] that the higher-rank webs and their skein relations provide a full description of the representation category of

quantum SL_d . Finding explicit skein relations which faithfully describe the representation categories of Lie groups for all Dynkin types is an active area of research; see Bodish, Elias, Rose, and Tatham [2] and Savage and Westbury [28] for recent examples. Many of these skein theories are first developed in the setting where q is a generic parameter. One can also study whether the same skein relations still faithfully describe, in some sense, the representation category when q is root of unity; see Bodish [1] and Elias [7] for instance.

The setting of our work here is of G_2 webs on surfaces, with particular focus on the case when q is a root of unity. Two webs on a surface can be multiplied by stacking web diagrams on the surface, yielding an algebra structure. These skein algebras of surfaces are rich algebras related to both quantum groups and character varieties. Skein algebras are generally noncommutative, and their centers depend on the parameter q . Our main goal is to use the skein relations of Kuperberg to find central elements in the G_2 skein algebras.

1.1 Power sum elements

There is substantial literature on the subject of the SL_2 case of Kauffman bracket skein algebras of surfaces. In the root of unity case, surprising central elements were discovered in the work of Bonahon and Wong [5]. These central elements arise from a topological operation of threading of certain polynomials along knots. The polynomials in question are the Chebyshev polynomials $T_n \in \mathbb{Z}[x]$ which satisfy the following fundamental property regarding the trace of a power of a matrix $A \in SL_2$:

$$T_n(\text{tr}(A)) = \text{tr}(A^n).$$

Since the trace of a power of a matrix is the power sum of the eigenvalues of the matrix, the polynomials T_n can be referred to as power sum polynomials. These central elements obtained by threading power sum polynomials along knots may be referred to as power sum elements in the skein algebra. We are interested in finding analogous elements for the case of G_2 .

The Chebyshev elements at roots of unity play a fundamental role in studying the representation theory of SL_2 skein algebras; see Bonahon and Wong [5], Frohman, Kania-Bartoszyńska, and Lê [10], and Ganev, Jordan, and Safronov [12]. For closed surfaces, the Chebyshev elements generate the whole center of the skein algebra [10]. We expect that the power sum elements found in this paper will play a similar role for the case of G_2 .

Power sum elements have also been studied in the setting of a generic parameter q , and they satisfy product-to-sum formulas which allow for an explicit presentation of the skein algebra of the torus. They have been studied for each of the settings of the SL_2 , HOMFLYPT, and Kauffman skein theories; see Frohman and Gelca [9], Morton and Samuelson [22], and Morton, Pokorny, and Samuelson [21]. For the SL_2 case, Chebyshev elements are also involved in the construction of positive bases of skein algebras of surfaces, a construction which is important in the program of categorification of surface skein algebras and also is related to the theory of cluster algebras; see Mandel and Qin [19], Queffelec [26], and Thurston [30].

1.2 Main results

We study the G_2 skein algebra, $\mathcal{S}_q^{G_2}(\Sigma)$, of a surface Σ using the webs of Kuperberg. For our ring of coefficients, we use a commutative integral domain $\mathcal{R}$ containing a distinguished invertible element $q \in \mathcal{R}$. Due to denominators appearing in the defining skein relations, we further assume that $[12]^{-1} \in \mathcal{R}$, where the quantum integer $[k]$ denotes the Laurent polynomial defined by $[k] = (q^k - q^{-k})/(q - q^{-1})$.

In [Section 4.2](#), we construct 2-variable polynomials $P_n(x, y)$ and $Q_n(x, y)$ in $\mathcal{R}[x, y]$ that are related to taking traces of powers of matrices in the two fundamental matrix representations of the Lie group G_2 . We hence call elements obtained by threading P_n or Q_n along knots *power sum elements* in the skein algebra. These elements satisfy a transparency property described in [Section 3.3](#), and the polynomials P_n and Q_n provide examples of so-called *transparent polynomials*.

Theorem 1.1 ([Theorem 5.7](#)) *If $q^{2n} = 1$ then the element obtained by threading the polynomial P_n or the polynomial Q_n along a knot diagram on a surface Σ is a central element of the G_2 skein algebra of the surface, $\mathcal{S}_q^{G_2}(\Sigma)$.*

We also obtain a result addressing the uniqueness of those polynomials which yield central elements by threadings.

Theorem 1.2 ([Theorem 5.9](#)) *Suppose that q^2 is a primitive root of unity of order n coprime to 3. Then any transparent polynomial $S(x, y)$ is generated by P_n and Q_n . In other words, $S(x, y)$ must be an element of the polynomial ring $\mathcal{R}[P_n, Q_n]$.*

Along the way, we study the confluence theory of the G_2 skein relations ([Theorem 2.4](#)) to describe bases of G_2 skein modules of surfaces. This had already been studied by Kuperberg [\[15\]](#) and Sikora and Westbury [\[29\]](#) using a smaller set of skein relations than the set that we use. We have incorporated the relations from the more recent works of Bodish and Wu [\[3\]](#) and Sakamoto and Yonezawa [\[27\]](#). See [Remark 2.5](#) comparing the two potential sets of skein relations.

The computations involved in establishing our main theorems take place inside of an algebra associated to the annulus $\mathcal{A}$, and an algebra associated to the twice-marked annulus $\mathcal{A}_{1,1}$. These computations are inspired by the earlier work of Lê in the SL_2 setting [\[16\]](#). Using our constructed bases, we show that $\mathcal{A} \cong \mathcal{R}[x, y]$ and that $\mathcal{A}_{1,1}$ is a commutative algebra, obtaining an explicit presentation for it in [Corollary 3.10](#). These are essential ingredients for us in connecting the theory of commutative polynomials to the skein theory of G_2 webs.

A similar approach in constructing transparent polynomials using the explicit skein relations for the setting of SL_d webs was carried out by Bonahon and the third author in [\[4\]](#). Informally, the strategy common between these papers is to attempt to match up certain elements in the commutative ring $\mathcal{R}[\lambda_1^{\pm 1}, \lambda_2^{\pm 1}]$ with certain diagrammatic elements in $\mathcal{A}_{1,1}$. In the SL_d case, this strategy goes exactly as planned. When the analogous approach was tried in this project for G_2 webs, it produced a surprising error term. In [Section 5.2](#) we discuss this error term and explain how to navigate around it.

Our approach here is elementary, and uses the explicit skein relations of Kuperberg. It would be quite interesting to extend this work to find a skein-theoretic proof that the threading operation respects the Kuperberg web relations and describes an algebra homomorphism $\mathcal{S}_1^{G_2}(\Sigma) \rightarrow \mathcal{S}_q^{G_2}(\Sigma)$ when q is an appropriate root of unity. There is an alternative approach to establishing the existence of these central elements, using the theory of quantum groups. It involves the Frobenius map of Lusztig at roots of unity; see Ganey, Jordan, and Safronov [12]. Kuperberg [15] showed that the G_2 webs and skein relations faithfully describe a full subcategory of the category of representations associated to the quantum group G_2 when the parameter q is generic. It is predicted that the same still holds when q is specialized to most roots of unity; see Bodish and Wu [3] and Morrison [20]. Such a result should provide an avenue for the translation of results from the algebraic setting of quantum groups at roots of unity to the skein-theoretic setting of Kuperberg webs. We expect that elements obtained by threading our polynomials P_n and Q_n along knots form a generating set for the image of the Frobenius map described in [12].

1.3 Multivariate Chebyshev polynomials

In the case of SL_2 , there are two families of single-variable polynomials T_n and S_n , which satisfy a linear relationship $T_n = S_n - S_{n-2}$ and play fundamental roles in the theory of SL_2 and its representation ring. The polynomials P_n and Q_n that we study here are 2-variable generalizations of the Chebyshev polynomial of the first kind, T_n , which satisfies the power sum property $T_n(\mathrm{tr}(A)) = \mathrm{tr}(A^n)$ for a matrix A in SL_2 . The closely related Chebyshev polynomial of the second kind, S_n , satisfies the equation $S_n([V_1]) = [V_n]$ in the representation ring of SL_2 , which expresses the representation V_n of highest weight n as a polynomial in the vector representation V_1 . The polynomial S_n can be obtained diagrammatically by taking the Jones–Wenzl idempotent in the Temperley–Lieb algebra, $\mathrm{JW}_n \in \mathrm{TL}_n$, and taking its annular closure, considered as an element of the SL_2 skein algebra of the annulus $\mathcal{S}_q^{\mathrm{SL}_2}(\mathbb{A}) \cong \mathcal{R}[x]$; see Lickorish [18, Chapter 13].

Nesterenko, Patera, and Tereszkievich [24] give recursive definitions for two families of polynomials, which they call C -polynomials and S -polynomials, generalizing the Chebyshev polynomials of the first and second kinds for compact Lie groups. In the case of G_2 , their polynomials $C_{(a,b)}$ are parametrized by dominant integral weights. We believe our polynomials P_n and Q_n are the same as their polynomials $C_{(n,0)}$ and $C_{(0,n)}$, although a change of variables is required. The basic variables of [24] correspond to the so-called C -functions of the two fundamental representations, while our basic variables are the characters of the two fundamental representations.

In [3], for each dominant integral weight (a, b) , Bodish and Wu construct the G_2 analogue of a Jones–Wenzl idempotent as a linear combination of G_2 webs. If we take the annular closure of an idempotent, we obtain an element of $\mathcal{S}_q^{G_2}(\mathbb{A})$, which is identified with a polynomial in $\mathcal{R}[x, y]$ by our Proposition 3.2. We expect that under this identification, the idempotent associated to the weight (a, b) yields the polynomial $S(a, b)$ described in [24]. Results analogous to this have been obtained in the case of SL_3 by Ohtsuki and Yamada [25] en route to a skein-theoretic construction of the Witten–Reshetikhin–Turaev 3-manifold

invariant. Those results were further studied by Yuasa [31] for the purpose of categorifying the SL_3 analogue of Chebyshev polynomials of the second kind.

Acknowledgments

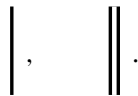
This work took place during the SURIEM REU program at Michigan State University, hosted by the MSU Lyman Briggs College during Summer 2023. We are grateful for funding for this project provided by NSA grant H98230-23-1-0005. The members of our project benefited from a synergistic environment resulting from a concurrently running undergraduate Summer Topology Program at MSU. We thank the organizers and participants of that program for stimulating conversations and for inviting us to daily tea times, funded by NSF RTG grant DMS-2135960. We would also like to thank F Bonahon for helpful conversations, as the idea for this project came out of a joint project between him and the third author. We also thank the referee for helpful comments and suggestions.

2 G_2 skein relations and reduction rules

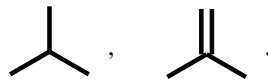
In this section, we give the defining skein relations for G_2 skein algebras of surfaces. Let Σ be an oriented surface of finite type. Let $\mathcal{R}$ denote a commutative integral domain containing an invertible element $q \in \mathcal{R}$ such that $[12]^{-1} \in \mathcal{R}$, where the quantum integer $[k]$ denotes the Laurent polynomial $[k] = (q^k - q^{-k})/(q - q^{-1})$.

2.1 G_2 skein algebras of surfaces

An abstract G_2 web is a trivalent graph where each edge is either a single-strand or double-strand,



and each vertex is one of the following two types:



We also allow vertexless loops of either strand type. The graph need not be connected. We will study G_2 webs on surfaces subject to certain skein relations of Kuperberg. Linear combinations of webs provide the structure of an $\mathcal{R}$ -module called a skein module, and it is equipped with a natural algebra structure. We now provide the necessary definitions.

Definition 2.1 A web W on an oriented surface Σ is a diagram on Σ which is an embedding of a G_2 web except for transverse double-points, which are labeled as overcrossings or undercrossings. Two webs W_1 and W_2 are isotopic on Σ if there is an isotopy of Σ taking W_1 to W_2 .

Definition 2.2 Let Σ be a smooth oriented surface of finite type. We define the G_2 skein module of Σ , denoted by $\mathcal{S}_q^{G_2}(\Sigma)$, to be the $\mathcal{R}$ -module spanned by isotopy classes of webs on Σ subject to the skein relations (1)–(12) listed later in this section.

We now describe the algebra structure for the G_2 skein algebra of a surface Σ , also denoted by $\mathcal{S}_q^{G_2}(\Sigma)$.

Definition 2.3 The module $\mathcal{S}_q^{G_2}(\Sigma)$ has an algebra structure arising from the bilinear extension of the following product of webs. If W_1 and W_2 are webs, then the product $W_1 \cdot W_2$ is obtained by isotoping the diagrams W_1 and W_2 on Σ so that they only intersect at transverse double points and then taking the union $W_1 \cup W_2$. At each intersection of W_1 with W_2 , the crossing data is assigned so that W_1 is drawn above W_2 using overcrossings.

The product operation on webs is well defined with respect to isotopy because the defining skein relations imply the following types of Reidemeister moves, which hold for all strand types and vertex types:

$$\alpha = \beta, \quad \gamma = \delta, \quad \epsilon = \zeta, \quad \eta = \theta.$$

In fact, there is an equivalent definition of the skein algebra of a surface, using the language of thickened surfaces and ribbon graphs, which satisfies the same Reidemeister moves as above. More generally, one may define skein modules spanned by ribbon graphs in oriented 3-manifolds. We focus on the more 2-dimensional perspective of webs on surfaces since that is the setting in which our computations take place.

The following skein relations serve two purposes. The first is that they are the defining relations for the skein modules. The second is that each relation can be interpreted as a reduction rule which replaces a web by a linear combination of simpler webs. We will soon explain how this leads to a basis for skein modules of surfaces, by using the method of confluence theory described in [29].

We use the following rules to remove crossings and internal double-edges:

$$\begin{aligned} (1) \quad & \diagup \diagdown = \frac{1}{[2]}(q^3) \left(+q^{-3} \frown + q \smile + q^{-1} \smile \right), \\ (2) \quad & \diagup \diagup = \frac{1}{[2]}(q^3 \smile + q^{-3} \smile + \frac{1}{[2]} \boxtimes), \\ (3) \quad & \diagdown \diagdown = \frac{1}{[2]}(q^{-3} \smile + q^3 \smile + \frac{1}{[2]} \boxtimes), \\ (4) \quad & \boxtimes = q^6 \left(\frown + q^{-6} \smile + \frac{2}{[2]^2} \boxtimes + \frac{q^3}{[2]^3} \boxtimes + \frac{q^{-3}}{[2]^3} \boxtimes \right), \\ (5) \quad & \smile = - \left(+ \frac{[4][6]}{[2][12]} \frown + \smile + \frac{1}{[3]} \smile \right). \end{aligned}$$

We call (5) the “double-edge removal” relation. Once we have applied the crossing removals and double-edge removals to a web diagram on a surface, we are left with webs which might have some polygonal faces bounded by single edges. The following defining reduction rules are used to remove small faces:

$$(6) \quad \bigcirc = \frac{[2][7][12]}{[4][6]}, \quad \bigodot = \frac{[7][8][15]}{[3][4][5]},$$

$$(7) \quad \bigcirc \downarrow = 0, \quad \bigodot \downarrow = 0,$$

$$(8) \quad \bigcirc \downarrow = -\frac{[3][8]}{[4]} \Big|, \quad \bigodot \downarrow = -[2]^2 \Big|, \quad \bigcirc \downarrow \downarrow = 0,$$

$$(9) \quad \triangle = \frac{[6]}{[2]} \downarrow, \quad \bigtriangleup = -[3] \downarrow, \quad \triangle \downarrow = 0,$$

$$(10) \quad \square = [3] \left(\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \right) + \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} - \frac{[4]}{[2]} \left(\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} \right),$$

$$(11) \quad \square \downarrow = \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array},$$

$$(12) \quad \star = \left(\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \right) - \left(\begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} + \begin{array}{c} \diagdown \diagup \\ \diagup \diagdown \end{array} + \begin{array}{c} \diagup \diagdown \\ \diagdown \diagup \end{array} \right).$$

All of these defining relations are local, in the sense that they hold in any disk on the surface. Some of the defining relations follow from the other defining relations. See [3; 27] for a smaller set of defining relations. We have listed the larger set because these are needed to form a set of relations which are confluent, in the following sense.

We can view each relation (1)–(12) as a reduction rule for rewriting a web as a linear combination of webs which are simpler when considering the triple of numbers (crossings, vertices, connected components), considered lexicographically. This allows us to find a spanning set of webs for $\mathcal{S}_q^{G_2}(\Sigma)$ consisting of the set of webs which cannot be reduced further. In order for this spanning set to be a basis, [29] explains that we must check that every overlap of the reduction rules can be resolved consistently. To this end, in addition to the defining relations, we need the following reduction rules. Even though these apply in annular regions instead of disk-shaped regions, they follow from the previous local rules. In the following pictures, the dot represents a region on the surface which may be noncontractible.

$$(13) \quad \bigodot \downarrow = \bigodot \downarrow - [2]^2 \bigodot \downarrow + \frac{[8]}{[4]} \bigcirc \downarrow - [7] \cdot,$$

$$(14) \quad \bigodot \downarrow = \bigodot \downarrow - \frac{[6]}{[2][3]} \bigcirc \downarrow.$$

The relation (13) is derived by considering an overlap of the middle relation of (8) and (5). The relation (14) is derived by considering an overlap of the last relation of (9) and of (5). We refer to the region bounded by the web in the left side of (13) as a *self-folded square*, and the region bounded by the web in the left side of (14) a *self-folded pentagon*.

By considering an overlap of the relation (14) with itself, we find we need the equivalence rule (15). This allows us to permute any two simple loops on the surface which bound an embedded annulus on the surface:

$$(15) \quad \text{Diagram 1} = \text{Diagram 2}$$


$$=$$


We do not consider it as a reduction rule, but we use it as an equivalence relation placed on the set of *irreducible diagrams*, meaning those diagrams which do not contain any of the subdiagrams appearing in the left sides of relations (1)–(14). We identify any two diagrams which are related by a sequence of moves of the form (15).

2.2 Bases of skein modules

When we apply a reduction rule, we replace a web with a linear combination of simpler descendant diagrams. Our relations are confluent if, whenever two reduction rules are applicable to a web, the element obtained by applying the first reduction rule and the element obtained by applying the second can both be reduced to the same element, up to surface isotopy and moves of the form (15). A set of reduction rules is terminal if no diagram has an infinite chain of descendants. If the defining set of relations for a module consists of a set of reduction rules which is terminal and confluent, then the irreducible diagrams form a linearly independent spanning set for the module. We encourage the reader unfamiliar with confluence theory (also called rewriting theory or the diamond lemma) to consult [29] for the description of this procedure in the context of skein theory. Other similar examples of the process of checking the confluence of skein relations can be found in [11; 13; 17].

Theorem 2.4 *The above defining skein relations (1)–(14) are terminal and confluent. Thus a basis for $S_q^{G_2}(\Sigma)$ consists of irreducible webs to which none of relations (1)–(14) apply. Any two irreducible webs representing the same basis element are related by an isotopy of the surface along with a sequence of loop switch moves (15).*

Proof Each reduction rule (1)–(14) strictly decreases the numbers of the following, in lexicographic order: (crossings, vertices, connected components). Thus the set of reduction rules is terminal. Then the proof follows from applying [29, Theorem 2.3], by checking that the relations are locally confluent for all nontrivial overlaps of the reduction rules. $\square$

In Proposition 2.8, we give a characterization of irreducible webs by analyzing their faces.

Remark 2.5 Our Theorem 2.4 fills in a small gap in the literature by listing the full set of reduction rules required to describe a basis of the G_2 skein module when both single-strands and double-strands are used. Although relations (5) and (6) appear in [15; 29], the other reduction rules involving double-strands were left out. They are required to complete the relations. For example, without the relation

$$\text{Diagram 1} = -[2]^2 \text{Diagram 2}$$


$$= -[2]^2$$


both Schur's lemma and the Reidemeister 2 move fail to hold. Hence the basis claimed in [29, Theorem 7.2] regarding G_2 is not a basis for the version of the skein module we study here. The more recent works [3; 27] do include all of the necessary defining relations. Our analysis of the overlaps of the relations led to the reduction rule (13), which we have not previously found in the literature.

On the other hand, if one considers only webs with single-strands and uses relations only concerning single-strands, then the relations appearing in [15] are complete, and from [29, Theorem 7.2] a basis of that version of the skein module can be deduced by leaving out the double-edge removal relation.

We will also consider skein modules of surfaces with marked points on the boundary. Here, however, the only marked surface we explicitly deal with is the twice-marked annulus $\mathcal{A}_{1,1}$ as described in Section 3.2.

Definition 2.6 Let $M = p_1, \dots, p_k \in \partial\Sigma$ be a finite set of marked points on the boundary of Σ . The skein module of the marked surface $S_q^{G_2}(\Sigma, M)$ is the $\mathcal{R}$ -module spanned by webs W having single-strand endpoints such that $\partial W = M$. Webs are considered up to isotopies of Σ which fix the boundary components.

Our Theorem 2.4 also applies when (Σ, M) is a surface with marked boundary. Following Kuperberg [14], we can identify irreducible webs by analyzing contributions of web faces to the Euler characteristic of the surface.

Definition 2.7 (Euler measure) Let W be a crossingless web on a surface Σ of finite type. We call a connected component of $\Sigma \setminus W$ a *face* of W . Each face U of W is bounded by some number of edges and vertices. The edges could be edges (or loops) of W or could be boundary edges (or boundary circles) of Σ , if W has endpoints on the boundary. If every edge bounding U is an edge of W , then U is an *interior face* of W . If at least one edge bounding U is a boundary edge of Σ , then we call U a *boundary face* of W . If the interior of the face U is homeomorphic to an open disk, then we call the face an *n -gon*, where n is the number of edges bounding the face, counted with multiplicity, so that an edge is counted twice if the face U appears on both sides of the edge. For example, the left sides of relations (10) and (13) depict faces which are both 4-gons.

For each face U of W , we define its *Euler measure* $\mu(U)$ to be

$$\mu(U) = \frac{1}{3}V + \frac{1}{2}V_{\partial} - \frac{1}{2}E - E_{\partial} + \chi(U),$$

where V and V_{∂} are the number of interior vertices and boundary vertices of W , E is the number of edges of U which belong to W , and E_{∂} is the number of edges of U which belong to $\partial\Sigma$. Vertices and edges are both counted with multiplicity. The quantity $\chi(U)$ is the ordinary Euler characteristic of the face U considered as a submanifold of the surface Σ . Loops of W and boundary circles of $\partial\Sigma$ do not contribute to the Euler measure.

A classical fact is that the Euler measure satisfies the following fundamental property. If the faces of W are labeled by $U_1, \dots, U_k$, then we have

$$\sum_{i=1}^k \mu(U_i) = \chi(\Sigma).$$

If U_n is an interior n -gon, then $\mu(U_n) = \frac{1}{3}n - \frac{1}{2}n + 1 = 1 - \frac{1}{6}n$. If U is a face that is not simply connected, then $\mu(U) \leq 0$. From this discussion we obtain the following characterization of terminal elements:

Proposition 2.8 *A web W on Σ is irreducible if and only if it contains no crossings, no interior double-edges, and if every interior face U of W satisfies $\mu(U) \leq 0$.*

Proof It follows from the discussion preceding Proposition 2.8 that the property of a web W having no interior face U of positive Euler measure is equivalent to W having no interior face an n -gon for $n \leq 5$. By analyzing our reduction rules (1)–(14), we will show that this property holds. Let W be an irreducible web diagram. By rules (1)–(5), W has no crossings and no interior double-edges. Furthermore, by Definition 2.6, W has no double-edge ending on the boundary of Σ , and so every interior vertex of W is of the form $\bigwedge$. The remaining reduction rules eliminate all interior n -gon faces for $n \leq 5$. These eliminated n -gons include self-folded ones by observing the following. Consider a sequence S of adjacent edges (considered with multiplicity) which bound an n -gon face U of W . Since each vertex of W is trivalent, any two instances of the same edge of W cannot be adjacent to each other in the sequence S . This implies that self-folded n -gon faces do not exist for $n \leq 3$ and that any self-folded interior 4-gon and interior 5-gon are eliminated by rules (13) and (14), respectively. $\square$

Kuperberg used the notion of Euler measure to prove the following:

Proposition 2.9 [14] *If Σ is a disk (with no marked points) or is a sphere, then $\mathcal{S}_q^{G^2}(\Sigma) \cong \mathcal{R}$.*

Proof It suffices to show that our basis of $\mathcal{S}_q^{G^2}(\Sigma)$ consists of only the empty diagram. Suppose that W is a basis web for $\mathcal{S}_q^{G^2}(\Sigma)$. Then W contains no internal double-edges, no crossings, and no n -gons for $n \leq 5$. Since Σ is a sphere or is a disk, the Euler characteristic of Σ is positive. If W is nonempty, then it must have a face of positive Euler measure. However, a face of positive Euler measure must be an n -gon with $n \leq 5$. Since W has no such face, W must be empty. $\square$

3 Annuli

We now study some algebras associated to skein modules of the annulus. In this section, we study these algebras from a purely skein-theoretic perspective. Later on, we will return to these algebras to perform computations with power sum polynomials.

3.1 The annulus and the algebra $\mathcal{A}$

Let $\mathbb{A} = S^1 \times [0, 1]$ denote the annulus.

Definition 3.1 Define $\mathcal{A}$ to be the $\mathcal{R}$ -algebra spanned by closed webs in $\mathbb{A}$. If $W_1, W_2 \in \mathcal{A}$ are two webs, we define their product, $W_1 W_2$ to be the web obtained by isotoping W_1 so it is contained in $S^1 \times (\frac{1}{2}, 1]$, isotoping W_2 so it is contained in $S^1 \times [0, \frac{1}{2})$, and taking the union of the webs. This product agrees with that of $\mathcal{S}_q^{G_2}(\mathbb{A})$.

The multiplicative unit $1 \in \mathcal{A}$ is the empty diagram. We define $x, y \in \mathcal{A}$ to be the following distinguished elements:

$$x = \text{diagram of a single loop with a central dot}, \quad y = \text{diagram of two concentric loops with a central dot}.$$

Proposition 3.2 The algebra $\mathcal{A}$ associated to the annulus is isomorphic to $\mathcal{R}[x, y]$ and has a basis of diagrams consisting of concentric loops around the core of the annulus.

Proof We first describe the basis of $\mathcal{A}$. Let W be a basis web. Thus W is a crossingless web which contains no internal double-edge, and has no interior n -gon face for $n \leq 5$. Any boundary face of W is not simply connected and must have nonpositive Euler measure. Consequently, W has no face of positive Euler measure. If W has a trivalent vertex, then as W must have a nonsimply connected face, W would have a face of negative Euler measure. Therefore, as the Euler characteristic of $\mathbb{A}$ is zero, W cannot contain a trivalent vertex. Thus W consists of concentric loops around the core of the annulus, and the description of the basis is as claimed.

Now we describe the algebra structure of $\mathcal{A}$. Due to the equivalence rule (15), the two kinds of loops in $\mathcal{A}$ commute. Consequently, there is a well-defined algebra homomorphism $\mathcal{R}[x, y] \rightarrow \mathcal{A}$ sending x to an essential single-strand loop and sending y to an essential double-strand loop. This map sends the standard basis of monomials of $\mathcal{R}[x, y]$ injectively onto the basis we have just described for $\mathcal{A}$. Thus this algebra map is an isomorphism $\mathcal{R}[x, y] \cong \mathcal{A}$. $\square$

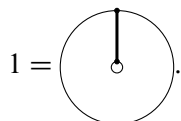
3.2 The twice-marked annulus and the algebra $\mathcal{A}_{1,1}$

Consider the annulus $\mathbb{A} = S^1 \times [0, 1]$ and a fixed point $p \in S^1$. Let $p_0 = (p, 0)$ and $p_1 = (p, 1)$ be two marked points on $\partial\mathbb{A}$. We consider webs with boundary equal to $M = \{p_0, p_1\}$.

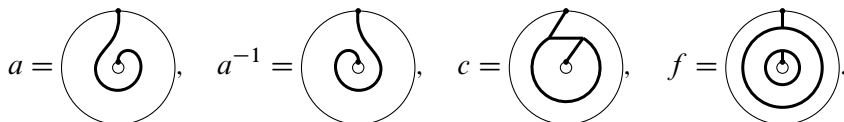
Definition 3.3 Define $\mathcal{A}_{1,1}$ to be the skein module $\mathcal{S}_q^{G_2}(\mathbb{A}, \{p_0, p_1\})$ of the marked annulus spanned by webs with one single-strand endpoint at p_0 on the interior boundary of the annulus and one single-strand endpoint at p_1 on the exterior boundary of the annulus. The algebra structure for $\mathcal{A}_{1,1}$ is given by concatenating annuli outward, as described for $\mathcal{A}$. Note that this product is not the same as a skein algebra product.

Remark 3.4 In the literature, $\mathcal{A}_{1,1}$ is also known as an endomorphism space in the *annular representation category*, *annular spider*, *skein category of the annulus*, or as a corner of the *tube algebra*. The product structure is to be thought of as the composition of morphisms. We will later show that $\mathcal{A}_{1,1}$ is commutative (Corollary 3.9).

The multiplicative unit of $\mathcal{A}_{1,1}$ is the following diagram:

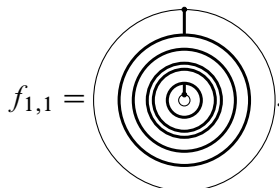


We identify the following distinguished elements $a, a^{-1}, c, f \in \mathcal{A}_{1,1}$:



In addition to the fundamental elements $a, a^{-1}, 1, c, f \in \mathcal{A}_{1,1}$ listed, we also need to provide notation for the elements $f_{i,j} \in \mathcal{A}_{1,1}$ which consist of the diagram f along with i concentric single-strand loops and j concentric double-strand loops inside of the open annular region bounded by the diagram of f .

For example, the following depicts the element $f_{1,1}$:



Using these elements, we can describe a basis for $\mathcal{A}_{1,1}$ over $\mathcal{R}$ (Proposition 3.6) and later a presentation for $\mathcal{A}_{1,1}$ as an $\mathcal{R}$ -algebra (Corollary 3.10).

We first record some topological facts about diagrams in $\mathcal{A}_{1,1}$. Recall that the mapping class group of the annulus, $\text{MCG}(\mathbb{A})$, is the group of isotopy classes of orientation-preserving diffeomorphisms of $\mathbb{A}$ which fix the boundary. It is a classical fact (see [8]) that $\text{MCG}(\mathbb{A}) \cong \mathbb{Z}$ and is generated by a Dehn twist of the annulus. The annulus is equipped with a natural orientation-reversing diffeomorphism $\mathbb{A} \rightarrow \mathbb{A}$, called the mirror image, which reverses the orientation of the S^1 factor of $S^1 \times [0, 1]$. From these facts we observe that if a web $W_1 \in \mathcal{A}_{1,1}$ is taken to a web $W_2 \in \mathcal{A}_{1,1}$ by an orientation-preserving diffeomorphism, then $W_1 = a^k W_2$ for some integer power k of the twist diagram a . Similarly, if W_1 is taken to W_2 by an orientation-reversing diffeomorphism, then W_1 and the mirror image of W_2 are related by a power of a . We use these observations to record the following:

Lemma 3.5 Suppose $W_1 \in \mathcal{A}_{1,1}$ is a web and $W_2 \in \mathcal{A}_{1,1}$ is one of $\{1, c^j, f_{ij}\}$. If W_1 and W_2 are diffeomorphic then there exists some $k \in \mathbb{Z}$ such that

$$W_1 = a^k W_2.$$

Moreover, if $W_2 = f_{ij}$, then we may use $k = 0$.

Proof The webs 1 and f_{ij} are invariant under the mirror-image operation, so the statement follows in these cases from the previous discussion. Moreover, it can be seen diagrammatically that $a^k f_{ij} = f_{ij}$ for all k . Now consider the case $W_2 = c^j$. It is an easy observation that the mirror image of the diagram c is

equal to $a^{-1}c$. From this, and the fact that the mirror image operation induces an algebra homomorphism $\mathcal{A}_{1,1} \rightarrow \mathcal{A}_{1,1}$, it follows that the mirror image of c^j is equal to $a^{-j}c^j$. Thus W_1 is related by a power of a to the diagram $a^{-j}c^j$, as required. $\square$

Proposition 3.6 *The algebra $\mathcal{A}_{1,1}$ associated to the twice-marked annulus has a basis of irreducible webs constructed in Theorem 2.4, and consists of the set of webs $\{a^i c^j \mid i \in \mathbb{Z}, j \in \mathbb{Z}_{\geq 0}\} \cup \{f_{i,j} \mid i, j \in \mathbb{Z}_{\geq 0}\}$.*

The proof of Proposition 3.6 follows from Lemmas 3.7 and 3.8.

Lemma 3.7 *The elements of $\{a^i c^j \mid i \in \mathbb{Z}, j \in \mathbb{Z}_{\geq 0}\} \cup \{f_{i,j} \mid i, j \in \mathbb{Z}_{\geq 0}\}$ are distinct and each element of the set is some basis element of $\mathcal{A}_{1,1}$.*

Proof Since the elements of $\{a^i c^j\} \cup \{f_{i,j}\}$ are all irreducible, they are basis elements. To show they are distinct, we must show that they are not related to each other by sequences of isotopies and moves of the form (15). Since isotopies and loop switch moves (15) preserve the number of single-strand and double-strand loops, the set $\{f_{i,j}\}$ consists of distinct elements. Since every element of $\{f_{i,j}\}$ is disconnected while every element of $\{a^i c^j\}$ is connected, $\{f_{i,j}\} \cap \{a^i c^j\} = \emptyset$.

We now show the elements of $\{a^i c^j\}$ are distinct. If $j \neq j'$, then $a^i c^j \neq a^{i'} c^{j'}$ for any i and i' , since the two diagrams have different numbers of vertices. If $i \neq i'$, then a^i is not isotopic to $a^{i'}$ since the diagrams have different winding numbers. Now suppose that $a^i c^j$ is isotopic to $a^{i'} c^{j'}$. Within the underlying web of c^j , there is a path from p_0 to p_1 which is isotopic rel endpoints to the identity element $1 \in \mathcal{A}_{1,1}$. Concatenating this path with a^i yields a path γ in $a^i c^j$ isotopic rel endpoints to a^i . On the other hand, if we concatenate the same path in c^j with $a^{i'}$, we find a path γ' in $a^{i'} c^{j'}$ isotopic rel endpoints to $a^{i'}$. An isotopy of $\mathbb{A}_{1,1}$ taking $a^{i'} c^{j'}$ to $a^i c^j$ must take γ to γ' , so we must have $i = i'$. Therefore elements of the set $\{a^i c^j\}$ are distinct. $\square$

Lemma 3.8 *If W is a basis web of $\mathcal{A}_{1,1}$, then W must be an element of*

$$\{a^i c^j \mid i \in \mathbb{Z}, j \in \mathbb{Z}_{\geq 0}\} \cup \{f_{i,j} \mid i, j \in \mathbb{Z}_{\geq 0}\}.$$

Proof Let W be a basis web. We consider two cases. If p_0 and p_1 belong to different connected components of W , we claim that $W = f_{i,j}$ for some i and j . Let W_0 be the connected component of W containing the endpoint p_0 and let W_1 be the connected component of W containing the endpoint p_1 . It will suffice to show the equality of webs $W_0 \cup W_1 = f$. Since W_0 has an endpoint at p_0 but not at p_1 , the web W_0 must contain a trivalent vertex. Consider the face of $\mathbb{A} \setminus W_0$ containing p_1 . It is not simply connected, and so must have negative Euler measure. Thus some other face U_0 of W_0 must have positive Euler measure. Since W_0 is irreducible, this face U_0 must be adjacent to the boundary edge of $\mathbb{A}$ containing p_0 . We now analyze U_0 . It is connected and simply connected, but not a sphere. Thus U_0 is topologically a disk and $\chi(U_0) = 1$. Furthermore, exactly one edge of U_0 is a boundary edge of $\mathbb{A}$, and exactly two vertices (both being p_0 , counted with multiplicity) of U_0 are on the boundary of $\mathbb{A}$. Since the edge of W_0 emerging from p_0 is seen to have U_0 on both of its sides, U_0 is a self-folded n -gon. So

far, we have deduced that $n - 1$ of the edges of U_0 are interior edges and $n - 2$ of the vertices of U_0 are interior vertices. We then compute its Euler measure via [Definition 2.7](#):

$$\mu(U_0) = \frac{1}{3}(n - 1) - \frac{1}{2}(n - 2) + \frac{1}{2}(2) - 1 + 1 = \frac{1}{6}(-n + 5).$$

Since $\mu(U_0) > 0$, we have that $n < 5$. Since self-folded n -gons for $n < 4$ do not exist for G_2 webs, we conclude that $n = 4$ and U_0 is a self-folded square. Hence W_0 is of the form claimed. The same argument shows that W_1 is of the form claimed, and we must have the equality of diagrams $W_0 \cup W_1 = f$. Since the region between W_0 and W_1 is an annulus, any irreducible diagram in that region must be a collection of concentric loops by [Proposition 3.2](#), so W must be equal to $f_{i,j}$ for some i and j .

Next, if W contains a path from p_0 to p_1 , we claim $W = a^i c^j$ for some $i \in \mathbb{Z}$ and $j \in \mathbb{Z}_{\geq 0}$. By [Lemma 3.5](#) it suffices to show that W is diffeomorphic to c^j for some $j \geq 0$. Let W' be the connected component of W containing p_0 and p_1 . If W' contains no *interior* trivalent vertex, then W' is diffeomorphic to $1 = c^0$. Otherwise, W' contains a trivalent vertex and, like above, we consider the boundary face U' of W' adjacent to p_0 . If the Euler measure of U' is negative, then the same analysis as in the first case shows that the boundary face of W adjacent to p_1 would be a self-folded square. However, this case is excluded since we are assuming p_0 and p_1 are connected by a path in W . Therefore $\mu(U') \geq 0$ and U' is a self-folded boundary n -gon. By the above computation of Euler measure, it must be either a self-folded square or a self-folded pentagon. The case of U' being a self-folded square is again excluded since p_0 is connected to p_1 . Therefore the boundary face U' must be a self-folded pentagon. Similarly, the boundary face adjacent to p_1 must be a self-folded pentagon as well. Since these boundary faces account for an Euler measure of zero, any other face of W' must also have Euler measure zero. Therefore all other faces of W' must be hexagons. By inspection, they must be self-folded hexagons and so the diagram W' must be diffeomorphic to some power of c . Since all faces of W' are topologically disks, any other component of W would be contained in the complement of W' and would thus be reducible by [Proposition 2.9](#). So $W = W' = a^i c^j$, as claimed. $\square$

Corollary 3.9 *The algebra $\mathcal{A}_{1,1}$ is commutative.*

Proof By the characterization of basis elements of $\mathcal{A}_{1,1}$ in [Proposition 3.6](#), one only needs to check that the elements a , c , and $f_{i,j}$ commute with each other. Some of these are checked by simple isotopies while others are checked using short computations similar to the ones in the proof of the next corollary. $\square$

Corollary 3.10 *The algebra $\mathcal{A}_{1,1}$ has the following presentation as an $\mathcal{R}$ -algebra. It is the quotient of the algebra generated by the commuting variables a , a^{-1} , c , and f_{ij} for $i, j \geq 0$, subject to the relations*

$$af_{i,j} = f_{i,j}, \quad a^{-1}f_{i,j} = f_{i,j}, \quad cf_{i,j} = f_{i+1,j} - \frac{[6]}{[2][3]}f_{i,j},$$

$$f_{i,j}f_{k,l} = f_{i+k+2,j+l} - [2]^2f_{i+k,j+l+1} + \frac{[8]}{[4]}f_{i+k+1,j+l} - [7]f_{i+k,j+l}.$$

Proof We first suppose the relations above hold. Since each element of our basis for $\mathcal{A}_{1,1}$ can be written as a product of these proposed generators, our list is a generating set for $\mathcal{A}_{1,1}$. On the other hand,

any element written as a linear combination of products of these generators can be reduced to a linear combination of basis elements using only the relations listed in the proposition. Therefore, if the relations hold, then they are a full set of relations for $\mathcal{A}_{1,1}$.

We now check this. The relation $af_{i,j} = f_{i,j} = a^{-1}f_{i,j}$ follows from observing a simple isotopy. The relation which rewrites $f_{i,j}f_{k,l}$ follows from the computation $f^2 = f_{2,0} - [2]^2 f_{0,1} + ([8]/[4])f_{1,0} - [7]f$, using the self-folded square relation (13). The relation which rewrites $cf_{i,j}$ follows from the computation $cf = f_{1,0} - ([6]/([2][3]))f$, using the self-folded pentagon relation (14). $\square$

The following will be used later on to identify a nicely behaved subalgebra of $\mathcal{A}_{1,1}$ with a subalgebra of the ring $\mathcal{R}[\lambda_1^{\pm 1}, \lambda_2^{\pm 1}]$.

Corollary 3.11 *The Laurent polynomial ring $\mathcal{R}[a^{\pm 1}, c]$ is a subalgebra of $\mathcal{A}_{1,1}$.*

Remark 3.12 The algebra $\mathcal{A}_{1,1}$ actually admits a finite set of generators $\{a, a^{-1}, c, f\}$. Indeed, in the next section we introduce elements x^* and y^* , and it can be seen that $f_{i,j} = (x^*)^i (y^*)^j f$. Then in the proof of Proposition 5.2, it is shown that x^* and y^* are generated by a, a^{-1}, c and f . However, it is more complicated to describe the algebra using the finite list of generators.

3.3 Transparency and actions of $\mathcal{A}$ on $\mathcal{A}_{1,1}$

We now describe two actions of $\mathcal{A}$ on $\mathcal{A}_{1,1}$ and use them to define transparent polynomials, analogous to the notion defined in [16].

Definition 3.13 We define an algebra homomorphism denoted by $(-)^* : \mathcal{A} \rightarrow \mathcal{A}_{1,1}$ such that if $D \in \mathcal{A}$ is a diagram, the element $D^* \in \mathcal{A}_{1,1}$ is the diagram of D placed over the identity strand $1 \in \mathcal{A}_{1,1}$. It can be checked diagrammatically that this is a well-defined algebra homomorphism.

Similarly, we define $(-)_* : \mathcal{A} \rightarrow \mathcal{A}_{1,1}$ to be the algebra homomorphism such that if $D \in \mathcal{A}$ is a diagram, the element $D_* \in \mathcal{A}_{1,1}$ is the diagram of D placed below the identity strand $1 \in \mathcal{A}_{1,1}$.

Recall that the algebra $\mathcal{A}$ is isomorphic to $\mathcal{R}[x, y]$. Below we show the images of the homomorphisms $(-)^*$ and $(-)_*$ on the generators x, y of $\mathcal{A}$:

$$x^* = \text{diagram of } x \text{ over identity strand}, \quad y^* = \text{diagram of } y \text{ over identity strand}, \quad x_* = \text{diagram of } x \text{ below identity strand}, \quad y_* = \text{diagram of } y \text{ below identity strand}.$$

The following definition uses the fact that $(-)^*$ and $(-)_*$ are algebra homomorphisms:

Definition 3.14 A polynomial $S(x, y) \in \mathcal{A}$ is said to be *transparent* if either of the following equivalent conditions hold:

- (i) $S(x, y)^* = S(x, y)_* \in \mathcal{A}_{1,1}$.
- (ii) $S(x^*, y^*) = S(x_*, y_*) \in \mathcal{A}_{1,1}$.

For example, the empty diagram in $\mathcal{A}$ is a transparent element regardless of the choice of $q \in \mathcal{R}$. If $q = 1$ then every element of $\mathcal{A}$ is transparent. When q is chosen to be a root of unity, we will find more interesting elements which are transparent.

We now briefly explain the process which takes a transparent polynomial and constructs central elements in the skein algebra $\mathcal{S}_q^{G_2}(\Sigma)$ of a surface Σ .

Definition 3.15 Let a knot diagram K on Σ represent a framed knot in the thickened surface $\Sigma \times [0, 1]$. Given a monomial $x^i y^j \in \mathcal{R}[x, y]$, the *threading* of $x^i y^j$ along K is denoted by $K^{[x^i y^j]}$. It is obtained by taking $i + j$ parallel copies of K , in the direction of the framing, such that i of these copies are colored by a single-strand and j of these copies are colored by a double-strand. We can extend this definition linearly to obtain the threading $K^{[S(x, y)]}$ of a polynomial $S(x, y) \in \mathcal{R}[x, y]$ along K .

Lemma 3.16 If an element $S(x, y) \in \mathcal{A}$ is transparent, then the threading $K^{[S]}$ of S along any knot K is a central element of $\mathcal{S}_q^{G_2}(\Sigma)$.

Proof Let K be a knot diagram on Σ , representing a framed knot which is an embedded annulus in the thickened surface given by an orientation-preserving embedding $k: \mathbb{A} \hookrightarrow \Sigma \times [0, 1]$. Let W be an arbitrary web on Σ . Consider the element $KW \in \mathcal{S}_q^{G_2}(\Sigma)$ which is represented by a diagram with finitely many transverse crossings of K and W , all of which involve K crossing over an edge of W . Consider $K^{[S]}$, which is the image of the embedding k applied to the element $S(x, y)$ on $\mathbb{A}$. To show that $K^{[S]}$ is central, we argue that in each spot that $K^{[S]}$ crosses over a strand of W , the overcrossings can be changed to undercrossings, which implies that $K^{[S]}W = WK^{[S]}$. Each crossing change involving a single-strand of W can be viewed as the image of k applied to the transparency property for $S(x, y)$. Since the skein relations are preserved under the embedding k , these crossing changes are valid in $\mathcal{S}_q^{G_2}(\Sigma)$.

To finish the proof of the lemma, it suffices that W can be replaced by an equivalent web so that all of the crossings of K and W involve single-strands of W . This follows by using the second relation of (8),

$$\parallel = -\frac{1}{[2]^2} \text{⦿},$$

which shows that a small interval of a double-strand can be replaced by two single-strands. $\square$

Threadings of polynomials can be defined similarly for links, and threadings of transparent polynomials along links yield central elements in the skein algebra as well.

4 Power sum polynomials for G_2

We now develop the ingredients to construct transparent polynomials.

4.1 Polynomials in $\mathcal{R}[\lambda_1^{\pm 1}, \lambda_2^{\pm 1}]$

Consider the 2-variable Laurent polynomial algebra $E = \mathcal{R}[\lambda_1^{\pm 1}, \lambda_2^{\pm 1}]$. Later, we will relate certain elements of this ring with elements in the algebra $\mathcal{A}_{1,1}$.

Definition 4.1 To a monomial $\lambda_1^i \lambda_2^j \in E$ we associate a degree d_1 and a bidegree d_2 given by

$$d_1(\lambda_1^i \lambda_2^j) = i + j \in \mathbb{Z}, \quad d_2(\lambda_1^i \lambda_2^j) = (i, j) \in \mathbb{Z}^2.$$

We note that E has a natural basis consisting of monomials of distinct bidegree:

$$\{\lambda_1^i \lambda_2^j \mid i, j \in \mathbb{Z}\}.$$

We can place a linear order on our monomials by using the lexicographic order and declaring $d_2(\lambda_1^i \lambda_2^j) > d_2(\lambda_1^{i'}, \lambda_2^{j'})$ if either $i > i'$, or $i = i'$ and $j > j'$.

Definition 4.2 This linear order on monomials allows us to extend d_1 and d_2 to all of E by the following. If $p = \sum a_{ij} \lambda_1^i \lambda_2^j$ is a finite $\mathcal{R}$ -linear combination of basis monomials, we define the (bi)degree of p to be the (bi)degree of the top monomial of p :

$$d_1(p) = \max\{d_1(\lambda_1^i \lambda_2^j) \mid a_{ij} \neq 0\}, \quad d_2(p) = \max\{d_2(\lambda_1^i \lambda_2^j) \mid a_{ij} \neq 0\}.$$

Since $\mathcal{R}$ is an integral domain, the degrees satisfy the following property:

Proposition 4.3 If $p_1, p_2 \in E$, then

$$d_1(p_1 p_2) = d_1(p_1) + d_1(p_2), \quad d_2(p_1 p_2) = d_2(p_1) + d_2(p_2).$$

Consider the two elements $\lambda_1 + \lambda_2$ and $\lambda_1 \lambda_2 \in E$. For $i \in \mathbb{Z}_{\geq 0}$ and $j \in \mathbb{Z}$ a monomial, $(\lambda_1 + \lambda_2)^i (\lambda_1 \lambda_2)^j$ is a monic element in E of bidegree $(i + j, j)$. Thus these monomials are linearly independent over $\mathcal{R}$. Equivalently, $\lambda_1 + \lambda_2$ and $\lambda_1 \lambda_2$ are algebraically independent over $\mathcal{R}$. These expressions generate a subalgebra $E^{\text{Sym}} \subseteq E$ of symmetric Laurent polynomials:

$$E^{\text{Sym}} = \mathcal{R}[\lambda_1 + \lambda_2, (\lambda_1 \lambda_2)^{\pm 1}].$$

Consequently, a basis of E^{Sym} is given by

$$\{(\lambda_1 + \lambda_2)^i (\lambda_1 \lambda_2)^j \mid i \in \mathbb{Z}_{\geq 0} \text{ and } j \in \mathbb{Z}\}.$$

We are interested in two distinguished elements $\mathbf{x}, \mathbf{y} \in E^{\text{Sym}}$ given by

$$\mathbf{x} = \lambda_1 + \lambda_2 + \lambda_1 \lambda_2 + 1 + (\lambda_1 \lambda_2)^{-1} + \lambda_2^{-1} + \lambda_1^{-1},$$

$$\begin{aligned} \mathbf{y} = & \lambda_1^2 \lambda_2 + \lambda_1 \lambda_2^2 + \lambda_1 \lambda_2 + \lambda_1 + \lambda_2 + \lambda_1 \lambda_2^{-1} + 1 \\ & + 1 + \lambda_1^{-1} \lambda_2 + \lambda_2^{-1} + \lambda_1^{-1} + (\lambda_1 \lambda_2)^{-1} + (\lambda_1 \lambda_2^2)^{-1} + (\lambda_1^2 \lambda_2)^{-1}. \end{aligned}$$

The expression defining the element $\mathbf{x}$ models the trace of a matrix in the 7-dimensional fundamental representation of G_2 , while $\mathbf{y}$ models the same for the 14-dimensional fundamental representation. When viewing the exponent (a, b) of a monomial $\lambda_1^a \lambda_2^b$ as a G_2 weight, elements $\mathbf{x}$ and $\mathbf{y}$ can be viewed as the characters of the first and second fundamental representations, recording the dimensions of weight spaces of the representations.

We have written $\mathbf{x}$ as a sum of 7 terms and $\mathbf{y}$ as a sum of 14 terms. We remark that $\mathbf{y} = e_2 - \mathbf{x}$, where e_2 is the second elementary symmetric sum (see [Definition 4.5](#) below) of the terms of $\mathbf{x}$.

Proposition 4.4 *The elements $\mathbf{x}$ and $\mathbf{y}$ are algebraically independent over $\mathcal{R}$.*

Proof We view each monomial $\mathbf{x}^i \mathbf{y}^j$ as a monic element of $E = \mathcal{R}[\lambda_1^{\pm 1}, \lambda_2^{\pm 1}]$ and compute its bidegree $d_2(\mathbf{x}^i \mathbf{y}^j) = d_2(\mathbf{x}^i) + d_2(\mathbf{y}^j) = i d_2(\mathbf{x}) + j d_2(\mathbf{y}) = (i, i) + (2j, j) = (i + 2j, i + j)$. We see the set of monomials $\{\mathbf{x}^i \mathbf{y}^j\}_{i,j \geq 0}$ consists of monic elements of E of distinct bidegrees. Therefore the monomials are linearly independent over $\mathcal{R}$, and $\mathbf{x}$ and $\mathbf{y}$ are algebraically independent. $\square$

More generally, we will be using the elements of E^{Sym} :

$$\begin{aligned} \mathbf{x}^{(i)} &= \lambda_1^i + \lambda_2^i + (\lambda_1 \lambda_2)^i + 1 + (\lambda_1 \lambda_2)^{-i} + \lambda_2^{-i} + \lambda_1^{-i}, \\ \mathbf{y}^{(i)} &= (\lambda_1^2 \lambda_2)^i + (\lambda_1 \lambda_2^2)^i + (\lambda_1 \lambda_2)^i + \lambda_1^i + \lambda_2^i + (\lambda_1 \lambda_2^{-1})^i + 1 \\ &\quad + 1 + (\lambda_1^{-1} \lambda_2)^i + \lambda_2^{-i} + \lambda_1^{-i} + (\lambda_1 \lambda_2)^{-i} + (\lambda_1 \lambda_2^2)^{-i} + (\lambda_1^2 \lambda_2)^{-i}. \end{aligned}$$

Using this notation, $\mathbf{x}^{(1)} = \mathbf{x}$ and $\mathbf{y}^{(1)} = \mathbf{y}$. We remark that $\mathbf{x}^{(i)}$ gives the i^{th} power sum ([Definition 4.5](#)) of the summands of $\mathbf{x}$, and similarly for $\mathbf{y}$.

4.2 Polynomials in $\mathcal{R}[\mathbf{x}, \mathbf{y}]$

Consider the polynomial ring in two variables $\mathcal{R}[\mathbf{x}, \mathbf{y}]$. The main goal of this section is to provide recursive definitions of two families of polynomials $P_i(\mathbf{x}, \mathbf{y})$ and $Q_i(\mathbf{x}, \mathbf{y})$ which will allow us to express $\mathbf{x}^{(i)}$ and $\mathbf{y}^{(i)}$ as polynomials of $\mathbf{x}$ and $\mathbf{y}$ defined in the previous section. We will study the polynomials from a purely algebraic perspective before we return to studying the diagrammatically defined algebra $\mathcal{A}$, which was shown in [Proposition 3.2](#) to be isomorphic to $\mathcal{R}[\mathbf{x}, \mathbf{y}]$.

Define the polynomials $P_k(\mathbf{x}, \mathbf{y})$ and $Q_k(\mathbf{x}, \mathbf{y})$ in $\mathcal{R}[\mathbf{x}, \mathbf{y}]$ by the recursive relations

$$\begin{aligned} P_k(\mathbf{x}, \mathbf{y}) &= \begin{cases} 7 & k = 0, \\ \sum_{i=1}^{k-1} (-1)^{i-1} e_i P_{k-i}(\mathbf{x}, \mathbf{y}) + (-1)^{k-1} k e_k & 0 < k < 7, \\ \sum_{i=1}^7 (-1)^{i-1} e_i P_{k-i}(\mathbf{x}, \mathbf{y}) & k \geq 7, \end{cases} \\ Q_k(\mathbf{x}, \mathbf{y}) &= \begin{cases} 14 & k = 0, \\ \sum_{i=1}^{k-1} (-1)^{i-1} f_i Q_{k-i}(\mathbf{x}, \mathbf{y}) + (-1)^{k-1} k f_k & 0 < k < 14, \\ \sum_{i=1}^{14} (-1)^{i-1} f_i Q_{k-i}(\mathbf{x}, \mathbf{y}) & k \geq 14, \end{cases} \end{aligned}$$

where the coefficients $e_i := e_i(x, y)$ and $f_i := f_i(x, y)$ are given by

$$\begin{aligned} e_0 &= 1, & f_0 &= 1, \\ e_1 &= x, & f_1 &= y, \\ e_2 &= x + y, & f_2 &= x^3 - x^2 - 2xy - x, \\ e_3 &= x^2 - y, & f_3 &= x^4 - x^3 - 3x^2y - x^2 + 2y^2 + x + y, \\ e_k &= e_{7-k} \quad \text{if } 0 \leq k \leq 7, & f_4 &= x^3y - x^3 - x^2y - 2xy^2 + x^2 + xy - y^2 + x + y, \end{aligned}$$

and

$$\begin{aligned} f_5 &= x^5 - 2x^4 - 5x^3y + 3x^2y + 6xy^2 + y^3 + 2x^2 + 5xy + 2y^2 - x, \\ f_6 &= x^4 - 3x^3y + x^2y^2 - x^2y + 4xy^2 - 2x^2 + 3xy + 2y^2 + y, \\ f_7 &= -2x^5 + 4x^4 + 6x^3y + 2x^2y^2 + 2x^3 - 4x^2y - 8xy^2 - 2y^3 - 6x^2 - 6xy - 6y^2 + 2, \\ f_k &= f_{14-k} \quad \text{if } 0 \leq k \leq 14. \end{aligned}$$

The first few terms of each sequence of polynomials are as follows:

$$\begin{aligned} P_0(x, y) &= 7, & Q_0(x, y) &= 14, \\ P_1(x, y) &= x, & Q_1(x, y) &= y, \\ P_2(x, y) &= x^2 - 2x - 2y, & Q_2(x, y) &= y^2 - 2x^3 + 2x^2 + 4xy + 2x. \end{aligned}$$

Definition 4.5 (elementary symmetric sums and power sums) Given terms $t_1, \dots, t_s$, the i^{th} elementary symmetric sum of the terms is defined to be

$$\sum_{1 \leq j_1 < j_2 < \dots < j_i \leq s} t_{j_1} t_{j_2} \cdots t_{j_i},$$

while the i^{th} power sum of the terms is defined to be

$$\sum_{j=1}^s t_j^i.$$

The following proposition relates the coefficients defining P_n and Q_n to elementary symmetric sums of the summands of $\mathbf{x}$ and $\mathbf{y}$. It can be verified by direct computation, which is made easier by use of mathematical software.

Proposition 4.6 For each $0 \leq i \leq 7$, the quantity $e_i(\mathbf{x}^{(1)}, \mathbf{y}^{(1)})$ is equal to the i^{th} elementary symmetric sum of the terms of $\mathbf{x}^{(1)}$. Similarly, for each $0 \leq i \leq 14$, the quantity $f_i(\mathbf{x}^{(1)}, \mathbf{y}^{(1)})$ is equal to the i^{th} elementary symmetric sum of the terms of $\mathbf{y}^{(1)}$.

We can now relate our polynomials P_k and Q_k to the k^{th} power sums of the terms of $\mathbf{x}$ and $\mathbf{y}$.

Proposition 4.7 For each $k \geq 0$,

$$P_k(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}) = \mathbf{x}^{(k)}, \quad Q_k(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}) = \mathbf{y}^{(k)}.$$

More generally, for each $k \geq 0$ and for each $i \geq 0$,

$$P_k(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}) = \mathbf{x}^{(ik)}, \quad Q_k(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}) = \mathbf{y}^{(ik)}.$$

Proof We consider P_k . The case of Q_k is handled similarly.

By Proposition 4.6, the quantity $e_l(\mathbf{x}^{(1)}, \mathbf{y}^{(1)})$ is equal to the l^{th} elementary symmetric sum of the terms of $\mathbf{x}^{(1)}$. Let A be a 7-by-7 diagonal matrix such that each diagonal entry A_{jj} is the j^{th} summand of $\mathbf{x}^{(1)}$. Each e_l is a coefficient of the characteristic polynomial of A . The Cayley–Hamilton theorem applied to the matrix A^k then gives

$$P_k(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}) = \text{tr } A^k = \mathbf{x}^{(k)}.$$

Next, consider the algebra endomorphism $\phi_i: E \rightarrow E$ defined on the generators by $\phi_i(\lambda_1) = \lambda_1^i$ and $\phi_i(\lambda_2) = \lambda_2^i$. Then $\phi_i(\mathbf{x}^{(1)}) = \mathbf{x}^{(i)}$ and $\phi_i(\mathbf{y}^{(1)}) = \mathbf{y}^{(i)}$.

We then use ϕ_i and the first part of the proposition to see that

$$P_k(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}) = P_k(\phi_i(\mathbf{x}^{(1)}), \phi_i(\mathbf{y}^{(1)})) = \phi_i(P_k(\mathbf{x}^{(1)}, \mathbf{y}^{(1)})) = \phi_i(\mathbf{x}^{(k)}) = \mathbf{x}^{(ik)}. \quad \square$$

Corollary 4.8 For each $k \geq 0$ and for each $i \geq 0$ the following hold in $\mathcal{R}[x, y]$:

$$P_k(P_i(x, y), Q_i(x, y)) = P_{ik}(x, y), \quad Q_k(P_i(x, y), Q_i(x, y)) = Q_{ik}(x, y).$$

Proof Since $\mathbf{x}^{(1)}$ and $\mathbf{y}^{(1)}$ are algebraically independent, there is an injective algebra homomorphism $\psi: \mathcal{R}[x, y] \rightarrow E$ defined on generators by $x \mapsto \mathbf{x}^{(1)}$ and $y \mapsto \mathbf{y}^{(1)}$. Applying ψ to the left side of the first equation in the proposition yields

$$\psi(P_k(P_i(x, y), Q_i(x, y))) = P_k(P_i(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), Q_i(\mathbf{x}^{(1)}, \mathbf{y}^{(1)})) = P_k(\mathbf{x}^{(i)}, \mathbf{y}^{(i)}) = P_{ki}(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}).$$

Applying ψ to the right side of the first equation in the proposition yields the same answer. The injectivity of ψ then proves the proposition for P_k . The case of Q_k is proven similarly. $\square$

Definition 4.9 For a monomial $x^i y^j \in \mathcal{R}[x, y]$, we define its bidegree to be

$$D_2(x^i y^j) = (i + 2j, i + j) \in \mathbb{Z}_{\geq 0}.$$

Using a lexicographic ordering of bidegrees, we extend D_2 to $\mathcal{R}[x, y]$ by defining for an element $p = \sum a_{ij} x^i y^j \in \mathcal{R}[x, y]$ its bidegree

$$D_2(p) = \max\{D_2(x^i y^j) \mid a_{ij} \neq 0\}.$$

The bidegree satisfies the following compatibility with multiplication in $\mathcal{R}[x, y]$.

Proposition 4.10 For any two polynomials $S(x, y), T(x, y) \in \mathcal{R}(x, y)$ we have

$$D_2(S(x, y)T(x, y)) = D_2(S(x, y)) + D_2(T(x, y)).$$

The following propositions provide support for the unusual choice of the bidegree D_2 :

Proposition 4.11 For $k \geq 1$, the polynomials $P_k(x, y)$ and $Q_k(x, y)$ are monic polynomials in $\mathcal{R}[x, y]$. Their bidegrees are $D_2(P_k) = (k, k)$ and $D_2(Q_k) = (2k, k)$.

Proof By inspecting the coefficients e_i and f_i in the recursive definitions of P_k and Q_k we observe inductively that $P_k = xP_{k-1} + (\text{lower terms})$ and $Q_k = yQ_{k-1} + (\text{lower terms})$. Thus $D_2(P_k) = D_2(x) + D_2(P_{k-1}) = (1, 1) + D_2(P_{k-1})$ and $D_2(Q_k) = D_2(y) + D_2(Q_{k-1}) = (2, 1) + D_2(Q_{k-1})$. The result follows by induction. $\square$

Thus P_0 and Q_0 are the only nonmonic polynomials in our sequences of polynomials. In the following proposition we renormalize P_0 and Q_0 so that $P_0(x, y) = 1 = Q_0(x, y)$ (but the other polynomials stay unchanged).

Proposition 4.12 The set of products of polynomials

$$\{P_k(x, y)Q_l(x, y)\}_{k,l \geq 0}$$

forms a basis for the 2-variable polynomial ring $\mathcal{R}[x, y]$.

Proof We observe that each $P_k Q_l$ is a monic polynomial of $\mathcal{R}[x, y]$ of distinct bidegree $D_2(P_k Q_l) = D_2(P_k) + D_2(Q_l) = (k, k) + (2l, l) = (k + 2l, k + l)$. The standard basis $\{x^k y^l\}$ of $\mathcal{R}[x, y]$ also consists of monic polynomials of distinct bidegrees $D_2(x^k y^l) = (k + 2l, k + l)$. Thus there is a unitriangular change of basis from the standard basis $\{x^k y^l\}_{k,l \geq 0}$ to $\{P_k(x, y)Q_l(x, y)\}_{k,l \geq 0}$. $\square$

Proposition 4.13 For $i, j \geq 0$ we have

$$D_2(x^i y^j) = d_2(\mathbf{x}^{(i)} \mathbf{y}^{(j)}).$$

Proof This follows from the compatibility of D_2 and d_2 with multiplication and the following observation:

$$D_2(x^i) = (i, i) = d_2(\mathbf{x}^{(i)}), \quad D_2(y^j) = (2j, j) = d_2(\mathbf{y}^{(j)}). \quad \square$$

5 Transparent elements

In this section we connect the algebra E^{Sym} to the algebra $\mathcal{A}_{1,1}$ and use the relationship to show that the polynomials P_n and Q_n are transparent when $q^{2n} = 1$. As a reminder, throughout the paper we have assumed that our ring of coefficients $\mathcal{R}$ contains $[12]^{-1}$, which gives some restriction on $\mathcal{R}$ and q .

5.1 The maps $F^\star$ and $F_\star$

We next define the maps $F^\star$ and $F_\star$, which are the key ingredients in connecting our algebra $E^{\text{Sym}} = \mathcal{R}[\lambda_1 + \lambda_2, (\lambda_1 \lambda_2)^{\pm 1}]$ and our diagrammatic algebra $\mathcal{A}_{1,1}$.

Definition 5.1 Below we define algebra maps $F^\star, F_\star: E^{\text{Sym}} \rightarrow \mathcal{A}_{1,1}$ on generators:

$$\begin{aligned} F^\star(\lambda_1 \lambda_2) &= q^2 a, & F_\star(\lambda_1 \lambda_2) &= q^{-2} a, \\ F^\star(\lambda_1 + \lambda_2) &= \frac{q}{[2]}(c - a - 1), & F_\star(\lambda_1 + \lambda_2) &= \frac{q^{-1}}{[2]}(c - a - 1). \end{aligned}$$

In particular, the map $F_\star$ is obtained from the map $F^\star$ by replacing q with q^{-1} . The following proposition provides the precise relationship between our distinguished elements of E^{Sym} and our distinguished elements of $\mathcal{A}_{1,1}$:

Proposition 5.2 $F^\star$ and $F_\star: E^{\text{Sym}} \rightarrow \mathcal{A}_{1,1}$ are well-defined algebra homomorphisms and satisfy

$$F^\star(\mathbf{x}^{(1)}) = x^\star, \quad F^\star(\mathbf{y}^{(1)}) = \bar{y} := y^\star + \frac{1}{[2]^2} f, \quad F_\star(\mathbf{x}^{(1)}) = x_\star, \quad F_\star(\mathbf{y}^{(1)}) = \underline{y} := y_\star + \frac{1}{[2]^2} f.$$

Proof We will prove the statement for $F^\star$. The statement for $F_\star$ is proved by replacing q with q^{-1} . As $\lambda_1 + \lambda_2$ and $(\lambda_1 \lambda_2)^{\pm 1}$ generate $E^{\text{Sym}} = \mathcal{R}[\lambda_1 + \lambda_2, (\lambda_1 \lambda_2)^{\pm 1}]$, we see that $F^\star$ is a well-defined algebra homomorphism by noting that $q^2 a$ is an invertible element of $\mathcal{A}_{1,1}$ and that $\mathcal{A}_{1,1}$ is a commutative algebra.

To check that $F^\star(\mathbf{x}^{(1)}) = x^\star$, we compute diagrammatically that

$$\begin{aligned} x^\star &= \text{diagram of a circle with a dot and a line connecting it to the top} = \frac{1}{[2]} \left(q^3 \text{diagram of a circle with a dot and a line connecting it to the top} + q^{-3} \text{diagram of a circle with a dot and a line connecting it to the top} + q \text{diagram of a circle with a dot and a line connecting it to the top} + q^{-1} \text{diagram of a circle with a dot and a line connecting it to the top} \right) \\ &= \frac{1}{[2]} (q^3 a + q^{-3} a^{-1} + qc + q^{-1} a^{-1} c). \end{aligned}$$

On the other hand, we compute that

$$\begin{aligned} F^\star(\mathbf{x}^{(1)}) &= F^\star(\lambda_1 + \lambda_2) + F^\star(\lambda_1 \lambda_2) + F^\star((\lambda_1 \lambda_2)^{-1}) + F^\star(1) + F^\star((\lambda_1 \lambda_2)^{-1}(\lambda_1 + \lambda_2)) \\ &= \frac{q}{[2]}(c - a - 1) + q^2 a + q^{-2} a^{-1} + 1 + q^{-2} a^{-1} \frac{q}{[2]}(c - a - 1) \\ &= \frac{1}{[2]} (q^3 a + q^{-3} a^{-1} + qc + q^{-1} a^{-1} c). \end{aligned}$$

Therefore $F^\star(\mathbf{x}^{(1)}) = x^\star$.

Now, to check that $F^*(y^{(1)}) = \bar{y} = y^* + (1/[2]^2)f$, we compute diagrammatically that

$$(16) \quad y^* = \text{diagram} = \frac{1}{[2]} \left(q^3 \text{diagram}_1 + q^{-3} \text{diagram}_2 + \frac{1}{[2]} \text{diagram}_3 \right).$$

Next we apply the double-edge removal relation to all three diagrams and then continue to fully simplify.

The following is the calculation for the final diagram of (16):

$$\begin{aligned} \text{diagram}_3 &= -\text{diagram}_1 + \frac{[4][6]}{[2][12]} \text{diagram}_4 + \text{diagram}_5 + \frac{1}{[3]} \text{diagram}_6 \\ &= -f - \frac{[6][3][8]}{[2][12]} 1 + a^{-1}c^2 + \frac{1}{[3]} ([3](a + a^{-1}) - \frac{[4]}{[2]}(c + a^{-1}c)). \end{aligned}$$

The other two diagrams of the right side of (16) are reduced using the same process and we record the results:

$$\text{diagram}_1 = -a^2 + \frac{[4][6]}{[2][12]} 1 + ac + \frac{1}{[3]}c, \quad \text{diagram}_2 = -a^{-2} + \frac{[4][6]}{[2][12]} 1 + a^{-2}c + \frac{1}{[3]}a^{-1}c.$$

Combining these computations and substituting them into (16), we write y^* in our basis for $\mathcal{A}_{1,1}$ as

$$\begin{aligned} y^* &= \frac{1}{[2]} \left(-q^3a^2 - q^{-3}a^{-2} + q^3ac + q^{-3}a^{-2}c + \frac{1}{[2]}a^{-1}c^2 \right. \\ &\quad \left. + \frac{1}{[2]}a + \frac{1}{[2]}a^{-1} + \frac{q^2-1}{[2]}c + \frac{q^{-2}-1}{[2]}a^{-1}c - \frac{q^2+q^{-2}}{[2]}1 - \frac{1}{[2]}f \right). \end{aligned}$$

On the other hand, we compute

$$\begin{aligned} F^*(y^{(1)}) &= F^*(\lambda_1 + \lambda_2) + F^*(\lambda_1^{-1} + \lambda_2^{-1}) + F^*(\lambda_1\lambda_2) + F^*(\lambda_1^{-1}\lambda_2^{-1}) + F^*(\lambda_1\lambda_2(\lambda_1 + \lambda_2)) \\ &\quad + F^*(\lambda_1^{-1}\lambda_2^{-1}(\lambda_1^{-1} + \lambda_2^{-1})) + F^*((\lambda_1 + \lambda_2)(\lambda_1^{-1} + \lambda_2^{-1})) \\ &= q \frac{c-a-1}{[2]} + q^{-1} \frac{a^{-1}(c-a-1)}{[2]} + q^2a + q^{-2}a^{-1} + q^2a \left(q \frac{c-a-1}{[2]} \right) \\ &\quad + q^{-2}a^{-1} \left(q^{-1} \frac{a^{-1}(c-a-1)}{[2]} \right) + \left(q \frac{c-a-1}{[2]} \right) \left(q^{-1} \frac{a^{-1}(c-a-1)}{[2]} \right) \\ &= \frac{1}{[2]} \left(-q^3a^2 - q^{-3}a^{-2} + q^3ac + q^{-3}a^{-2}c + \frac{1}{[2]}a^{-1}c^2 \right. \\ &\quad \left. + \frac{1}{[2]}a + \frac{1}{[2]}a^{-1} + \frac{q^2-1}{[2]}c + \frac{q^{-2}-1}{[2]}a^{-1}c - \frac{q^2+q^{-2}}{[2]}1 \right). \end{aligned}$$

Thus $F^*(y^{(1)}) = y^* + (1/[2]^2)f$, as claimed. $\square$

Proposition 5.3 *The maps $F^\star$ and $F_\star$ each give an isomorphism $E^{\text{Sym}} \cong \mathcal{R}[a^{\pm 1}, c] \subseteq \mathcal{A}_{1,1}$. In particular, $F^\star$ and $F_\star$ are injective.*

Proof We check the statement for $F^\star$. The other case is similar. By recalling the definition (Definition 5.1) of $F^\star$ on generators of E^{Sym} , we see that the image of $F^\star$ is contained in the subalgebra $\mathcal{R}[a^{\pm 1}, c] \subseteq \mathcal{A}_{1,1}$. Since $[2]$ is invertible in $\mathcal{R}$ we deduce from the definition of $F^\star$ on the generators of E^{Sym} that $a^{\pm 1}$ and c are in the image of $F^\star$. Thus $F^\star$ is surjective.

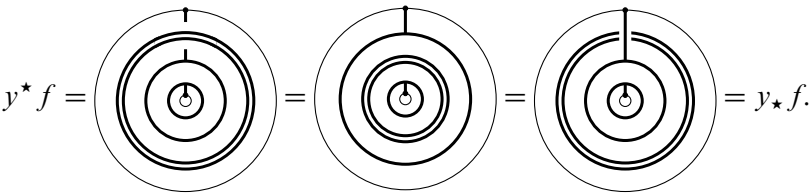
On the subalgebra $\mathcal{R}[a^{\pm 1}, c]$, for $k \in \mathbb{Z}$ and $j \in \mathbb{Z}_{\geq 0}$, we can assign the bidegree (j, k) to the monomial $a^k c^j$ and give the set of all bidegrees the lexicographic ordering. We then see that $F^\star$ sends the monomial $(\lambda_1 \lambda_2)^i (\lambda_1 + \lambda_2)^j$ to an element of $\mathcal{R}[a^{\pm 1}, c]$ whose top bidegree is $(j, i + j)$. Thus $F^\star$ sends a spanning set of E^{Sym} to a linearly independent subset of $\mathcal{A}_{1,1}$. Therefore $F^\star$ is injective. $\square$

5.2 The error term

Our original strategy was to attempt to define an isomorphism from a subalgebra of $E = \mathcal{R}[\lambda_1^{\pm 1}, \lambda_2^{\pm 1}]$ to $\mathcal{A}_{1,1}$ whose image on $\mathbf{x}^{(1)}$ was $x^\star$ and whose image on $\mathbf{y}^{(1)}$ was $y^\star$. We did not find it possible to achieve this goal on the nose since the algebra $\mathcal{A}_{1,1}$ contains zero divisors while E does not. Indeed, as observed earlier, we have the identity $af = a$ and so $(a - 1)f = 0$ while both $a - 1$ and f are nonzero by consideration of our basis for $\mathcal{A}_{1,1}$.

Instead, we found an injective algebra homomorphism $F^\star$ from the subalgebra $E^{\text{Sym}} \subset E$ onto the subalgebra $\mathcal{R}[a^{\pm 1}, c] \subset \mathcal{A}_{1,1}$. We did succeed in defining $F^\star$ so that its image on $\mathbf{x}^{(1)}$ is $x^\star$. However, its image on $\mathbf{y}^{(1)}$ is not quite $y^\star$. It is off by an error term: $F^\star(\mathbf{y}^{(1)}) = y^\star + (1/[2]^2)f$. Below we show that the error term is essentially negligible in our context since the diagram of f allows loops to pass through it.

For example, we observe the identity $y^\star f = y_\star f$ in $\mathcal{A}_{1,1}$:



More generally, we will make use of the following observation:

Lemma 5.4 *The following identities hold in $\mathcal{A}_{1,1}$:*

$$x^\star f = x_\star f, \quad y^\star f = y_\star f, \quad \bar{y}f = \underline{y}f.$$

Proof The first two identities follow diagrammatically by using a 3-dimensional isotopy to pass a loop through the empty annular region in the diagram of f . The third identity follows from the second and from the defining equations $\bar{y} := y^\star + (1/[2]^2)f$ and $\underline{y} := y_\star + (1/[2]^2)f$. $\square$

Proposition 5.5 An element $S(x, y) \in \mathcal{A}$ is transparent (Definition 3.14) if and only if

$$S(x^\star, \bar{y}) = S(x_\star, \underline{y}) \in \mathcal{A}_{1,1}.$$

Proof It will suffice to show that

$$S(x^\star, y^\star) - S(x_\star, y_\star) = S(x^\star, \bar{y}) - S(x_\star, \underline{y}).$$

Write $S(x, y) \in \mathcal{A}$ as $S(x, y) = \sum a_{ij} x^i y^j$ for $a_{ij} \in \mathcal{R}$.

Then

$$\begin{aligned} S(x^\star, y^\star) &= \sum a_{ij} (x^\star)^i (y^\star)^j = \sum a_{ij} (x^\star)^i \left(\bar{y} - \frac{1}{[2]^2} f \right)^j \\ &= \sum a_{ij} (x^\star)^i \left(\sum_{k=0}^j \binom{j}{k} (\bar{y})^k \left(-\frac{1}{[2]^2} f \right)^{j-k} \right) \\ &= \sum a_{ij} (x^\star)^i (\bar{y})^j + \sum_{k < j} a_{ij} \binom{j}{k} (x^\star)^i (\bar{y})^k \left(-\frac{1}{[2]^2} f \right)^{j-k} \\ &= S(x^\star, \bar{y}) + \sum_{k < j} a_{ij} \binom{j}{k} (x_\star)^i (\underline{y})^k \left(-\frac{1}{[2]^2} f \right)^{j-k}. \end{aligned}$$

The last equality follows from Lemma 5.4. Similarly, we compute that

$$S(x_\star, y_\star) = S(x_\star, \underline{y}) + \sum_{k < j} a_{ij} \binom{j}{k} (x_\star)^i (\underline{y})^k \left(-\frac{1}{[2]^2} f \right)^{j-k},$$

and the result follows. $\square$

5.3 The transparent elements

Recall the definition of the degree d_1 from Definition 4.1.

Proposition 5.6 Suppose $\lambda \in E^{\text{Sym}} \subseteq E = \mathcal{R}[\lambda_1^{\pm 1}, \lambda_2^{\pm 1}]$ is a homogeneous element of degree $d_1(\lambda) = k$. Then $F^\star(\lambda) = q^{2k} F_\star(\lambda)$.

Proof Since the generators $\lambda_1 + \lambda_2$ and $(\lambda_1 \lambda_2)^{\pm 1}$ are homogenous elements of degree 1 and degree ± 2 , respectively, it suffices to check the statement on these generators. We observe that $F^\star(\lambda_1 \lambda_2) = q^2 a = q^4 (q^{-2} a) = q^4 F_\star(\lambda_1 \lambda_2)$. Similarly, $F^\star(\lambda_1 + \lambda_2) = (q/[2])(c - a - 1) = q^2 (q^{-1}/[2])(c - a - 1) = q^2 F_\star(\lambda_1 + \lambda_2)$. $\square$

Recall the polynomials P_n and Q_n and their power sum properties from Section 4.2. We are now able to show that P_n and Q_n are transparent when q is a $2n^{\text{th}}$ root of unity.

Theorem 5.7 If $q \in \mathcal{R}$ satisfies $q^{2n} = 1$, then both the elements $P_n(x, y)$ and $Q_n(x, y)$ in $\mathcal{A}$ are transparent elements.

Proof We compute

$$\begin{aligned} P_n(x^\star, \bar{y}) &= P_n(F^\star(x^{(1)}), F^\star(y^{(1)})) = F^\star(P_n(x^{(1)}, y^{(1)})) = F^\star(x^{(n)}) \\ &= F^\star(\lambda_1^n + \lambda_2^n) + F^\star(\lambda_1^{-n} + \lambda_2^{-n}) + F^\star((\lambda_1 \lambda_2)^n) + F^\star((\lambda_1 \lambda_2)^{-n}) + F^\star(1) \\ &= q^{2n} F_\star(\lambda_1^n + \lambda_2^n) + q^{-2n} F_\star(\lambda_1^{-n} + \lambda_2^{-n}) + q^{4n} F_\star((\lambda_1 \lambda_2)^n) + q^{-4n} F_\star((\lambda_1 \lambda_2)^{-n}) + F_\star(1). \end{aligned}$$

Similarly, we compute

$$P_n(x_\star, \bar{y}) = F_\star(\lambda_1^n + \lambda_2^n) + F_\star(\lambda_1^{-n} + \lambda_2^{-n}) + F_\star((\lambda_1 \lambda_2)^n) + F_\star((\lambda_1 \lambda_2)^{-n}) + F_\star(1).$$

Thus when $q^{2n} = 1$ we have $P_n(x^\star, \bar{y}) = P_n(x_\star, \bar{y})$, and P_n is transparent by [Proposition 5.5](#). For the case of Q_n we perform a similar computation

$$\begin{aligned} Q_n(x^\star, \bar{y}) &= F^\star(y^{(n)}) \\ &= F^\star((\lambda_1^2 \lambda_2)^n + (\lambda_1 \lambda_2^2)^n) + F^\star((\lambda_1 \lambda_2)^n) + F^\star(\lambda_1^n + \lambda_2^n) + F^\star((\lambda_1 \lambda_2^{-1})^n + (\lambda_1^{-1} \lambda_2)^n) \\ &\quad + F^\star(1) + F^\star(1) + F^\star(\lambda_2^{-n} + \lambda_1^{-n}) + F^\star((\lambda_1 \lambda_2)^{-n}) + F^\star((\lambda_1 \lambda_2^2)^{-n} + (\lambda_1^2 \lambda_2)^{-n}) \\ &= q^{6n} F_\star((\lambda_1^2 \lambda_2)^n + (\lambda_1 \lambda_2^2)^n) + q^{4n} F_\star((\lambda_1 \lambda_2)^n) + q^{2n} F_\star(\lambda_1^n + \lambda_2^n) + F_\star((\lambda_1 \lambda_2^{-1})^n \\ &\quad + (\lambda_1^{-1} \lambda_2)^n) + F_\star(1) + F_\star(1) + q^{-2n} F_\star(\lambda_2^{-n} + \lambda_1^{-n}) + q^{-4n} F_\star((\lambda_1 \lambda_2)^{-n}) \\ &\quad + q^{-6n} F_\star((\lambda_1 \lambda_2^2)^{-n} + (\lambda_1^2 \lambda_2)^{-n}), \end{aligned}$$

to see that Q_n is transparent when $q^{2n} = 1$. □

Notation Inspired by the computations in the previous proof, and to make the proof of the upcoming [Theorem 5.9](#) cleaner, we will use the following notation:

$$\begin{aligned} \tilde{x}^{(i)} &= q^{2i}(\lambda_1^i + \lambda_2^i) + q^{-2i}(\lambda_1^{-i} + \lambda_2^{-i}) + q^{4i}(\lambda_1 \lambda_2)^i + q^{-4i}(\lambda_1 \lambda_2)^{-i} + 1, \\ \tilde{y}^{(j)} &= q^{6j}((\lambda_1^2 \lambda_2)^j + (\lambda_1 \lambda_2^2)^j) + q^{4j}(\lambda_1 \lambda_2)^j + q^{2j}(\lambda_1^j + \lambda_2^j) + (\lambda_1 \lambda_2^{-1})^j + (\lambda_1^{-1} \lambda_2)^j \\ &\quad + 1 + 1 + q^{-2j}(\lambda_2^{-j} + \lambda_1^{-j}) + q^{-4j}(\lambda_1 \lambda_2)^{-j} + q^{-6j}((\lambda_1 \lambda_2^2)^{-j} + (\lambda_1^2 \lambda_2)^{-j}). \end{aligned}$$

By the injectivity of $F_\star$, the elements $\tilde{x}^{(i)}, \tilde{y}^{(j)} \in E^{\text{Sym}}$ are the unique elements satisfying

$$F^\star(x^{(i)}) = F_\star(\tilde{x}^{(i)}), \quad F^\star(y^{(j)}) = F_\star(\tilde{y}^{(j)}).$$

We now discuss some terms of $x^{(i)} y^{(j)}$ which behave like leading terms. Recall the bidegree d_2 from [Definition 4.2](#). The top two terms of $x^{(i)} y^{(j)}$, with respect to the bidegree d_2 , are $\lambda_1^{i+2j} \lambda_2^{i+j}$ and $\lambda_1^{i+2j} \lambda_2^j$. The following lemma shows that the top terms of a monomial $x^{(s)} y^{(t)}$ cannot be canceled out by any terms from a monomial $x^{(i)} y^{(j)}$ of lower bidegree.

Lemma 5.8 Suppose $d_2(\mathbf{x}^{(i)}\mathbf{y}^{(j)}) < d_2(\mathbf{x}^{(s)}\mathbf{y}^{(t)})$ for some $s, t, i, j \in \mathbb{Z}_{\geq 0}$. Then $\mathbf{x}^{(i)}\mathbf{y}^{(j)}$ contains no term $\lambda_1^{s+2t}\lambda_2^{s+t}$ and no term $\lambda_1^{s+2t}\lambda_2^t$.

Proof Suppose the statement is false. By inspection, the highest power of λ_1 in any of the terms of $\mathbf{x}^{(i)}\mathbf{y}^{(j)}$ occurs in $\lambda_1^{i+2j}\lambda_2^{i+j}$ and $\lambda_1^{i+2j}\lambda_2^j$. By assumption, $i+2j \leq s+2t$. Thus, if $\mathbf{x}^{(i)}\mathbf{y}^{(j)}$ contains such a term, $i+2j = s+2t$. Then by assumption $s+t > i+j \geq j$ and so neither the term $\lambda_1^{i+2j}\lambda_2^{i+j}$ nor the term $\lambda_1^{i+2j}\lambda_2^j$ can be equal to $\lambda_1^{s+2t}\lambda_2^{s+t}$. For the remaining case, suppose one of the terms is equal to $\lambda_1^{s+2t}\lambda_2^t$. If $\lambda_1^{i+2j}\lambda_2^j = \lambda_1^{s+2t}\lambda_2^t$ then $j = t$ and $i = s$, which contradicts our hypothesis. If $\lambda_1^{i+2j}\lambda_2^{i+j} = \lambda_1^{s+2t}\lambda_2^t$ then $i+2j = s+2t$ and $i+j = t$. This system implies that $s = -i$, which implies that $i = s = 0$ and $j = t$, and is a contradiction. $\square$

Theorem 5.9 If $S(x, y) \in \mathcal{A}$ is a transparent element, then q^2 is a root of unity of some order n . Furthermore, if n is coprime to 3, then $S(x, y) \in \mathcal{R}[P_n, Q_n]$, where $P_n := P_n(x, y)$ and $Q_n := Q_n(x, y)$.

Proof Suppose $S(x, y) \in \mathcal{A}$ is transparent. By Proposition 4.12, $\{P_i Q_j\}_{i,j \geq 0}$ is a basis for $\mathcal{A}$. We write $S(x, y) = \sum a_{ij} P_i Q_j$.

Since $S(x, y)$ is transparent, we have by Proposition 5.5 that $S(x^*, \bar{y}) - S(x_*, \underline{y}) = 0 \in \mathcal{A}_{1,1}$.

We compute

$$\begin{aligned} S(x^*, \bar{y}) - S(x_*, \underline{y}) &= \sum a_{ij} (P_i(x^*, \bar{y}) Q_j(x^*, \bar{y}) - P_i(x_*, \underline{y}) Q_j(x_*, \underline{y})) \\ &= \sum a_{ij} (F_\star(\mathbf{x}^{(i)}\mathbf{y}^{(j)}) - F_\star(\mathbf{x}^{(i)}\mathbf{y}^{(j)})) = \sum a_{ij} (F_\star(\tilde{\mathbf{x}}^{(i)}\tilde{\mathbf{y}}^{(j)}) - F_\star(\mathbf{x}^{(i)}\mathbf{y}^{(j)})) \\ &= F_\star\left(\sum a_{ij} (\tilde{\mathbf{x}}^{(i)}\tilde{\mathbf{y}}^{(j)} - \mathbf{x}^{(i)}\mathbf{y}^{(j)})\right). \end{aligned}$$

Let $P_s Q_t$ be the unique term of $S(x, y)$ of maximal bidegree $(s+2t, s+t)$.

Using the injectivity of $F_\star$ and considering the leading terms of Lemma 5.8, we have

$$a_{st}(q^{4s+6t} - 1) = 0, \quad a_{st}(q^{2s+6t} - 1) = 0,$$

where $a_{st}(q^{4s+6t} - 1)$ is the coefficient of $\lambda_1^{s+2t}\lambda_2^{s+t}$ in $\sum a_{ij}(\tilde{\mathbf{x}}^{(i)}\tilde{\mathbf{y}}^{(j)} - \mathbf{x}^{(i)}\mathbf{y}^{(j)})$ and $a_{st}(q^{2s+6t} - 1)$ is the coefficient of $\lambda_1^{s+2t}\lambda_2^t$.

Since $a_{st} \neq 0$ and $\mathcal{R}$ is an integral domain, q^2 is a root of unity of some order n such that

$$2s + 3t \equiv 0 \pmod{n}, \quad s + 3t \equiv 0 \pmod{n}.$$

So $s, 3t \equiv 0 \pmod{n}$. If n is not divisible by 3, then $s, t \equiv 0 \pmod{n}$. In this case, there exist $k_s, k_t \in \mathbb{Z}_{\geq 0}$ such that $s = k_s n$ and $t = k_t n$.

By Corollary 4.8,

$$P_s(x, y) = P_{k_s n}(x, y) = P_{k_s}(P_n(x, y), Q_n(x, y)) \quad \text{and} \quad Q_t(x, y) = Q_{k_t}(P_n(x, y), Q_n(x, y)).$$

Thus the top term of $S(x, y)$ is in $\mathcal{R}[P_n, Q_n]$. Consequently, $S(x, y) - a_{st} P_s Q_t$ is a transparent element of lower bidegree than $S(x, y)$. A repeated application of this argument shows that $S(x, y) \in \mathcal{R}[P_n, Q_n]$. $\square$

References

- [1] **E Bodish**, *Web calculus and tilting modules in type C_2* , Quantum Topol. 13 (2022) 407–458 [MR](#) [Zbl](#)
- [2] **E Bodish**, **B Elias**, **D E V Rose**, **L Tatham**, *Type C webs*, preprint (2021) [arXiv 2103.14997](#)
- [3] **E Bodish**, **H Wu**, *Triple clasp formulas for G_2* , preprint (2021) [arXiv 2112.01007](#)
- [4] **F Bonahon**, **V Higgins**, *Central elements in the SL_d -skein algebra of a surface*, Math. Z. 308 (2024) art. id. 1 [MR](#) [Zbl](#)
- [5] **F Bonahon**, **H Wong**, *Representations of the Kauffman bracket skein algebra, I: Invariants and miraculous cancellations*, Invent. Math. 204 (2016) 195–243 [MR](#) [Zbl](#)
- [6] **S Cautis**, **J Kamnitzer**, **S Morrison**, *Webs and quantum skew Howe duality*, Math. Ann. 360 (2014) 351–390 [MR](#) [Zbl](#)
- [7] **B Elias**, *Light ladders and clasp conjectures*, preprint (2015) [arXiv 1510.06840](#)
- [8] **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Math. Ser. 49, Princeton Univ. Press (2012) [MR](#) [Zbl](#)
- [9] **C Frohman**, **R Gelca**, *Skein modules and the noncommutative torus*, Trans. Amer. Math. Soc. 352 (2000) 4877–4888 [MR](#) [Zbl](#)
- [10] **C Frohman**, **J Kania-Bartoszyńska**, **T Lê**, *Unicity for representations of the Kauffman bracket skein algebra*, Invent. Math. 215 (2019) 609–650 [MR](#) [Zbl](#)
- [11] **C Frohman**, **A S Sikora**, *$SU(3)$ -skein algebras and webs on surfaces*, Math. Z. 300 (2022) 33–56 [MR](#) [Zbl](#)
- [12] **I Ganev**, **D Jordan**, **P Safronov**, *The quantum Frobenius for character varieties and multiplicative quiver varieties*, J. Eur. Math. Soc. (JEMS) 27 (2025) 3023–3084 [MR](#)
- [13] **V Higgins**, *Triangular decomposition of SL_3 skein algebras*, Quantum Topol. 14 (2023) 1–63 [MR](#) [Zbl](#)
- [14] **G Kuperberg**, *The quantum G_2 link invariant*, Internat. J. Math. 5 (1994) 61–85 [MR](#) [Zbl](#)
- [15] **G Kuperberg**, *Spiders for rank 2 Lie algebras*, Comm. Math. Phys. 180 (1996) 109–151 [MR](#) [Zbl](#)
- [16] **T T Q Lê**, *On Kauffman bracket skein modules at roots of unity*, Algebr. Geom. Topol. 15 (2015) 1093–1117 [MR](#) [Zbl](#)
- [17] **T T Q Lê**, *Triangular decomposition of skein algebras*, Quantum Topol. 9 (2018) 591–632 [MR](#) [Zbl](#)
- [18] **W B R Lickorish**, *An introduction to knot theory*, Graduate Texts in Math. 175, Springer (1997) [MR](#) [Zbl](#)
- [19] **T Mandel**, **F Qin**, *Bracelets bases are theta bases*, preprint (2023) [arXiv 2301.11101](#)
- [20] **S Morrison**, *The braid group surjects onto G_2 tensor space*, Pacific J. Math. 249 (2011) 189–198 [MR](#) [Zbl](#)
- [21] **H Morton**, **A Pokorný**, **P Samuelson**, *The Kauffman skein algebra of the torus*, Int. Math. Res. Not. 2023 (2023) 855–900 [MR](#) [Zbl](#)
- [22] **H Morton**, **P Samuelson**, *The HOMFLYPT skein algebra of the torus and the elliptic Hall algebra*, Duke Math. J. 166 (2017) 801–854 [MR](#) [Zbl](#)
- [23] **H Murakami**, **T Ohtsuki**, **S Yamada**, *HOMFLY polynomial via an invariant of colored plane graphs*, Enseign. Math. 44 (1998) 325–360 [MR](#) [Zbl](#)
- [24] **M Nesterenko**, **J Patera**, **A Tereszkiewicz**, *Orthogonal polynomials of compact simple Lie groups*, Int. J. Math. Math. Sci. (2011) art. id. 969424 [MR](#) [Zbl](#)

- [25] **T Ohtsuki, S Yamada**, *Quantum $SU(3)$ invariant of 3-manifolds via linear skein theory*, J. Knot Theory Ramifications 6 (1997) 373–404 [MR](#) [Zbl](#)
- [26] **H Queffelec**, *gl_2 foam functoriality and skein positivity*, preprint (2022) [arXiv 2209.08794](#)
- [27] **T Sakamoto, Y Yonezawa**, *Link invariant and G_2 web space*, Hiroshima Math. J. 47 (2017) 19–41 [MR](#) [Zbl](#)
- [28] **A Savage, B W Westbury**, *Quantum diagrammatics for F_4* , J. Pure Appl. Algebra 228 (2024) art. id. 107731 [MR](#) [Zbl](#)
- [29] **A S Sikora, B W Westbury**, *Confluence theory for graphs*, Algebr. Geom. Topol. 7 (2007) 439–478 [MR](#) [Zbl](#)
- [30] **D P Thurston**, *Positive basis for surface skein algebras*, Proc. Natl. Acad. Sci. USA 111 (2014) 9725–9732 [MR](#) [Zbl](#)
- [31] **W Yuasa**, *A graphical categorification of the two-variable Chebyshev polynomials of the second kind*, preprint (2019) [arXiv 1903.01099](#)

BB-G: Department of Mathematics, Lewis and Clark College
Portland, OR, United States

EB, VH, JMM: Department of Mathematics, Michigan State University
East Lansing, MI, United States

AH: Department of Mathematics, Grand Valley State University
Allendale, MI, United States

JP: Department of Mathematics and Computer Science, Wheaton College
Wheaton, IL, United States

bodiebg@lclark.edu, brodskye@msu.edu, higginsv@math.ucla.edu, hoganal@mail.gvsu.edu,
melbyjosephm@gmail.com, josh.piazza@my.wheaton.edu

Received: 22 January 2024 Revised: 15 May 2024

On contact mapping classes of prequantizations

SOUHEIB ALLOUT

MURAT SAĞLAM

We present examples of prequantizations over integral symplectic manifolds which admit infinitely many smoothly trivial contact mapping classes. These classes are given by the connected components of the strict contactomorphism group which project to the identity component of the symplectomorphism group of the base manifold. Along the way, we study the lifting problem of symplectomorphisms of the base manifold to strict contactomorphisms of the prequantization.

[53D10](#), [53D12](#), [53D22](#)

1 Introduction

One of the fundamental questions in contact topology is to understand the group $\text{Cont}(V, \xi)$ of coorientation preserving contactomorphisms of a given cooriented contact structure ξ on a closed manifold V . Understanding $\text{Cont}(V, \xi)$ requires studying its identity component $\text{Cont}^0(V, \xi)$ and the contact mapping class group $\pi_0(\text{Cont}(V, \xi))$. The aim of this paper is to study $\pi_0(\text{Cont}(V, \xi))$ for a particular class of contact manifolds called prequantizations. These are total spaces of principal circle bundles $V \rightarrow M$, with a connection 1-form λ which is a contact form whose Reeb vector field generates the circle action and, therefore, the curvature form of λ is an integral symplectic form ω on M .

Concerning the contact mapping class group, it is interesting to study the homomorphism

$$\mathcal{J}: \pi_0(\text{Cont}(V, \xi)) \rightarrow \pi_0(\text{Diff}(V))$$

where $\text{Diff}(V)$ is the group of diffeomorphisms of V which preserve the orientation determined by the coorientation of ξ . Our main results, which rely solely on algebraic topological methods, provide a large class of examples of prequantizations (V, λ) for which the kernel of $\mathcal{J}$ contains an infinite subgroup of $\pi_0(\text{Cont}(V, \xi))$. This subgroup is given as follows. Given a prequantization $(V, \lambda) \rightarrow (M, \omega)$, we get a natural homomorphism

$$(1) \quad \pi_0(\text{SCont}(V, \lambda)) \rightarrow \pi_0(\text{Symp}(M, \omega))$$

where $\text{SCont}(V, \lambda)$ is the group of strict contactomorphisms, that is, contactomorphisms which preserve the contact form λ , and $\text{Symp}(M, \omega)$ is the group of symplectomorphisms of (M, ω) . It turns out that

the kernel of the above homomorphism is isomorphic to $H_{dR}^1(M, \mathbb{Z}) / \Gamma_\omega$, where Γ_ω is the flux group of (M, ω) ; see [Theorem 2.10](#). Then we show that for a large class of integral symplectic manifolds for which Γ_ω turns out to be trivial, the composition

$$\mathcal{J}: H_{dR}^1(M, \mathbb{Z}) \rightarrow \pi_0(\text{SCont}(V, \lambda)) \rightarrow \pi_0(\text{Cont}(V, \xi))$$

is injective. More precisely, we prove the following.

Theorem 1.1 *Let (M, ω) be a closed integral symplectic manifold such that either $[\omega]$ vanishes on $\pi_2(M)$ or (M, ω) is monotone on $\pi_2(M)$ with a nonvanishing monotonicity constant. Suppose, in addition, that any abelian subgroup of $\pi_1(M)$ has rank at most one. If $(V, \lambda) \rightarrow (M, \omega)$ is a prequantization, then the homomorphism $\mathcal{J}$ is injective.*

The idea of the proof is as follows. For a given class $\alpha \in H_{dR}^1(M, \mathbb{Z})$, the associated strict contactomorphism L_α acts on a loop δ in the conformally symplectic frame bundle $\text{CSp}(\xi)$, up to free homotopy, by concatenation with an iterate of the *Reeb frame loop*, which is defined by picking a symplectic frame of a contact hyperplane and pushing it forward by the Reeb flow along a circle fiber. The amount of iteration of the Reeb frame loop is $\langle [\alpha], [\bar{\delta}] \rangle$, where $\bar{\delta}$ is the projection of δ to M . Then, we associate a Maslov-type index to the Reeb frame loop which we use to show that the resulting loop $L_\alpha \circ \delta$ is not freely homotopic to δ whenever $\langle [\alpha], [\bar{\delta}] \rangle \neq 0$. Using this Maslov-type index, we show the following statement:

Proposition 1.2 *Let $(V, \lambda) \rightarrow (M, \omega)$ be a prequantization. Suppose that either $[\omega]$ vanishes on $\pi_2(M)$ or (M, ω) is monotone on $\pi_2(M)$ with nonvanishing monotonicity constant. Then the Reeb frame loop is of infinite order in $\pi_1(\text{CSp}(\xi))$. In particular, the Reeb flow, seen as a loop in $\text{Cont}(V, \xi)$, is of infinite order in $\pi_1(\text{Cont}(V, \xi))$.*

We note that the Maslov index that we associate to the Reeb frame loop can be seen as a reformulation of the first Maslov index, which is a homomorphism from $\pi_1(\text{Cont}(V, \xi))$ to $\mathbb{Z}$, of Casals, Ginzburg and Presas [\[7\]](#), applied to the Reeb loop (see also [\[8; 9\]](#)). A similar result was also obtained by Albers, Shelukhin, and Zapolsky, as an application (among other applications) of some spectral invariants which they associate to contactomorphisms of prequantization bundles (see [\[1\]](#) for more details).

Theorem 1.3 *Let (Σ_g, ω_0) be a surface of genus $g \geq 2$ with an integral symplectic form and ω_{FS} be the Fubini–Study form on $\mathbb{CP}^n$, which is normalized so that the standard $\mathbb{CP}^1 \subset \mathbb{CP}^n$ has unit area. Consider a prequantization $(V, \lambda) \rightarrow (\mathbb{CP}^n \times \Sigma_g, \omega_{\text{FS}} \oplus \omega_0)$. Then the homomorphism*

$$\mathcal{J}: H_{dR}^1(\Sigma_g, \mathbb{Z}) \rightarrow \pi_0(\text{Cont}(V, \xi))$$

is injective. Moreover, $\ker(\mathcal{J})$ contains the image of $\mathcal{J}$ for n odd, and contains $\mathcal{J}(2H_{dR}^1(\Sigma_g, \mathbb{Z}))$ for n even.

More generally, we describe a setting for which the composition $\mathcal{J} \circ \mathcal{J}$ fails to be injective; see [Proposition 3.12](#). In principle, this provides examples other than the ones given in the above theorem.

In what follows we give an account of previous results in the direction of this work. The first result we want to mention is given by Giroux [16], which says that $\mathcal{J}$ is injective for $V = \mathbb{T}^3$ and ξ the standard tight contact structure (see also [13, Proposition 2]). The first example, according to our knowledge, of a closed contact manifold for which the kernel of $\mathcal{J}$ is nontrivial is attributed to Chekanov; see Eliashberg and Fraser [12] and Vogel [25]. The statement here is that for any overtwisted contact structure ξ on $\mathbb{S}^3$, $\pi_0(\text{Cont}(\mathbb{S}^3, \xi))$ is either $\mathbb{Z}_2$ or $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ depending on the Hopf invariant. Later in [11], Dymara introduced an invariant for overtwisted contact structures on $\mathbb{S}^3$, which enabled her to show that a homomorphism from $\pi_0(\text{Cont}(\mathbb{S}^3, \xi))$ to $\mathbb{Z}_2$ that she constructed is surjective for every overtwisted ξ .

The first example where the contact structure is tight and also the kernel of $\mathcal{J}$ is infinite was given by Ding and Geiges in [10], following an observation made by Gompf in [18]. Here the contact manifold is $\mathbb{S}^1 \times \mathbb{S}^2$ with the standard tight contact structure. Later in [17], Giroux and Massot showed that the kernel of $\mathcal{J}$ is $\mathbb{Z}_d$ for V the d -fold fiber cover of the unit cotangent bundle of a suitable surface of genus at least 2. In [21], Massot and Niederkrüger constructed hypertight contact manifolds in every odd dimension for which the kernel of $\mathcal{J}$ contains nontrivial but finite-order elements. Lantat and Zapolsky [20] then constructed an embedding of certain (braid) subgroups G of the symplectic mapping class group of the A_m -Milnor fiber, which is noncompact and exact, to the contact mapping class group of the associated contactization and showed that G is contained in the kernel of $\mathcal{J}$.

The last two works we want to mention are due to Gironella. In [14] he constructed examples of overtwisted contact manifolds, in every odd dimension, admitting smoothly trivial finite order contact mapping classes. Lastly, in [15] Gironella generalized the constructions in [10] to provide examples of tight contact manifolds in all odd dimensions, which admit smoothly trivial contact mapping classes of infinite order.

Questions and further discussion

Given a prequantization bundle (V, λ) over an integral symplectic manifold (M, ω) , the kernel of the homomorphism (1) is naturally identified with $H_{dR}^1(M, \mathbb{Z}) / \Gamma_\omega$ — see Theorem 2.10 — and one obtains a homomorphism

$$\mathcal{J}: H_{dR}^1(M, \mathbb{Z}) / \Gamma_\omega \rightarrow \pi_0(\text{Cont}(V, \xi)).$$

We do not know any example for which the homomorphism $\mathcal{J}$ fails to be injective.

One may perhaps find situations where the image of $\mathcal{J}$ contains elements which are isotopic to the identity in the group of formal contactomorphisms, namely the group of pairs (φ, Φ) , where $\varphi \in \text{Diff}(V)$ and Φ is a bundle map of TV covering φ and preserving ξ conformally symplectically. A candidate setting for this is when the Reeb frame loop is contractible, in other words, when M admits an element in $\pi_2(M)$ of unit symplectic area on which $c_1(TM)$ vanishes.

On the other hand, the noninjectivity of $\mathcal{J}$ has the potential to provide examples of strict contactomorphisms that are contact isotopic to the identity but without translated points in the sense of Sandon [23]. The

setting of such an example could be that there is some $\alpha \in \ker \mathcal{F}$ represented by a nonsingular closed 1-form on M such that the time-one-map of its symplectic gradient flow is without fixed points. Recently, Cant [6] provided examples of contactomorphisms on standard contact spheres that are contact isotopic to identity yet without translated points. However there is no known example of a strict contactomorphism with these properties.

Outline of the paper

In Section 2, we discuss generalities about prequantizations, and study the homomorphism (1). In Section 3, we give our main results, namely Theorem 3.11, Proposition 3.12 and Corollary 3.13.

Acknowledgements

We thank Felix Schlenk for initiating this work, and for many useful discussions and comments on this manuscript. We thank Alberto Abbondandolo, Stefan Nemirovski and Stefan Suhr for very helpful discussions and comments on this manuscript and also for providing encouragement. We also thank Hansjörg Geiges for his comments during the presentation of this work in a series of lectures in the Arbeitsgemeinschaft Symplektische Topologie at University of Cologne. Lastly, we thank the referee for the comments on this paper.

Allout is fully and Sağlam is partially supported by the SFB/TRR 191 “Symplectic Structures in Geometry, Algebra and Dynamics”, funded by the DFG (Projektnummer 281071066 - TRR 191).

2 Generalities on strict contactomorphisms of prequantizations

In this section we fix our setting, recall some generalities on prequantizations and determine the kernel of the homomorphism (1). Most of the results presented in this section can be found in the classical books of Banyaga [2] and Souriau [24]. We try to keep our presentation minimal and we refer to [2; 24] for further details.

Let (M, ω) be a $2n$ -dimensional closed integral symplectic manifold, that is, the value of $[\omega]$ over any integral homology 2-cycle is an integer. Then a classical fact due to Kobayashi provides our main object of study. We fix $\mathbb{S}^1 := \mathbb{R}/\mathbb{Z}$ for the rest of this manuscript.

Theorem 2.1 [4; 19] *Let $\pi: V \rightarrow M$ be a principal $\mathbb{S}^1$ -bundle whose Euler class is a lift of $-[\omega]$. Then there is a contact form λ on the total space V such that*

- (p1) *the vector field that generates the $\mathbb{S}^1$ -action is the Reeb vector field of λ ;*
- (p2) *$\pi^*\omega = d\lambda$.*

We call such a bundle (V, λ) , together with a contact form as above, a *prequantization* over (M, ω) .

Given a prequantization (V, λ) over (M, ω) we want to understand the induced homomorphism

$$(2) \quad \text{SCont}(V, \lambda) \rightarrow \text{Symp}(M, \omega), \quad \varphi \mapsto \bar{\varphi},$$

where $\text{SCont}(V, \lambda)$ is the group of strict contactomorphisms of (V, λ) , that is, diffeomorphisms of V that preserve λ , and $\text{Symp}(M, \omega)$ is the group of symplectomorphisms of (M, ω) . Notice that any strict contactomorphism φ of (V, λ) is in particular an $\mathbb{S}^1$ -equivariant diffeomorphism of V and therefore descends to a diffeomorphism $\bar{\varphi}$ of M , which preserves ω due to (p2).

Our main objective is to characterize the image and the kernel of the map (2) and of the induced map

$$(3) \quad \pi_0(\text{SCont}(V, \lambda)) \rightarrow \pi_0(\text{Symp}(M, \omega))$$

on mapping class groups. The surjectivity of the map (2), and of the map (3), relies on the answer to the question whether a given $\bar{\varphi} \in \text{Symp}(M, \omega)$ can be lifted (maybe after an isotopy) to a map $\varphi \in \text{SCont}(V, \lambda)$.

In order to study this lifting problem, we fix the $\mathbb{S}^1$ -bundle $\pi: V \rightarrow M$ as in Theorem 2.1 and consider the space $\mathcal{P}$ of all contact forms on V that satisfy (p1) and (p2). Notice that $\mathcal{P}$ is an affine space given by

$$(4) \quad \mathcal{P} = \{\lambda + \beta \mid \beta \in \Omega_{\text{cl}}^1(V) \text{ and } \beta(X) = 0\} = \{\lambda + \pi^* \alpha \mid \alpha \in \Omega_{\text{cl}}^1(M)\}$$

for any fixed $\lambda \in \mathcal{P}$, where Ω_{cl}^* denotes the space of closed forms. Let $\text{Diff}_{\mathbb{S}^1}(V)$ be the group of $\mathbb{S}^1$ -equivariant diffeomorphisms of V . We consider the action of the subgroups

$$(5) \quad \text{Diff}_{\mathbb{S}^1}^\omega(V) = \{\varphi \in \text{Diff}_{\mathbb{S}^1}(V) \mid \bar{\varphi}^* \omega = \omega\}$$

and

$$(6) \quad \text{Diff}_{\mathbb{S}^1}^{\text{Id}}(V) = \{\varphi \in \text{Diff}_{\mathbb{S}^1}(V) \mid \bar{\varphi} = \text{Id}\}.$$

on $\mathcal{P}$ by the pull-back. We note that $\varphi \in \text{Diff}_{\mathbb{S}^1}^\omega(V)$ if and only if $\varphi \in \text{Diff}_{\mathbb{S}^1}(V)$ and $\varphi^*(d\lambda) = d\lambda$ for all $\lambda \in \mathcal{P}$ and the stabilizer of any $\lambda \in \mathcal{P}$ in $\text{Diff}_{\mathbb{S}^1}^\omega(V)$ is precisely $\text{SCont}(V, \lambda)$. On the other hand we have the identification

$$(7) \quad \text{Diff}_{\mathbb{S}^1}^{\text{Id}}(V) \cong C^\infty(M, \mathbb{S}^1)$$

since any $\varphi \in \text{Diff}_{\mathbb{S}^1}^{\text{Id}}(V)$ is given by

$$(8) \quad \varphi(p) = \phi_f^X(\pi(p))(p)$$

for a unique smooth map $f: M \rightarrow \mathbb{S}^1$, where ϕ_f^X denotes the flow of the vector field X that generates the $\mathbb{S}^1$ -action.

Theorem 2.2 *Let (V, λ) be a prequantization over (M, ω) and assume that $H^2(M, \mathbb{Z})$ is torsion-free. Then the homomorphism (3) is surjective.*

Proof We first show that the identity component of $\text{Diff}_{\mathbb{S}^1}^\omega(V)$ acts transitively on $\mathcal{P}$. In fact, given the line segment $\lambda_t = \lambda + t\beta$, $t \in [0, 1]$, joining λ and $\lambda + \beta$, we define the vector field Y on V by the

equations $\lambda(Y) = 0$ and $\beta = -\iota_Y d\lambda$. Note that $\dot{\lambda}_t = \beta$, $d\lambda_t = d\lambda$ and $\lambda_t(Y) = 0$ for all $t \in [0, 1]$. Now let ϕ_t be the flow of Y . We get

$$\frac{d}{dt}\phi_t^*\lambda_t = \phi_t^*(\dot{\lambda}_t + \mathcal{L}_Y\lambda_t) = \phi_t^*(\beta + \iota_Y d\lambda_t + d(\lambda_t(Y))) = \phi_t^*(\beta + \iota_Y d\lambda) = 0.$$

Hence $\phi_t^*(\lambda + t\beta) = \lambda$ and $\phi_t \in \text{Diff}_{\mathbb{S}^1}^\omega(V)$ for all $t \in [0, 1]$. In fact $[X, Y] = 0$ since

$$0 = \lambda([X, Y]) = X(\lambda(Y)) - Y(\lambda(X)) - d\lambda(X, Y)$$

and

$$0 = \iota_{[X, Y]}d\lambda = \mathcal{L}_X\iota_Y d\lambda - \iota_Y \mathcal{L}_X d\lambda = -\mathcal{L}_X\beta = 0.$$

Note that any $\varphi \in \text{Symp}(M, \omega)$ lifts to some $\tilde{\varphi} \in \text{Diff}_{\mathbb{S}^1}^\omega(V)$ since φ pulls back the circle bundle $V \rightarrow M$ to a circle bundle on M , which has the same Euler class since $H^2(M, \mathbb{Z})$ is torsion-free. The resulting bundle equivalence provides $\tilde{\varphi}$. Then there is some ψ in the identity component of $\text{Diff}_{\mathbb{S}^1}^\omega(V)$ such that $\psi^*(\tilde{\varphi}^*\lambda) = \lambda$. So $\tilde{\varphi} \circ \psi \in \text{SCont}(V, \lambda)$ and $\overline{\tilde{\varphi} \circ \psi}$ is isotopic to φ in $\text{Symp}(M, \omega)$. $\square$

Lemma 2.3 *Let $\lambda \in \mathcal{P}$ and $\varphi \in \text{SCont}(M, \lambda)$ be such that $\bar{\varphi} = \text{Id}$. Then $\varphi = \phi_t^X$ for some $t \in \mathbb{S}^1$.*

Proof Given $\varphi \in \text{Diff}_{\mathbb{S}^1}^{\text{dd}}(V)$ as in (8), an easy computation shows that

$$(9) \quad \varphi^*\lambda = \lambda + \pi^*df$$

where df is the differential of $f: M \rightarrow \mathbb{S}^1$ viewed as the 1-form f^*dt . Now if φ preserves λ , then f is a constant map and therefore $\varphi = \phi_t^X$ for some $t \in \mathbb{S}^1$. $\square$

Theorem 2.4 *Let (V, λ) be a prequantization over (M, ω) and assume that $H^2(M, \mathbb{Z})$ is torsion-free. Then the following are equivalent.*

- (1) *The action of $\text{Diff}_{\mathbb{S}^1}^{\text{dd}}(V)$ on $\mathcal{P}$ is transitive.*
- (2) *The first cohomology group $H_{dR}^1(M, \mathbb{R})$ is trivial.*
- (3) *The homomorphism (2) is surjective. In other words, the sequence*

$$0 \rightarrow \mathbb{S}^1 \rightarrow \text{SCont}(V, \lambda) \rightarrow \text{Symp}(M, \omega) \rightarrow 0$$

is exact.

Proof We recall that every class in $H_{dR}^1(M, \mathbb{Z})$ is represented by the differential of a circle valued function. Combining this fact with (9) gives the noncanonical identification

$$\mathcal{P} / \text{Diff}_{\mathbb{S}^1}^{\text{dd}}(V) \simeq H_{dR}^1(M, \mathbb{R}) / H_{dR}^1(M, \mathbb{Z}),$$

which shows the equivalence of (1) and (2).

Now assume that $\text{Diff}_{\mathbb{S}^1}^{\text{dd}}(V)$ acts transitively on $\mathcal{P}$ and let $\varphi \in \text{Symp}(M, \omega)$. Then φ lifts to some $\tilde{\varphi} \in \text{Diff}_{\mathbb{S}^1}^\omega(V)$ as in the proof of Theorem 2.2. We have some $\psi \in \text{Diff}_{\mathbb{S}^1}^{\text{dd}}(V)$ such that $\psi^*(\tilde{\varphi}^*\lambda) = \lambda$. Hence $\tilde{\varphi} \circ \psi$ is in $\text{SCont}(V, \lambda)$ and lifts φ . Together with Lemma 2.3 we get the exact sequence above. For the converse statement, we take $\lambda' \in \mathcal{P}$ and by the proof of Theorem 2.2, there is some φ in the

identity component of $\text{Diff}_{\mathbb{S}^1}^\omega(V)$ such that $\varphi^*\lambda' = \lambda$ and $\bar{\varphi} \in \text{Symp}(M, \omega)$. By assumption there is a lift $\tilde{\varphi} \in \text{SCont}(V, \lambda)$ of $\bar{\varphi}$. Then we get $\varphi \circ \tilde{\varphi}^{-1} \in \text{Diff}_{\mathbb{S}^1}^{\text{Id}}(V)$ and $(\varphi \circ \tilde{\varphi}^{-1})^*\lambda' = \lambda$. $\square$

Remark 2.5 The assumption that $H^2(M, \mathbb{Z})$ is torsion-free is essential for Theorems 2.2 and 2.4. It is possible to construct an integral symplectic manifold (M, ω) with nontrivial torsion in $H^2(M, \mathbb{Z})$ and a symplectomorphism that does not preserve a class $e \in H^2(M, \mathbb{Z})$ that lifts $-\omega$. Such a symplectomorphism cannot be lifted even to an $\mathbb{S}^1$ -equivariant map of the circle bundle with Euler class e .

Given a prequantization $(V, \lambda) \rightarrow (M, \omega)$, we now want to understand the kernel of the map (3). To this end we define the subgroups

$$\mathcal{L} := \{\varphi \in \text{Symp}(M, \omega) \mid \varphi \text{ lifts to } \text{SCont}(V, \lambda)\}$$

and

$$\mathcal{H} := \mathcal{L} \cap \text{Symp}^0(M, \omega)$$

where $\text{Symp}^0(M, \omega)$ is the identity component.

By Lemma 2.3, the map (2) now fits into the exact sequence

$$0 \rightarrow \mathbb{S}^1 \rightarrow \text{SCont}(V, \lambda) \rightarrow \mathcal{L} \rightarrow 0$$

which defines a Serre fibration. In fact, it is easy to see that once restricted to the identity components, the above sequence reads

$$0 \rightarrow \mathbb{S}^1 \rightarrow \text{SCont}^0(V, \lambda) \rightarrow \text{Ham}(M, \omega) \rightarrow 0$$

where $\text{Ham}(M, \omega)$ is the group of Hamiltonian diffeomorphisms. Then the homotopy lifting property for 0-cells follows from the nontrivial fact that any path of Hamiltonian diffeomorphism starting at Id is a Hamiltonian isotopy. For the details, we refer to [3].

An immediate consequence of the above lemma is the following exact sequence of mapping class groups:

$$0 \rightarrow \pi_0(\text{SCont}(V, \lambda)) \rightarrow \pi_0(\mathcal{L}) \rightarrow 0.$$

On the other hand if $H^2(M, \mathbb{Z})$ is torsion-free we have the exact sequence

$$0 \rightarrow \pi_0(\mathcal{H}) \rightarrow \pi_0(\mathcal{L}) \rightarrow \pi_0(\text{Symp}(M, \omega)) \rightarrow 0.$$

It turns out that the elements of $\mathcal{L}$ are characterized by their effect on holonomy and this effect is explicitly determined for the elements of $\mathcal{H}$. Let $\gamma: [0, 1] \rightarrow M$ be a piecewise smooth curve and $p \in \pi^{-1}(\gamma(0))$. Then there exists a unique curve $\hat{\gamma}: [0, 1] \rightarrow V$ such that

$$\hat{\gamma}(0) = p, \quad \pi \circ \hat{\gamma} = \gamma, \quad \lambda(\dot{\hat{\gamma}}) = 0.$$

We call $\hat{\gamma}$, suppressing its initial condition, the *horizontal lift* of γ . If γ is a loop, that is $\gamma(0) = \gamma(1)$, then the *holonomy* $\text{Hol}(\gamma) \in \mathbb{S}^1$ of γ is defined by

$$\phi_{\text{Hol}(\gamma)}^X(\hat{\gamma}(1)) = \hat{\gamma}(0).$$

We note that the holonomy does not depend on $\tilde{\gamma}(0)$ and it is invariant under positive reparameterizations of γ . Moreover, $\text{Hol}(\gamma^{-1}) = -\text{Hol}(\gamma)$ where γ^{-1} is some (hence any) negative reparameterization of γ . Now given a diffeomorphism φ of M one considers its effect on holonomy, namely the map

$$(10) \quad E_\varphi: \text{Loop}(M) \rightarrow \mathbb{S}^1, \quad \gamma \mapsto \text{Hol}(\varphi \circ \gamma) - \text{Hol}(\gamma),$$

where $\text{Loop}(M)$ is the space of piecewise smooth maps from $\mathbb{S}^1$ to M . The following fact, which we spell out in accordance with our purposes, is due to Kostant and says that the holonomy effect determines whether a symplectomorphism φ lifts to $\text{SCont}(V, \lambda)$.

Theorem 2.6 (Kostant [22, Proposition 3.3]) *A symplectomorphism $\varphi \in \text{Symp}(M, \omega)$ admits a lift to $\text{SCont}(V, \lambda)$ if and only if $E_\varphi = 0$.*

Proof We fix a point $x \in M$ and put $y = \varphi(x)$. After trivializing the circle bundle near x and y , we fix a map $\eta: \pi^{-1}(x) \rightarrow \pi^{-1}(y)$, which reads as a rotation with respect to any other choice of such trivializations. Now given $p \in V$ we take a path γ such that $\gamma(0) = \pi(p)$ and $\gamma(1) = x$ and we define $\tilde{\varphi}(p)$ by the equation

$$\widehat{\varphi \circ \gamma}(1) = \eta(\hat{\gamma}(1))$$

where $\widehat{\varphi \circ \gamma}$ is the horizontal lift of $\varphi \circ \gamma$ with $\widehat{\varphi \circ \gamma}(0) = \tilde{\varphi}(p)$ and $\hat{\gamma}$ is the horizontal lift of γ with $\hat{\gamma}(0) = p$.

Note that $\tilde{\varphi}$ is well defined if and only if $E_\varphi = 0$. Moreover $\tilde{\varphi}$, whenever is well defined, is an $\mathbb{S}^1$ -equivariant lift of φ , which maps horizontal lifts of curves in M to horizontal lifts of their images under φ . It is not difficult to show that this property together with the fact that φ is symplectic implies that $\tilde{\varphi}$ is in $\text{SCont}(V, \lambda)$. $\square$

Given $\varphi \in \text{Symp}^0(M, \omega)$, it is easy to understand its effect on holonomy. We begin with the following observation.

Lemma 2.7 *Let $\gamma_1, \gamma_2, \dots, \gamma_k: \mathbb{S}^1 \rightarrow M$ be a family of smooth loops. Suppose there is an oriented surface Σ with boundary $\partial\Sigma = S_1 \cup S_2 \cup \dots \cup S_k$ consisting of k circles and a smooth map $\sigma: \Sigma \rightarrow M$ extending each γ_j , that is, for each j , once S_j is oriented as the boundary of Σ and identified with $\mathbb{S}^1$, $\sigma|_{S_j} = \gamma_j$. Then we have*

$$\sum_{j=1}^k \text{Hol}(\gamma_j) = \int_{\Sigma} \sigma^* \omega \mod \mathbb{Z}.$$

Proof We first note that for any smooth $\gamma: \mathbb{S}^1 \rightarrow M$,

$$\text{Hol}(\gamma) = \int_{\tilde{\gamma}} \lambda \mod \mathbb{Z}$$

where $\tilde{\gamma}$ is any loop that lifts γ . In fact, given a lift $\tilde{\gamma}: \mathbb{S}^1 \rightarrow V$, we consider the horizontal lift $\hat{\gamma}$ of γ with $\hat{\gamma}(0) = \tilde{\gamma}(0)$ and define a smooth map $\tau: [0, 1] \rightarrow \mathbb{S}^1$ such that

$$\phi_{\tau(s)}^X(\hat{\gamma}(s)) = \tilde{\gamma}(s).$$

Note that $\tau(1) = \text{Hol}(\gamma)$ modulo $\mathbb{Z}$. We lift τ to a map into $\mathbb{R}$, still denoted by τ , such that $\tau(0) = 0$ and consider the smooth map

$$\eta: \Delta \rightarrow V, \quad (s, t) \mapsto \phi_t^X(\hat{\gamma}(s))$$

on the domain $\Delta := \{(s, t) \in \mathbb{R}^2 \mid 0 \leq t \leq \tau(s), s \in [0, 1]\}$. Note that $\eta^* d\lambda = 0$. By Stokes' theorem,

$$0 = \int_{\Delta} \eta^* d\lambda = \int_{\hat{\gamma}} \lambda + \tau(1) - \int_{\tilde{\gamma}} \lambda$$

and the claim follows.

Now let $\sigma: \Sigma \rightarrow M$ be as in the statement. Since Σ has the homotopy type of a graph, there exists a lift $\tilde{\sigma}: \Sigma \rightarrow V$ of σ . By the above observation we get

$$\int_{\Sigma} \sigma^* \omega = \int_{\Sigma} \tilde{\sigma}^* d\lambda = \int_{\partial \Sigma} \tilde{\sigma}^* \lambda = \sum_{j=1}^k \text{Hol}(\gamma_j) \pmod{\mathbb{Z}}. \quad \square$$

An immediate consequence of this lemma is the following.

Proposition 2.8 *For any $\varphi \in \text{Symp}(M, \omega)$, the map (10) descends to a homomorphism*

$$E_{\varphi}: H_1(M, \mathbb{Z}) \rightarrow \mathbb{S}^1, \quad [\gamma] \mapsto E_{\varphi}(\gamma).$$

Moreover the map

$$(11) \quad \text{Symp}^0(M, \omega) \rightarrow \text{Hom}(H_1(M, \mathbb{Z}), \mathbb{S}^1), \quad \varphi \mapsto E_{\varphi}$$

is a homomorphism.

Proof Given smooth loops γ_1 and γ_2 representing the same class in $H_1(M, \mathbb{Z})$, there is an oriented surface Σ with two oriented boundary components S_1 and S_2 and a smooth map $\sigma: \Sigma \rightarrow M$ such that $\sigma|_{S_1} = \gamma_1$ and $\sigma|_{S_2} = \gamma_2^{-1}$. We get

$$\text{Hol}(\gamma_1) + \text{Hol}(\gamma_2^{-1}) = \text{Hol}(\gamma_1) - \text{Hol}(\gamma_2) = \int_{\Sigma} \sigma^* \omega$$

and therefore

$$E_{\varphi}(\gamma_1) - E_{\varphi}(\gamma_2) = \int_{\Sigma} (\varphi \circ \sigma)^* \omega - \int_{\Sigma} \sigma^* \omega = \int_{\Sigma} \sigma^* (\varphi^* \omega) - \int_{\Sigma} \sigma^* \omega = 0.$$

Now given $\varphi_1, \varphi_2 \in \text{Symp}^0(M, \omega)$ and $[\gamma] \in H_1(M, \mathbb{Z})$, we have $[\varphi_2 \circ \gamma] = [\gamma]$ in $H_1(M, \mathbb{Z})$ and

$$E_{\varphi_1 \circ \varphi_2}(\gamma) = E_{\varphi_1}(\varphi_2 \circ \gamma) + E_{\varphi_2}(\gamma) = E_{\varphi_1}(\gamma) + E_{\varphi_2}(\gamma). \quad \square$$

Notice that $\mathcal{H}$ is precisely the kernel of the homomorphism (11). Our next objective is to relate the homomorphism (11) to the flux homomorphism

$$(12) \quad \text{Flux}: \text{Symp}^0(M, \omega) \rightarrow H_{dR}^1(M, \mathbb{R}) / \Gamma_{\omega}.$$

Recall that if $(\varphi_t)_{t \in [0,1]}$ is a symplectic isotopy of a closed symplectic manifold (M, ω) , not necessarily integral, generated by the time dependent symplectic vector field X_t , the flux of $(\varphi_t)_{t \in [0,1]}$ is defined by

$$\text{Flux}(\varphi_t) = \int_0^1 [\iota_{X_t} \omega] dt \in H_{dR}^1(M, \mathbb{R}).$$

It turns out that $\text{Flux}(\varphi_t)$ stays invariant under any homotopy of the path $(\varphi_t)_{t \in [0,1]}$ in $\text{Symp}^0(M, \omega)$, which fixes the end points Id and φ_1 and consequently defines a homomorphism

$$\text{Flux}: \widetilde{\text{Symp}}^0(M, \omega) \rightarrow H_{dR}^1(M, \mathbb{R})$$

where $\widetilde{\text{Symp}}^0(M, \omega)$ is the universal cover of $\text{Symp}^0(M, \omega)$, seen as the space of homotopy classes, in the above sense, of paths in $\text{Symp}^0(M, \omega)$ starting from the identity. Viewing $\pi_1(\text{Symp}^0(M, \omega), \text{Id})$ as a subgroup of $\widetilde{\text{Symp}}^0(M, \omega)$, one defines the *flux group*

$$\Gamma_\omega := \text{Flux}(\pi_1(\text{Symp}^0(M, \omega), \text{Id})) \subseteq H_{dR}^1(M, \mathbb{R})$$

and obtains the homomorphism (12). It turns out that $\text{Flux}(\varphi_t)$ vanishes if and only if (φ_t) is homotopic to a Hamiltonian isotopy. Hence the kernel of (12) is precisely the group $\text{Ham}(M, \omega)$ of Hamiltonian diffeomorphisms.

The following observation relates Flux to the homomorphism (11) if ω is integral.

Lemma 2.9 *Let (φ_t) be a symplectic isotopy starting at Id . Then for any smooth loop $\gamma: \mathbb{S}^1 \rightarrow M$, we have*

$$\int_\gamma \text{Flux}(\varphi_t) = E_{\varphi_1}(\gamma) \mod \mathbb{Z}.$$

Proof Given γ and φ_t as above, one defines the smooth map $\sigma: [0, 1] \times \mathbb{S}^1 \rightarrow M$ by $\sigma(s, t) = \varphi_t(\gamma(s))$. An easy computation shows that

$$(13) \quad \int_{[\gamma]} \text{Flux}(\varphi_t) = \int_{[0,1] \times \mathbb{S}^1} \sigma^* \omega$$

and from Lemma 2.7 it follows that

$$\int_{[0,1] \times \mathbb{S}^1} \sigma^* \omega = \text{Hol}(\gamma^{-1}) + \text{Hol}(\varphi_1 \circ \gamma) = E_{\varphi_1}(\gamma) \mod \mathbb{Z}. \quad \square$$

We have the following characterization of the mapping class group $\pi_0(\mathcal{H})$.

Theorem 2.10 *We have the following canonical isomorphisms of groups:*

$$(14) \quad \pi_0(\mathcal{H}) = \mathcal{H} / \text{Ham}(M, \omega) \cong H_{dR}^1(M, \mathbb{Z}) / \Gamma_\omega.$$

Proof The statement follows immediately from Lemma 2.9 and the fact that Γ_ω is contained in $H_{dR}^1(M, \mathbb{Z})$ as ω is integral. $\square$

3 $H_{dR}^1(M, \mathbb{Z})$ in $\pi_0(\text{Cont}(V, \xi))$ and in $\pi_0(\text{Diff}(V))$

The aim of this section is to provide a class of examples of prequantizations $(V, \lambda) \rightarrow (M, \omega)$ such that $\pi_0(\mathcal{H})$ is an infinite subgroup of $\pi_0(\text{SCont}(V, \lambda))$ and the homomorphism

$$(15) \quad \pi_0(\mathcal{H}) \rightarrow \pi_0(\text{Cont}(V, \xi))$$

is injective, while the homomorphism

$$(16) \quad \pi_0(\mathcal{H}) \rightarrow \pi_0(\text{Diff}(V))$$

is trivial. Here ξ stands for the kernel of λ , $\text{Cont}(V, \xi)$ is the group of coorientation (given by λ) preserving contactomorphisms of (V, ξ) and $\text{Diff}(V)$ is the group of orientation (determined by λ) preserving diffeomorphisms of V .

We begin with an explicit description of the inverse of the isomorphism (14). Let α be a closed integral 1-form on M given by $\alpha = df$ for some $f \in C^\infty(M, \mathbb{S}^1)$. We define the symplectic gradient X_α of α via $-\alpha = \iota_{X_\alpha} \omega$ and we let $(\bar{\varphi}_t)$ denote its flow. Then the time-one-map $\bar{\varphi}_1$ belongs to $\mathcal{H}$ with a lift

$$(17) \quad L_\alpha: V \rightarrow V, \quad L_\alpha = \phi_{f \circ \pi}^X \circ \varphi_1$$

in $\text{SCont}(V, \lambda)$, where φ_1 is the time-one-map of the flow (φ_t) of the horizontal lift X_h of X_α and the map $\phi_{f \circ \pi}^X$ is given by (8). Indeed as we saw in the proof of Theorem 2.2,

$$(18) \quad (\varphi_t)^* \lambda = \lambda - t \pi^* \alpha.$$

Together with (9) we see that L_α preserves λ . We also note that the maps $\phi_{f \circ \pi}^X$ and φ_t commute for all $t \in \mathbb{R}$. Indeed as in the proof of Theorem 2.2, the vector fields X and X_h commute and the map $f \circ \pi$ is invariant under the flow of X_h . We have the following observation on the action of L_α on loops in V .

Lemma 3.1 *Let $\delta: \mathbb{S}^1 \rightarrow V$ be a loop. Then $L_\alpha \circ \delta$ is freely homotopic to the loop $\delta * \gamma_V^k$ where $\gamma_V: \mathbb{S}^1 \rightarrow V$ denotes the parameterized Reeb orbit starting at the point $\delta(0)$ and $k = \int_\delta \pi^* \alpha$.*

Proof Note that the loop $L_\alpha \circ \delta$ is freely homotopic to the loop $\phi_{f \circ \pi}^X \circ \delta$ via the isotopy (φ_t) so it is enough to show that $\delta * \gamma_V^k$ is freely homotopic to the loop $\phi_{f \circ \pi}^X \circ \delta$. To this end we consider the pull-back bundle $\bar{\delta}^* V \rightarrow \mathbb{S}^1$, where $\bar{\delta} = \pi \circ \delta$ and trivialize it via the section δ . After identifying the total space with $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$, where the bundle projection π corresponds to the projection to the first circle, the loop $\delta * \gamma_V^k$ reads $(\mathbb{S}^1 \times \{0\}) * (\{0\} \times \mathbb{S}^1)^k$ and the loop $\phi_{f \circ \pi}^X \circ \delta$ reads $t \mapsto (t, f(t))$, where $f: \mathbb{S}^1 \rightarrow \mathbb{S}^1$ is of degree k . It is clear that these loops are homotopic in $\mathbb{T}^2$. $\square$

As a consequence of the lemma we get the following class of prequantizations where the map (15) is injective due to topological reasons; more precisely, due to the fact that (16) is injective.

Proposition 3.2 *Let (M, ω) be a closed integral symplectic manifold such that*

- (1) $[\omega]$ vanishes on $\pi_2(M)$;
- (2) any abelian subgroup of $\pi_1(M)$ has rank at most one.

Let $(V, \lambda) \rightarrow (M, \omega)$ be a prequantization. Then for any closed integral 1-form α on M , L_α is isotopic to Id in $\text{Diff}(V)$ if and only if $[\alpha]$ is trivial in $H_{dR}^1(M, \mathbb{Z})$.

Lemma 3.3 *Let H be an abelian group of rank at most one. Then every $\mathbb{Z}$ -central extension of H is trivial. More precisely, if G is a group with a central subgroup C isomorphic to $\mathbb{Z}$ and G/C is isomorphic to H then G is isomorphic to $H \times \mathbb{Z}$. In particular, G is abelian.*

Proof The commutator subgroup $[G, G]$ is contained in $C = \mathbb{Z}$. The map $G \times G \rightarrow \mathbb{Z}$ that sends a couple (a, b) to the commutator $aba^{-1}b^{-1}$ factors to a skew-symmetric bilinear map $B: H \times H \rightarrow \mathbb{Z}$. Since H has rank at most one, it follows that B is identically zero. $\square$

Proof of Proposition 3.2 The homotopy exact sequence associated to the bundle $V \rightarrow M$ reads

$$(19) \quad \cdots \rightarrow \pi_2(V) \rightarrow \pi_2(M) \rightarrow \pi_1(\mathbb{S}^1) \rightarrow \pi_1(V) \rightarrow \pi_1(M) \rightarrow 0$$

where the connecting homomorphism is given by integrating ω over spheres in M . By the first assumption, the fiber γ_V is of infinite order in $\pi_1(V)$. We note also that the image of $\pi_1(\mathbb{S}^1)$ is always central in $\pi_1(V)$, that is γ_V commutes with every element of $\pi_1(V)$. In fact, given any loop η in V (having the same base point with γ_V) and its projection $\bar{\eta}$ in M , γ_V and η can be seen as loops in the total space of the pull-back bundle $\bar{\eta}^*V \rightarrow \mathbb{S}^1$, which is diffeomorphic to $\mathbb{T}^2$ and therefore has abelian fundamental group.

Now if L_α is isotopic to Id and δ is a loop in V , then by Lemma 3.1, δ is freely homotopic to $\delta * \gamma_V^k$. Then there is a loop β in V such that $\beta * \delta * \beta^{-1} = \delta * \gamma_V^k$ in $\pi_1(V)$. We write $\bar{\delta} = \pi \circ \delta$ and $\bar{\beta} = \pi \circ \beta$. Then $\bar{\delta}$ commutes with $\bar{\beta}$ in $\pi_1(M)$ and by assumption $\bar{\delta}$ and $\bar{\beta}$ are contained in a subgroup H of rank at most one. Then, by Lemma 3.3, the subgroup $\pi_*^{-1}(H)$ is abelian. As δ and β are contained in $\pi_*^{-1}(H)$, we conclude that γ_V^k is contractible and therefore $k = \int_{\bar{\delta}} \alpha = 0$. Since δ is arbitrary we get $[\alpha] = 0$. $\square$

Remark 3.4 The above statement shows that the flux group Γ_ω is trivial in this case and $\pi_0(\mathcal{H})$, being isomorphic to $H_{dR}^1(M, \mathbb{Z})$, injects into $\pi_0(\text{Diff}(V))$.

Example 3.5 A typical example for the above proposition is an oriented surface of genus at least two.

Our next objective is to use Proposition 3.2 as a springboard to describe prequantizations for which the homomorphism (15) is injective for contact topological reasons. We want to replace the role of the Reeb orbit with the loop induced by the Reeb flow in the bundle $\text{CSp}(\xi) \rightarrow V$ of conformally symplectic frames of ξ .

The bundle $\text{CSp}(\xi) \rightarrow V$ is a principal $\text{CSp}(2n, \mathbb{R})$ -bundle, where $\text{CSp}(2n, \mathbb{R})$ denotes the group of conformally symplectic matrices. As the symplectic group $\text{Sp}(2n, \mathbb{R})$ is a deformation retract of $\text{CSp}(2n, \mathbb{R})$, it is natural to consider the isomorphism, called the *Maslov index*, given by

$$\text{Index: } \pi_1(\text{CSp}(2n, \mathbb{R})) = \pi_1(\text{Sp}(2n, \mathbb{R})) \rightarrow \pi_1(\text{U}(n)) \rightarrow \pi_1(\mathbb{S}^1) \cong \mathbb{Z}$$

where the first arrow is given by the polar decomposition and the second arrow is induced by the determinant map. We have the homotopy exact sequence

$$(20) \quad \cdots \rightarrow \pi_2(V) \rightarrow \pi_1(\mathrm{CSp}(2n, \mathbb{R})) \rightarrow \pi_1(\mathrm{CSp}(\xi)) \rightarrow \pi_1(V) \rightarrow 0$$

and it is easy to see that the connecting homomorphism is given by integrating the first Chern class $c_1(\xi)$ over spheres in V . Here we view ξ as a complex vector bundle of rank n , by choosing a complex structure on ξ compatible with $d\lambda$. We note that the space of such complex structures is nonempty and contractible. In the same vein, $c_1(TM)$ is well defined and we have $c_1(\xi) = \pi^*c_1(TM)$. Combining this with (20), we conclude that the homomorphism

$$(21) \quad \pi_1(\mathrm{CSp}(2n, \mathbb{R})) \rightarrow \pi_1(\mathrm{CSp}(\xi))$$

is injective if and only if $c_1(TM)$ vanishes on every sphere of vanishing symplectic area in M . Note that this is equivalent to assuming that (M, ω) is *monotone on $\pi_2(M)$* , that is, there is some $\kappa \in \mathbb{R}$ (in fact $\kappa \in \mathbb{Q}$), which we call the *monotonicity constant*, such that $c_1(TM) = \kappa\omega$ on $\pi_2(M)$. We also remark that the image of $\pi_1(\mathrm{CSp}(2n, \mathbb{R}))$ in $\pi_1(\mathrm{CSp}(\xi))$ is always central. This follows from the argument given in the proof of [Proposition 3.2](#) for the bundle $V \rightarrow M$ and the fact that $\pi_1(\mathrm{CSp}(2n, \mathbb{R}))$ is abelian.

Our aim now is to study the loop in $\mathrm{CSp}(\xi)$ that is associated to the Reeb flow and to determine sufficient conditions such that it is of infinite order in $\pi_1(\mathrm{CSp}(\xi))$. Let $p \in V$ and $\mathcal{F}_p$ be a symplectic frame of ξ_p . We define the *Reeb frame loop* in the frame bundle by

$$(22) \quad \Gamma_V: \mathbb{S}^1 \rightarrow \mathrm{CSp}(\xi), \quad t \mapsto (\phi_t^X)_* \mathcal{F}_p,$$

which defines an element in $\pi_1(\mathrm{CSp}(\xi), \mathcal{F}_p)$. Next we assume that the Reeb orbit γ_V is contractible. Let $j: \mathbb{D} \rightarrow V$ be a capping disc, that is $j = \gamma_V$ on $\partial\mathbb{D}$. Let $\mathcal{F}$ be a symplectic framing of $j^*\xi$. Writing Γ_V with respect to the framing $\mathcal{F}|_{\partial\mathbb{D}}$, we get the loop

$$\Gamma_V^j: \mathbb{S}^1 \rightarrow \mathrm{CSp}(2n, \mathbb{R})$$

to which we can associate its Maslov index.

Lemma 3.6 *Let $\bar{j}: S^2 := \mathbb{D}/\partial\mathbb{D} \rightarrow M$ be the sphere given by $\bar{j} = \pi \circ j$. Then we have*

$$\mathrm{Index}(\Gamma_V^j) = -\langle c_1(\bar{j}^*TM), S^2 \rangle = -\langle c_1(TM), \bar{j}(S^2) \rangle.$$

In particular, $\mathrm{Index}(\Gamma_V^j)$ does not depend on $\mathcal{F}$. Moreover, $\mathrm{Index}(\Gamma_V^j)$ is independent of the capping disc if and only if (M, ω) is monotone on $\pi_2(M)$.

Proof Let $x \in S^2$ denote the image of $\partial\mathbb{D}$ in the quotient. We take a collar neighborhood U of $\partial\mathbb{D}$ in $\mathbb{D}$ and let $\bar{U}$ be the corresponding disc neighborhood of $x \in S^2$. We pick a symplectic framing $\bar{\mathcal{G}}$ of $\bar{j}^*TM|_{\bar{U}}$, say extending $\pi_*\mathcal{F}_p$, and lift it to a Reeb invariant framing $\mathcal{G}$ of $j^*\xi|_U$ extending $\mathcal{F}_p$. Note that this is possible since $\bar{U}$ is contractible. On the other hand, $\mathcal{F}$ descends to a symplectic framing $\bar{\mathcal{F}}$ of

$\bar{j}^*TM|_{S^2 \setminus \{x\}}$. Viewing $c_1(\bar{j}^*TM)$ as the homotopy class of the clutching function that is used to glue the trivial bundles $(\bar{j}^*TM|_{S^2 \setminus \bar{U}}, \bar{\mathcal{F}})$ and $(\bar{j}^*TM|_{\bar{U}}, \bar{\mathcal{G}})$ along $\partial\bar{U}$, we get

$$\langle c_1(\bar{j}^*TM), S^2 \rangle = \text{Index}(\bar{\mathcal{F}}|_{\partial\bar{U}}) = -\text{Index}(\bar{\mathcal{G}}|_{\partial\bar{U}})$$

where we view $\bar{\mathcal{F}}|_{\partial\bar{U}}$ (resp. $\bar{\mathcal{G}}|_{\partial\bar{U}}$) as a loop in $\text{CSp}(2n, \mathbb{R})$ using the framing $\bar{\mathcal{G}}|_{\partial\bar{U}}$ (resp. $\bar{\mathcal{F}}|_{\partial\bar{U}}$). Using the symplectic isomorphism π_* , we get

$$\langle c_1(\bar{j}^*TM), S^2 \rangle = -\text{Index}(\mathcal{G}|_\eta)$$

where η denotes the preimage of $\partial\bar{U}$ in $\mathbb{D}$ and the loop $\mathcal{G}|_\eta$ is obtained via the framing $\mathcal{F}$. Since $\mathcal{G}|_\eta$ is homotopic to the loop $\mathcal{G}|_{\partial\mathbb{D}}$, which is defined via the framing $\mathcal{F}$, and since Γ_V^j reads as the constant loop with respect to the framing $\mathcal{G}|_{\partial\mathbb{D}}$, we get

$$\langle c_1(\bar{j}^*TM), S^2 \rangle = -\text{Index}(\mathcal{G}|_{\partial\mathbb{D}}) = -\text{Index}(\Gamma_V^j),$$

where the left hand side is clearly independent of the framing $\mathcal{F}$.

Recall that (M, ω) is monotone on $\pi_2(M)$ if and only if $c_1(\xi)$ vanishes on $\pi_2(V)$. Now given two capping discs $j_1, j_2: \mathbb{D} \rightarrow V$, we glue them together, say after reversing the orientation of the second, to get a sphere $j: S^2 \rightarrow V$. We have

$$\text{Index}(\Gamma_V^{j_1}) - \text{Index}(\Gamma_V^{j_2}) = \langle c_1(j^*\xi), S^2 \rangle.$$

Hence $\text{Index}(\Gamma_V^{j_1}) = \text{Index}(\Gamma_V^{j_2})$ if $c_1(\xi)$ vanishes on $\pi_2(V)$. On the other hand, given a sphere S in V , one can choose some capping disc $j_1: \mathbb{D} \rightarrow V$ with $j_1(0) \in S$ and define the capping disc j_2 to be the concatenation of j_1 and S . Repeating the above construction for these two capping discs leads to a sphere j in the homotopy class of S . Then the above formula says that $c_1(\xi)$ vanishes on S if the index is independent of the capping discs. $\square$

Remark 3.7 The above construction of the Maslov index can also be carried out in the case where γ_V is null homologous. In this case the capping disc is replaced by a capping surface $j: \Sigma \rightarrow V$ and the relative index is given by evaluating $c_1(TM)$ over the closed surface $\bar{j}: \bar{\Sigma} := \Sigma/\partial\Sigma \rightarrow M$.

An immediate consequence of the above lemma is the following.

Proposition 3.8 *Let $(V, \lambda) \rightarrow (M, \omega)$ be a prequantization. Suppose that either $[\omega]$ vanishes on $\pi_2(M)$ or (M, ω) is monotone on $\pi_2(M)$ with nonvanishing monotonicity constant. Then the Reeb frame loop Γ_V is of infinite order in $\pi_1(\text{CSp}(\xi))$.*

Proof Notice that if $[\omega]$ is trivial on $\pi_2(M)$ then the Reeb orbit is of infinite order in $\pi_1(V)$. Since Γ_V projects to γ_V the statement follows.

Now assume that there is sphere in M with unit symplectic area. Then γ_V is contractible. Given a capping disc $j: \mathbb{D} \rightarrow V$ with $p = j(1)$, with a framing $\mathcal{F}$ of $j^*\xi$ extending $\mathcal{F}_p$, the loop Γ_V based at

$\mathcal{F}_p$ is homotopic to a loop in $\mathrm{CSp}(\xi_p)$ based at $\mathcal{F}_p$, whose homotopy class is given by $\mathrm{Index}(\Gamma_V^j)$ after $(\mathrm{CSp}(\xi_p), \mathcal{F}_p)$ is identified by $(\mathrm{CSp}(2n, \mathbb{R}), \mathrm{Id})$. By monotonicity, the homomorphism (21) is injective. Hence Γ_V is of infinite order in $\pi_1(\mathrm{CSp}(\xi), \mathcal{F}_p)$ if and only if $\mathrm{Index}(\Gamma_V^j)$ is nonzero.

On the other hand,

$$1 = \int_{\gamma_V} \lambda = \int_{\mathbb{D}} j^* d\lambda = \int_{S^2} \bar{j}^* \omega.$$

Since the monotonicity constant is assumed to be nonzero, $\mathrm{Index}(\Gamma_V^j)$ does not vanish.

Finally, if γ_V is torsion, say of order d , in $\pi_1(V)$, then we consider the loop Γ_V^d in $\pi_1(\mathrm{CSp}(\xi))$ and apply the Maslov index construction to Γ_V^d . The argument above applies word by word to Γ_V^d and we conclude that Γ_V^d and consequently Γ_V is of infinite order. $\square$

Given $p \in V$ and a symplectic frame $\mathcal{F}_p$ of ξ_p , we define a map

$$\mathrm{Cont}(V, \xi) \rightarrow \mathrm{CSp}(\xi), \quad \varphi \mapsto \varphi_* \mathcal{F}_p$$

and get the induced homomorphisms $\pi_1(\mathrm{Cont}(V, \xi), \mathrm{Id}) \rightarrow \pi_1(\mathrm{CSp}(\xi), \mathcal{F}_p)$ which maps the homotopy class of the Reeb loop $(\phi_t^X)_{t \in \mathbb{S}^1}$ to the homotopy class of Γ_V . Hence the Reeb loop is of infinite order in $\pi_1(\mathrm{Cont}(V, \xi), \mathrm{Id})$ if Γ_V is of infinite order in $\pi_1(\mathrm{CSp}(\xi), \mathcal{F}_p)$. Hence we get the following corollary.

Corollary 3.9 *Let $(V, \lambda) \rightarrow (M, \omega)$ be a prequantization. Suppose that either $[\omega]$ vanishes on $\pi_2(M)$ or (M, ω) is monotone on $\pi_2(M)$ with nonvanishing monotonicity constant. Then the Reeb loop $(\phi_t^X)_{t \in \mathbb{S}^1}$ is of infinite order in $\pi_1(\mathrm{Cont}(V, \xi))$.*

After establishing the setting in which Γ_V is suitable to take the role of γ_V in the proof of Proposition 3.2, we are ready to prove the lemma that is analogous to Lemma 3.1.

Lemma 3.10 *Let $\hat{\delta}: \mathbb{S}^1 \rightarrow \mathrm{CSp}(\xi)$ be a loop with its projection $\delta: \mathbb{S}^1 \rightarrow V$. Then $(L_\alpha)_* \circ \hat{\delta}$ is freely homotopic to the loop $\hat{\delta} * \Gamma_V^k$ where $\Gamma_V: \mathbb{S}^1 \rightarrow \mathrm{CSp}(\xi)$ is the Reeb frame loop based at $\hat{\delta}(0)$ and $k = \int_{\delta} \pi^* \alpha$.*

Proof Recall that $L_\alpha = \varphi_1 \circ \phi_{f \circ \pi}^X$ where $\alpha = df$ for some $f: M \rightarrow \mathbb{S}^1$ and φ_1 is the time-one-map of the flow (φ_t) of the horizontal vector field X_h . From (18) it follows that

$$(\varphi_t)_*: (\xi_0, d\lambda) \rightarrow (\xi_t, d\lambda), \quad t \in [0, 1]$$

is a smooth path of isomorphisms of rotating contact structures

$$\xi_t := \ker(\lambda + (t - 1)\pi^* \alpha), \quad t \in [0, 1],$$

seen as symplectic vector bundles. Note that $\xi_1 = \ker \lambda = \xi$ and by (9) we have that

$$(\phi_{f \circ \pi}^X)_*: (\xi_1, d\lambda) \rightarrow (\xi_0, d\lambda)$$

is also an isomorphism. Notice that the vector field X is the Reeb vector field for every contact form $\lambda + (t - 1)\pi^*\alpha$ and by projecting along X , we get a smooth path

$$P_t: (\xi_t, d\lambda) \rightarrow (\xi_1, d\lambda)$$

of isomorphisms of symplectic vector bundles. Observe that P_1 is the identity map. Now we define a path of loops

$$\sigma: \mathbb{S}^1 \times [0, 1] \rightarrow \mathrm{CSp}(\xi), \quad (s, t) \mapsto P_t \circ (\varphi_t)_* \circ (\phi_{f \circ \pi}^X)_* \circ \hat{\delta}(s),$$

which defines a free homotopy between $\sigma(\cdot, 0) = P_0 \circ (\phi_{f \circ \pi}^X)_* \circ \hat{\delta}$ and $\sigma(\cdot, 1) = (L_\alpha)_* \circ \hat{\delta}$.

We claim that $P_0 \circ (\phi_{f \circ \pi}^X)_* \circ \hat{\delta}$ is freely homotopic to $\hat{\delta} * \Gamma_V^k$ in $\mathrm{CSp}(\xi)$. To see this, we write $\bar{\delta} := \pi \circ \delta$ and consider the pull-back bundle $\bar{\delta}^*V \rightarrow \mathbb{S}^1$. Let $\Delta: \bar{\delta}^*V \rightarrow V$ be the bundle map covering $\bar{\delta}$. Next we take a symplectic trivialization of $\bar{\delta}^*TM$ and lift it to an $\mathbb{S}^1$ -invariant trivialization of the pull-back bundle $\Delta^*\xi \rightarrow \bar{\delta}^*V$. Pairing this trivialization with the trivialization of $\bar{\delta}^*V \rightarrow \mathbb{S}^1$ described in the proof of [Lemma 3.1](#), we identify $\mathrm{CSp}(\Delta^*\xi)$ with $\mathbb{T}^2 \times \mathrm{CSp}(2n, \mathbb{R})$ in such a way that $\hat{\delta}$ reads as $(\mathbb{S}^1 \times \{0\})$ in $\mathbb{T}^2$ paired with a loop η in $\mathrm{CSp}(2n, \mathbb{R})$ based at Id , and Γ_V^k reads as $(\{0\} \times \mathbb{S}^1)^k$ in $\mathbb{T}^2$ paired with the constant loop Id in $\mathrm{CSp}(2n, \mathbb{R})$ since $\Delta^*\xi$ is trivialized in an $\mathbb{S}^1$ -invariant fashion. Hence $\hat{\delta} * \Gamma_V^k$ reads as $(\mathbb{S}^1 \times \{0\}) * (\{0\} \times \mathbb{S}^1)^k$ paired with η . On the other hand since P_0 is given by the projection along X , for any $s \in \mathbb{S}^1$ and $u \in \hat{\delta}(s)$, we have

$$P_0 \circ (\phi_{f \circ \pi}^X)_* u = P_0(d(\phi_{f(\bar{\delta}(s))}^X)_{\delta(s)}[u] + \alpha_{\bar{\delta}(s)}(\pi_* u)X) = d(\phi_{f(\bar{\delta}(s))}^X)_{\delta(s)}[u].$$

Again since $\Delta^*\xi$ is trivialized in an $\mathbb{S}^1$ -invariant fashion, $P_0 \circ (\phi_{f \circ \pi}^X)_* \circ \hat{\delta}$ reads

$$s \mapsto ((s, f(s)), \eta(s)) \in \mathbb{T}^2 \times \mathrm{CSp}(2n, \mathbb{R}),$$

where $f: \mathbb{S}^1 \rightarrow \mathbb{S}^1$ is of degree k . Therefore, $P_0 \circ (\phi_{f \circ \pi}^X)_* \circ \hat{\delta}$ is freely homotopic to $\hat{\delta} * \Gamma_V^k$ in $\mathbb{T}^2 \times \mathrm{CSp}(2n, \mathbb{R})$. $\square$

Now we are ready to state our first main result.

Theorem 3.11 *Let (M, ω) be a closed integral symplectic manifold such that either $[\omega]$ vanishes on $\pi_2(M)$ or (M, ω) is monotone on $\pi_2(M)$ with a nonvanishing monotonicity constant. Suppose that any abelian subgroup of $\pi_1(M)$ has rank at most one. If $(V, \lambda) \rightarrow (M, \omega)$ is a prequantization, then for any closed integral 1-form α on M , L_α is isotopic to Id in $\mathrm{Cont}(V, \xi)$ if and only if $[\alpha]$ is trivial in $H_{dR}^1(M, \mathbb{Z})$.*

Proof If $[\omega]$ vanishes on $\pi_2(M)$ then this follows from [Proposition 3.2](#). Now, suppose $[\omega]$ does not vanish on $\pi_2(M)$ and (M, ω) is monotone on $\pi_2(M)$ with a nonvanishing monotonicity constant. If L_α is isotopic to Id in $\mathrm{Cont}(V, \xi)$ and $\hat{\delta}$ is a loop in $\mathrm{CSp}(\xi)$, then by [Lemma 3.10](#) $\hat{\delta}$ is freely homotopic to $\hat{\delta} * \Gamma_V^k$. So there is a loop $\hat{\beta}$ in $\mathrm{CSp}(\xi)$ such that $\hat{\beta} * \hat{\delta} * \hat{\beta}^{-1} = \hat{\delta} * \Gamma_V^k$ in $\pi_1(\mathrm{CSp}(\xi))$. Let δ and β be the projections of $\hat{\delta}$ and $\hat{\beta}$ to V respectively. Then $\delta^{-1}\beta\delta\beta^{-1} = \gamma_V^k$ where γ_V^k is the projection of Γ_V^k . But, we can assume γ_V^k is contractible by iterating L_α (or equivalently, by considering a suitable multiple

$n\alpha$ of α). In this case, δ commutes with β in $\pi_1(V)$. It follows that δ and β are contained in an abelian subgroup H of rank at most one since the $\mathbb{S}^1$ -fiber of V is torsion in $\pi_1(V)$ (by the assumption that $[\omega]$ does not vanish on $\pi_2(M)$) and every abelian subgroup of $\pi_1(M)$ has rank at most one. It follows, by Lemma 3.3, that the subgroup $\pi_*^{-1}(H)$ is an abelian subgroup of $\pi_1(\mathrm{CSp}(\xi))$ since it is a $\mathbb{Z}$ -central extension of H (the Reeb frame loop is of infinite order by Proposition 3.8). Hence $\hat{\delta}$ and $\hat{\beta}$ are contained in the abelian subgroup $\pi_*^{-1}(H)$ and therefore Γ_V^k is contractible. But by Proposition 3.8 this is possible only if $k = \int_{\hat{\delta}} \alpha = 0$. Since δ is arbitrary we get $\alpha = 0$ in $H_{dR}^1(M, \mathbb{Z})$. $\square$

Now let (W, ω_W) be a simply connected integral symplectic manifold. Consider the product symplectic manifold $M := W \times \Sigma_g$, $\omega := \omega_W \oplus \omega_0$, where ω_0 is any integral symplectic form on the surface Σ_g of genus $g \geq 2$ and let $\pi: (V, \lambda) \rightarrow (M, \omega)$ be a prequantization. Then V restricted to $W \times \{\mathrm{pt}\}$, denoted by $V|_W$, defines a prequantization of (W, ω_W) . Note that $\pi_1(M) = \pi_1(\Sigma_g)$ and Γ_ω is trivial (see, for example, [5, Theorem 3]). Then by Theorem 2.10, the kernel of the homomorphism (which is surjective in this case)

$$\pi_0(\mathrm{SCont}(V, \lambda)) \rightarrow \pi_0(\mathrm{Symp}(M, \omega))$$

is given by $H_{dR}^1(M, \mathbb{Z}) = H_{dR}^1(\Sigma_g, \mathbb{Z}) \cong \mathbb{Z}^{2g}$. So we get the homomorphism

$$\mathcal{J}: H_{dR}^1(\Sigma_g, \mathbb{Z}) \rightarrow \pi_0(\mathrm{Cont}(V, \xi))$$

and we have the natural homomorphism $\mathcal{J}: \pi_0(\mathrm{Cont}(V, \xi)) \rightarrow \pi_0(\mathrm{Diff}(V))$.

Proposition 3.12 *Suppose that the $\mathbb{S}^1$ -action on $V|_W$, given by the Reeb flow, defines a torsion loop in $\pi_1(\mathrm{Diff}_0(V|_W))$ of order k . Then every mapping class in the subgroup $kH_{dR}^1(\Sigma_g, \mathbb{Z}) \subset \pi_0(\mathrm{SCont}(V, \lambda))$ is smoothly trivial.*

Proof Let $\sigma: \mathbb{S}^1 \times [0, \varepsilon] \rightarrow \Sigma_g$ be a symplectic embedding of a small cylinder and let $f: [0, \varepsilon] \rightarrow \mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ be a smooth function that vanishes near 0 and ε . Now, f induces a function $\tilde{f}: \Sigma_g \rightarrow \mathbb{S}^1$ as follows: extend f to $\mathbb{S}^1 \times [0, \varepsilon]$ in the obvious way, so it induces a function on the image of σ which can be extended to be identically zero outside. The differential form $d\tilde{f}$ represents an integral cohomology class and, conversely, every integral cohomology is represented by such a function (with a suitable choice of σ and the degree of f). Therefore, every element of $H_{dR}^1(M, \mathbb{Z})$ can be represented as the differential of a function $\tilde{f} \circ p$ where $p: M \rightarrow \Sigma_g$ is the projection and $\tilde{f}$ is as above.

Recall that by (17), the lift of $d(\tilde{f} \circ p)$ to $\mathrm{SCont}(V, \lambda)$ is given by $\phi_{(\tilde{f} \circ p) \circ \pi}^X \circ \varphi_1$ where φ_1 is the time-one-map of the flow (φ_t) of the horizontal lift of the symplectic gradient of $d(\tilde{f} \circ p)$. Since φ_1 is isotopic to the identity, one needs only to show that if the degree of f is k (where $f: [0, \varepsilon] \rightarrow \mathbb{S}^1$ is seen as a loop) then $\phi_{(\tilde{f} \circ p) \circ \pi}^X$ is isotopic to the identity.

Let $\sigma: \mathbb{S}^1 \times [0, \varepsilon] \rightarrow \Sigma_g$ be as above and $f: [0, \varepsilon] \rightarrow \mathbb{S}^1$ be such a map of degree k . Denote by C the image of σ . Let $V|_C$ denote the restriction of the $\mathbb{S}^1$ -bundle V to $W \times C$ and $Z = V|_W$ denote the restriction of V to $W \times \{\mathrm{pt}\}$. Then the map $p \circ \pi: V|_C \rightarrow C$ defines a fiber bundle with fiber Z and the

$\mathbb{S}^1$ -action preserves each fiber. Our claim is that $p \circ \pi: V|_C \rightarrow C$ is equivalent, in an $\mathbb{S}^1$ -equivariant way, to $Z \times C \rightarrow C$ where Z is seen as the prequantization of (W, ω_W) .

We first show that the bundle $p \circ \pi: V|_C \rightarrow C$ is trivial, in an $\mathbb{S}^1$ -equivariant way, when restricted to $\mathbb{S}^1 \times \{0\} \subset C$, where C is identified with $\mathbb{S}^1 \times [0, \varepsilon]$. Fix $y_0 \in W$ and suppose, without loss of generality, that the loop $S_{y_0} := \{y_0\} \times \mathbb{S}^1 \times \{0\}$ has zero holonomy. Then for every $y \in W$, the loop $S_y := \{y\} \times \mathbb{S}^1 \times \{0\}$ also has zero holonomy since for some $\varphi \in \text{Ham}(W, \omega_W)$ with $\varphi(y_0) = y$, the Hamiltonian diffeomorphism $\varphi \times \text{Id}$ of M maps S_{y_0} to S_y . Now, horizontal lifts of the loops S_y for $y \in W$ produce an $\mathbb{S}^1$ -equivariant identification of all the fibers $(p \circ \pi)^{-1}(z)$ for $z \in \mathbb{S}^1 \times \{0\}$. Now we can argue in the same fashion and identify all the fibers $(p \circ \pi)^{-1}(z)$, for $z \in \mathbb{S}^1 \times [0, \varepsilon]$, $\mathbb{S}^1$ -equivariantly just by considering the foliation of $\mathbb{S}^1 \times [0, \varepsilon]$ by intervals $\{s\} \times [0, \varepsilon]$ and their associated horizontal lifts.

Now, the map $\phi_{(\tilde{f} \circ p) \circ \pi}^X$ reads

$$\phi_{(\tilde{f} \circ p) \circ \pi}^X: Z \times C \rightarrow Z \times C, \quad (q, (x, y)) \mapsto (f(y) \cdot q, (x, y))$$

where $f(y) \cdot q$ means the action of $f(y) \in \mathbb{S}^1$ on q . So, $\phi_{(\tilde{f} \circ p) \circ \pi}^X$ has the form

$$(q, (x, y)) \mapsto (\phi_y(q), (x, y))$$

where $\phi: [0, \varepsilon] \rightarrow \text{Diff}^0(Z)$ is a loop based at Id . In our case, ϕ is contractible if f has degree k (since, by assumption, the Reeb flow defines a torsion loop of order k in $\pi_1(\text{Diff}^0(Z))$). This shows that $\phi_{(\tilde{f} \circ p) \circ \pi}^X$ is isotopic to the identity. $\square$

Corollary 3.13 *Let (Σ_g, ω_0) be a surface of genus $g \geq 2$ with an integral symplectic form and ω_{FS} be the Fubini–Study form on $\mathbb{C}P^n$, which is normalized so that the standard $\mathbb{C}P^1 \subset \mathbb{C}P^n$ has unit area. Consider a prequantization $(V, \lambda) \rightarrow (\mathbb{C}P^n \times \Sigma_g, \omega_{\text{FS}} \oplus \omega_0)$. Then the homomorphism*

$$\mathcal{J}: H_{dR}^1(\Sigma_g, \mathbb{Z}) \rightarrow \pi_0(\text{Cont}(V, \xi)),$$

is injective. Moreover, $\ker(\mathcal{J})$ contains the image of $\mathcal{J}$ for n odd, and contains $\mathcal{J}(2H_{dR}^1(\Sigma_g, \mathbb{Z}))$ for n even.

Proof The injectivity of $\mathcal{J}$ follows from [Theorem 3.11](#). Note that the prequantization $V|_{\mathbb{C}P^n}$ is the Hopf bundle $\mathbb{S}^{2n+1} \rightarrow \mathbb{C}P^n$ and the Reeb loop is given by the loop $t \mapsto \text{diag}(e^{2\pi i t}, \dots, e^{2\pi i t})$ in $U(n+1)$. This defines a loop in $\text{SO}(2n+2)$ which is contractible if n is odd and of order 2 if n is even. The rest is a direct consequence of [Proposition 3.12](#). $\square$

References

- [1] **P Albers, E Shelukhin, F Zapolsky**, *Spectral invariants for contactomorphisms of prequantization bundles and applications*, lecture video, IAS Symplectic Geometry Seminar (2016) Available at <https://www.ias.edu/video/puias/2016/0218-Zapolsky>

- [2] **A Banyaga**, *The structure of classical diffeomorphism groups*, Math. Appl. 400, Kluwer, Dordrecht (1997) [MR](#) [Zbl](#)
- [3] **S Bates, A Weinstein**, *Lectures on the geometry of quantization*, Berkeley Math. Lect. Notes 8, Amer. Math. Soc., Providence, RI (1997) [MR](#) [Zbl](#)
- [4] **W M Boothby, H C Wang**, *On contact manifolds*, Ann. of Math. 68 (1958) 721–734 [MR](#) [Zbl](#)
- [5] **C Campos-Apanco, A Pedroza**, *Bounded symplectic diffeomorphisms and split flux groups*, Proc. Amer. Math. Soc. 140 (2012) 2883–2892 [MR](#) [Zbl](#)
- [6] **D Cant**, *Contactomorphisms of the sphere without translated points*, J. Symplectic Geom. 22 (2024) 1–10
- [7] **R Casals, V Ginzburg, F Presas**, *Higher Maslov indices*, J. Geom. Phys. 115 (2017) 167–177 [MR](#) [Zbl](#)
- [8] **R Casals, F Presas**, *A remark on the Reeb flow for spheres*, J. Symplectic Geom. 12 (2014) 657–671 [MR](#) [Zbl](#)
- [9] **R Casals, O Spáčil**, *Chern–Weil theory and the group of strict contactomorphisms*, J. Topol. Anal. 8 (2016) 59–87 [MR](#) [Zbl](#)
- [10] **F Ding, H Geiges**, *The diffeotopy group of $S^1 \times S^2$ via contact topology*, Compos. Math. 146 (2010) 1096–1112 [MR](#) [Zbl](#)
- [11] **K Dymara**, *Legendrian knots in overtwisted contact structures on S^3* , Ann. Global Anal. Geom. 19 (2001) 293–305 [MR](#) [Zbl](#)
- [12] **Y Eliashberg, M Fraser**, *Topologically trivial Legendrian knots*, J. Symplectic Geom. 7 (2009) 77–127 [MR](#) [Zbl](#)
- [13] **H Geiges, M Klukas**, *The fundamental group of the space of contact structures on the 3-torus*, Math. Res. Lett. 21 (2014) 1257–1262 [MR](#) [Zbl](#)
- [14] **F Gironella**, *Examples of nontrivial contact mapping classes for overtwisted contact manifolds in all dimensions*, Algebr. Geom. Topol. 19 (2019) 1207–1227 [MR](#) [Zbl](#)
- [15] **F Gironella**, *Examples of contact mapping classes of infinite order in all dimensions*, Math. Res. Lett. 28 (2021) 707–727 [MR](#) [Zbl](#)
- [16] **E Giroux**, *Sur les transformations de contact au-dessus des surfaces*, from “Essays on geometry and related topics, II”, Monogr. Enseign. Math. 38, Enseignement Math., Geneva (2001) 329–350 [MR](#) [Zbl](#)
- [17] **E Giroux, P Massot**, *On the contact mapping class group of Legendrian circle bundles*, Compos. Math. 153 (2017) 294–312 [MR](#) [Zbl](#)
- [18] **R E Gompf**, *Handlebody construction of Stein surfaces*, Ann. of Math. 148 (1998) 619–693 [MR](#) [Zbl](#)
- [19] **S Kobayashi**, *Principal fibre bundles with the 1-dimensional toroidal group*, Tohoku Math. J. 8 (1956) 29–45 [MR](#) [Zbl](#)
- [20] **S Lanzat, F Zapolsky**, *On the contact mapping class group of the contactization of the A_m -Milnor fiber*, Ann. Math. Qué. 42 (2018) 79–94 [MR](#) [Zbl](#)
- [21] **P Massot, K Niederkrüger**, *Examples of non-trivial contact mapping classes in all dimensions*, Int. Math. Res. Not. 2016 (2016) 4784–4806 [MR](#) [Zbl](#)
- [22] **T S Ratiu, R Schmid**, *The differentiable structure of three remarkable diffeomorphism groups*, Math. Z. 177 (1981) 81–100 [MR](#) [Zbl](#)

- [23] **S Sandon**, *A Morse estimate for translated points of contactomorphisms of spheres and projective spaces*, Geom. Dedicata 165 (2013) 95–110 [MR](#) [Zbl](#)
- [24] **J-M Souriau**, *Structure of dynamical systems: a symplectic view of physics*, Progr. Math. 149, Birkhäuser, Boston, MA (1997) [MR](#) [Zbl](#)
- [25] **T Vogel**, *Non-loose unknots, overtwisted discs, and the contact mapping class group of S^3* , Geom. Funct. Anal. 28 (2018) 228–288 [MR](#) [Zbl](#)

*Fakultät für Mathematik, Ruhr-Universität Bochum
Bochum, Germany*

*Department Mathematik/Informatik, Abteilung Mathematik, Universität zu Köln
Köln, Germany*

souheib.allout@rub.de, msaglam@math.uni-koeln.de

Received: 7 March 2024 Revised: 7 May 2024

Guidelines for Authors

Submitting a paper to Algebraic & Geometric Topology

Papers must be submitted using the upload page at the [AGT website](#). You will need to choose a suitable editor from the list of editors' interests and to supply MSC codes.

The normal language used by the journal is English. Articles written in other languages are acceptable, provided your chosen editor is comfortable with the language and you supply an additional English version of the abstract.

Preparing your article for Algebraic & Geometric Topology

At the time of submission you need only supply a PDF file. Once accepted for publication, the paper must be supplied in $\text{\LaTeX}$, preferably using the journal's class file. More information on preparing articles in $\text{\LaTeX}$ for publication in AGT is available on the [AGT website](#).

arXiv papers

If your paper has previously been deposited on the arXiv, we will need its arXiv number at acceptance time. This allows us to deposit the DOI of the published version on the paper's arXiv page.

References

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited at least once in the text. Use of Bib $\text{\TeX}$ is preferred but not required. Any bibliographical citation style may be used, but will be converted to the house style (see a current issue for examples).

Figures

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Fuzzy or sloppily drawn figures will not be accepted. For labeling figure elements consider the [pinlabel](#) $\text{\LaTeX}$ package, but other methods are fine if the result is editable. If you're not sure whether your figures are acceptable, check with production by sending an email to graphics@msp.org.

Proofs

Page proofs will be made available to authors (or to the designated corresponding author) in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ALGEBRAIC & GEOMETRIC TOPOLOGY

Volume 25

Issue 4 (pages 1917–2526)

2025

The zero stability for the one-row colored $s\ell_3$ -Jones polynomial	1917
WATARU YUASA	
Quillen homology of spectral Lie algebras with application to mod p homology of labeled configuration spaces	1945
ADELA YIYU ZHANG	
Coarse Alexander duality for pairs and applications	1999
G CHRISTOPHER HRUSKA, EMILY STARK and HÙNG CÔNG TRẦN	
K -cwaist on complete foliated manifolds	2037
GUANGXIANG SU and XIANGSHENG WANG	
Line bundle twists for unitary bordism are ghosts	2053
THORSTEN HERTL	
The generalized Kauffman–Harary conjecture is true	2067
RHEA PALAK BAKSHI, HUIZHENG GUO, GABRIEL MONTOYA-VEGA, SUJOY MUKHERJEE and JÓZEF H PRZYTICKI	
Rigidity of elliptic genera for nonspin manifolds	2083
MICHAEL WIEMELER	
Shrinking without doing much at all	2099
MICHAEL FREEDMAN and MICHAEL STARBIRD	
Action of the Mazur pattern up to topological concordance	2115
ALEX MANCHESTER	
Kauffman bracket intertwiners and the volume conjecture	2143
ZHIHAO WANG	
Horizontal decompositions, II	2179
PAOLO LISCA and ANDREA PARMA	
On the nonorientable four-ball genus of torus knots	2209
FRASER BINNS, SUNGKYUNG KANG, JONATHAN SIMONE and PAULA TRUÖL	
Generalised Baumslag–Solitar groups and hierarchically hyperbolic groups	2253
JACK O BUTTON	
Geometric and arithmetic properties of Löbell polyhedra	2281
NIKOLAY BOGACHEV and SAMI DOUBA	
Formality of sphere bundles	2297
JIAWEI ZHOU	
A Quillen stability criterion for bounded cohomology	2317
CARLOS DE LA CRUZ MENGUAL and TOBIAS HARTNICK	
T -equivariant motives of flag varieties	2343
CAN YAYLALI	
Small Heegaard genus and $SU(2)$	2369
JOHN A BALDWIN and STEVEN SIVEK	
Harmonic measures and rigidity for surface group actions on the circle	2391
MASANORI ADACHI, YOSHIFUMI MATSUDA and HIRAKU NOZAWA	
Finite groups of untwisted outer automorphisms of RAAGs	2413
COREY BREGMAN, RUTH CHARNEY and KAREN VOGTMANN	
Computations on cobordism groups of projected immersions	2441
ANDRÁS CSÉPAI	
Rank-preserving additions for topological vector bundles, after a construction of Horrocks	2451
MORGAN P OPIE	
Power sum elements in the G_2 skein algebra	2477